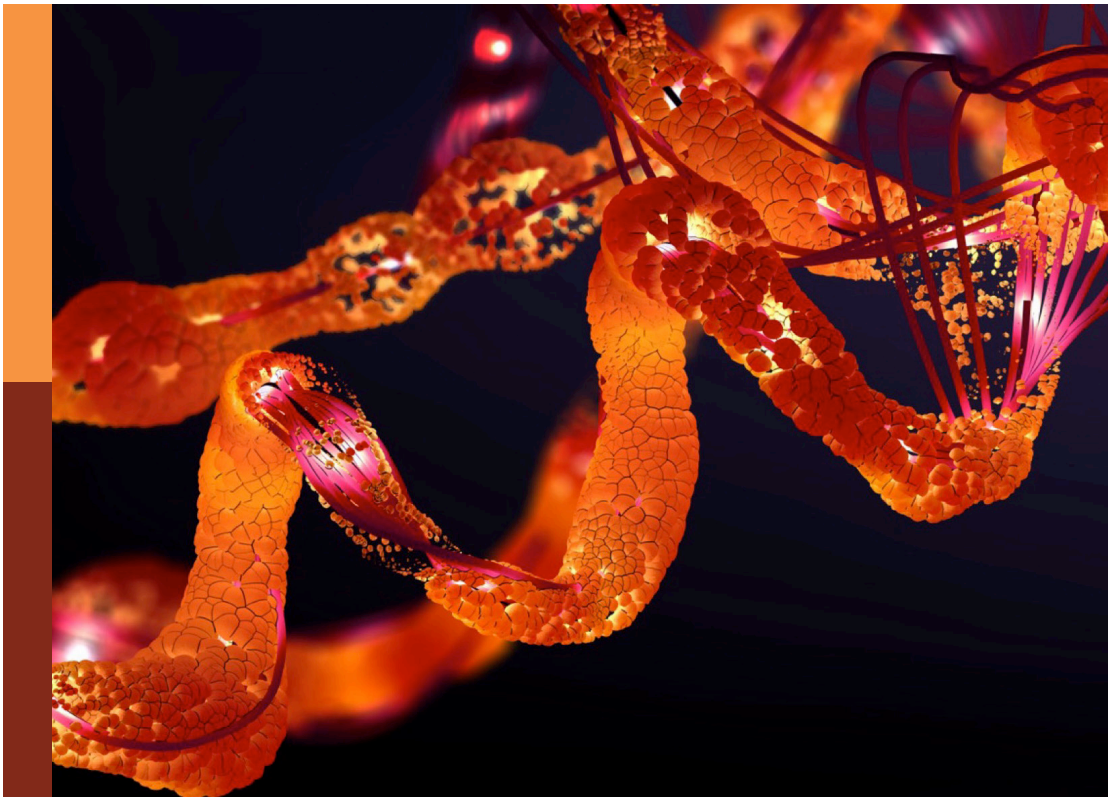
frontiers | Research Topics

Biomolecular solid-state NMR: Methods and applications

Edited by
Amir Goldbourt, Loren B. Andreas and
Józef Romuald Lewandowski

Published in
Frontiers in Molecular Biosciences



ISSN 1664-8714
ISBN 978-2-83251-740-6
DOI 10.3389/978-2-83251-740-6

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Biomolecular solid-state NMR: Methods and applications

Topic editors
Amir Goldbourt — Tel Aviv University, Israel
Loren B. Andreas — Max Planck Institute for Biophysical Chemistry, Germany
Józef Romuald Lewandowski — University of Warwick, United Kingdom

Citation
Goldbourt, A., Andreas, L. B., Lewandowski, J. R., eds. (2023). *Biomolecular solid-state NMR: Methods and applications*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-83251-740-6

Table of contents

frontiers | Frontiers in Molecular Biosciences

TYPE Editorial
PUBLISHED 08 December 2022
DOI 10.3389/fmolb.2022.1082067

OPEN ACCESS

EDITED AND REVIEWED BY
Annalisa Pastore,
King's College London, United Kingdom

*CORRESPONDENCE
Amir Goldbourt,
amirgo@tauex.tau.ac.il

SPECIALTY SECTION
This article was submitted to Structural Biology, a section of the journal Frontiers in Molecular Biosciences

RECEIVED 27 October 2022
ACCEPTED 29 November 2022
PUBLISHED 08 December 2022

CITATION
Goldbourt A, Andreas LB and Lewandowski JR (2022), Editorial: Biomolecular solid-state NMR: Methods and applications.
Front. Mol. Biosci. 9:1082067.
doi: 10.3389/fmolb.2022.1082067

Editorial: Biomolecular solid-state NMR: Methods and applications

Amir Goldbourt[1]*, Loren B. Andreas[2] and Józef R. Lewandowski[3]

[1]School of Chemistry, Tel Aviv University, Tel Aviv, Israel, [2]Department for NMR-Based Structural Biology, Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany, [3]Department of Chemistry, University of Warwick, Coventry, United Kingdom

KEYWORDS

magic angle spinning (MAS), nuclear magnetic resonance (NMR), spectroscopy, protein assemblies, protein dynamics, carbohydrates, dipolar recoupling methods, amyloid fibrils, RNA, membrane proteins

Editorial on the Research Topic
Biomolecular solid- state NMR: Methods and applications

The field of biomolecular solid-state nuclear magnetic resonance (ssNMR) has matured in the last two decades allowing the structural and dynamical characterization of highly complex systems down to atomic resolution. This special issue shows a collection of original studies as well as reviews from leading experts in the field that both advance the methodology, and also span many of the topics where ssNMR makes an impact in structural and functional biology.

A real strength of ssNMR is an ability to structurally characterize not only samples with long range molecular order such as crystals but also systems with short range order like fibrils and membrane proteins or even disordered systems. The applicability of ssNMR is continuously extended by active method development. One of the breakthroughs in ssNMR was the advent of fast sample spinning, which is constantly redefined with maximum frequencies exceeding 60 kHz, then 100 kHz and, recently, even 170 kHz. In this context, Duong et al. showed how spinning between 60–70 kHz allows the measurement of selective ^{1}H-^{14}N distances by exciting and detecting protons that are coupled to ^{14}N, which provides very useful information for biological systems without isotopic labelling. In the study dipolar recoupling is achieved by saturating the overtone transitions of the natural abundant nitrogen spins. In the example of Paluch et al. several ^{13}C homonuclear mixing schemes are systematically compared with the aim to facilitate assignment of methyl groups in high-molecular weight proteins. The authors demonstrated dramatic improvements in methyl to alpha carbon transfer efficiencies upon increasing from 55 to 95 kHz spinning rates. In another contribution to this issue, Zadorozhnyi et al. described a spectral editing technique to determine histidine protonation states, which play important functional roles in proton transfer, metal binding *etc.* The authors demonstrated how highly resolved information can be

obtained using ^{1}H detection within 2D experiments based on selective inversion of ring nitrogen atoms. To contrast the methodology relying on fast spinning, van der Wel comprehensively reviews moderate spinning rates based methods for measuring dihedral angles in peptides and proteins as a complement to more commonly recorded distance restraints. The emphasis is on direct measurements *via* correlations of anisotropic interactions including ϕ and ψ backbone dihedral angles *via* HNCH and NCCN experiments, sidechain angles *via* HCCH, the peptide bond angle ω, and long-range angle restraints between backbone amides. The review contains various examples on real systems, including amyloids, and details on the particularity of the pulse programs.

Several interesting applications can be viewed in this special issue. A field to which ssNMR has been continuously contributing key structural and functional information is that of amyloids and other protein aggregation phenomena. Qi et al. provide an excellent example for the complementarity and contribution of ssNMR studies to our understanding of fibril formation. Since ssNMR does not require long-range crystallinity of the samples, they have been able to study variants of the Y145Stop mutant of the human prion protein, which is associated with hereditary prionopathy. The variants were made of different deletions in the flexible N-terminal tail, and the authors have been able to show how those deletions affect or maintain the aggregation properties of the protein on an atomic level. In the review of Kitoka et al., many aspects of the tau protein, as viewed by NMR and in particular ssNMR spectroscopy, are discussed. The tau protein forms intracellular neurofibrillary tangles in neurons, and is a major drug target to treat Alzheimer's disease. The review describes solution NMR efforts to study the monomeric form, its secondary structure, and the effects of phosphorylation on aggregation properties. Solid-state NMR studies, including dynamic nuclear polarization (DNP), contributed to our understanding of oligomer and filamentous structures of various tau constructs including a three-dimensional structural model of its fold in the fibrillar form.

An additional application area to which ssNMR techniques have significantly contributed is the study of membrane proteins in close to native environments, and in particular it is possible to study them in membrane bilayers. Amani et al. contributed a structural study of the potassium channel KirBac1.1 from the bacterium *Burkholderia pseudomallei* that causes Melioidosis. An original X-ray structure lacked 85 residues of the total of the 333, mainly in the N- and C-terminus regions. Using T2-filtered ssNMR experiments, the authors generated surface accessibility potentials based on the assumption that only those residues in the vicinity of water can be detected.

Yet another field where ssNMR has increasing impact is study of biomolecular complexes. Recent progress in the study of RNA and ribonucleic acid–protein complexes (RNPs) are discussed in detail by Aguion and Marchanka. This unique review discusses strategies to label synthetic RNA, means to assign RNA polynucleotides (including ^{1}H spins), and dedicates attention to discussions on the complexities and possible opportunities. Discussions on the particular stages in assigning ribose, base, their linkage, and sequential contacts are highly detailed and accompanied by many examples, pulse sequence details, and spectra. One of the key protein-DNA complexes in the cell is chromatin, consisting of the DNA wrapped with histone proteins. Ackermann and Debelouchina describe the emerging contributions of ssNMR to understanding this complex and highly important gene expression system. They discuss the details of histone preparation and isotopic labeling strategies including DNA and the four histones, as well as post translational modifications, techniques and results from studies of the rigid core and flexible histone tails, and chromatin modulators. While chromatin studies by NMR are highly complex, the set of biochemical and NMR tools presented in this review will help to advance further understanding of chromatin structural biology.

Given that a majority of the applications in the field of biomolecular ssNMR focus on proteins, and recently more studies on polynucleotides emerge (see review below), it is interesting to see that significant progress is also achieved to study the complex polysaccharide networks making up cell walls. Fernando et al. studied the polymorphism of carbohydrates making up the fungal cell wall. They find that the chitin moiety shows similarity to the α- and γ-allomorphs and is not significantly altered in the presence of anti-fungal treatment. In addition, statistical analysis revealed that chitosan (a deacetylation product of chitin) from *R. delemar* and *A. sydowii* share some similarity to Type-II chitosan (a relaxed two-fold helix conformation) but is completely different from Type-I.

Besides being a tool for structural characterization ssNMR is also important for its ability to probe dynamics spanning several orders of magnitude in time scale on challenging systems including fibrils, membrane proteins and large biomolecular complexes. To obtain a comprehensive view of the molecular motions in different regimes requires different complementary methods. For example, Vugmeyster et al. demonstrated how slow motions (in the order or 10^{4}–10^{5} s^{-1}) can be probed by utilizing ^{2}H Chemical Exchange Saturation Transfer (CEST) techniques, both at slow and fast MAS rates. In another contribution, Franks et al. show how at fast spinning NH dipolar couplings can be measured using newly optimized symmetry-based pulses, previously utilized mostly at moderate spinning of 10–30 kHz, to enable such measurements for large protein complexes requiring high sensitivity afforded by proton detected experiments at fast spinning. Dipolar couplings report on cumulative amplitudes of motion for picosecond to microsecond motions and thus valuable parameters for characterizing dynamics on their own but also often employed to restrain overall motional amplitudes in model-free types of

analyses of relaxation rates. In a related context, Zumpfe and Smith provide an insightful review of methods to quantify protein dynamics based on relaxation rate measurements in the solid state. They consider the model-free, extended model-free, spectral density mapping, and the LeMaster's approaches highlighting their advantages, disadvantages and pitfalls that can lead to erroneous interpretation of molecular motions. The authors then show the advantage of the detectors method, in particular its generality and its use to describe molecular dynamics (MD) and thus extend our correlation of NMR and MD simulation data.

Two additional contributions demonstrate the strength of combining ssNMR with dynamic nuclear polarization (DNP) to study in-cell NMR. While this field is still in its infancy, studies slowly reveal both the technicalities and the advantages of such experiments that have to be performed at low temperatures (~100 K) and with radicals. For example, Overall and Barnes discuss the effects of DNP radicals and cryoprotectants on cell viability (using human Jurkat cells) and signal enhancement showing that 10% d_6-DMSO maintains the same enhancement as "DNP juice" (60/30/10 d_8-glycerol/D_2O/H_2O) motivated by the superiority of DMSO with respect to the conditions of the cells. Very similarly, Xiao et al. has shown that in Human embryonic kidney 293 (HEK293) cells, incubation with the radical AMUPol and using 10% DMSO as a cryoprotectant along with slow cooling, were essential for cell integrity and provided similar enhancements as 15% glycerol. It was also shown that distribution of the radicals within the cells was non-uniform.

Overall, this special issue covers a large variety of topics providing insight into the diversity of applications of ssNMR, the state-of-the-art technology, and the wide range of experimental approaches that are available and that continuously extend to fit new applications and new needs.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

frontiers in Molecular Biosciences

METHODS
published: 08 April 2021
doi: 10.3389/fmolb.2021.645347

Selective ^{1}H-^{14}N Distance Measurements by ^{14}N Overtone Solid-State NMR Spectroscopy at Fast MAS

*Nghia Tuan Duong[1], Zhehong Gan[2] and Yusuke Nishiyama[1,3]**

[1]NMR Science and Development Division, RIKEN SPring-8 Center, Nano-Crystallography Unit, RIKEN-JEOL Collaboration Center, Yokohama, Japan, [2]Centre of Interdisciplinary Magnetic Resonance, National High Magnetic Field Laboratory, Tallahassee, FL, United States, [3]JEOL RESONANCE Inc., Tokyo, Japan

Accurate distance measurements between proton and nitrogen can provide detailed information on the structures and dynamics of various molecules. The combination of broadband phase-modulated (PM) pulse and rotational-echo saturation-pulse double-resonance (RESPDOR) sequence at fast magic-angle spinning (MAS) has enabled the measurement of multiple ^{1}H-^{14}N distances with high accuracy. However, complications may arise when applying this sequence to systems with multiple inequivalent ^{14}N nuclei, especially a single ^{1}H sitting close to multiple ^{14}N atoms. Due to its broadband characteristics, the PM pulse saturates all ^{14}N atoms; hence, the single ^{1}H simultaneously experiences the RESPDOR effect from multiple ^{1}H-^{14}N couplings. Consequently, no reliable H-N distances are obtained. To overcome the problem, selective ^{14}N saturation is desired, but it is difficult because ^{14}N is an integer quadrupolar nucleus. Alternatively, ^{14}N overtone (OT) NMR spectroscopy can be employed owing to its narrow bandwidth for selectivity. Moreover, owing to the sole presence of two energy levels ($m = \pm 1$), the ^{14}N OT spin dynamics behaves similarly to that of spin-1/2. This allows the interchangeability between RESPDOR and rotational-echo double-resonance (REDOR) since their principles are the same except the degree of ^{14}N OT population transfer; saturation for the former whereas inversion for the latter. As the ideal saturation/inversion is impractical due to the slow and orientation-dependent effective nutation of ^{14}N OT, the working condition is usually an intermediate between REDOR and RESPDOR. The degree of ^{14}N OT population transfer can be determined from the results of protons with short distances to ^{14}N and then can be used to obtain long-distance determination of other protons to the same ^{14}N site. Herein, we combine the ^{14}N OT and REDOR/RESPDOR to explore the feasibility of selective ^{1}H-^{14}N distance measurements. Experimental demonstrations on simple biological compounds of L-tyrosine.HCl, N-acetyl-L-alanine, and L-alanyl-L-alanine were performed at 14.1 T and MAS frequency of 62.5 kHz. The former two consist of a single ^{14}N site, whereas the latter consists of two ^{14}N sites. The experimental optimizations and reliable fittings by the universal curves are described. The extracted ^{1}H-^{14}N distances by OT-REDOR are in good agreement with those determined by PM-RESPDOR and diffraction techniques.

Keywords: ^{1}H-^{14}N distances, ^{14}N overtone spectroscopy, PM-S-RESPDOR, REDOR, fast MAS frequency

OPEN ACCESS

Edited by:
Amir Goldbourt,
Tel Aviv University, Israel

Reviewed by:
Luke O'Dell,
Deakin University, Australia
Aaron Rossini,
Iowa State University, United States

Correspondence:
Yusuke Nishiyama
yunishiy@jeol.co.jp

Specialty section:
This article was submitted to Structural Biology, a section of the journal Frontiers in Molecular Biosciences

Received: *23 December 2020*
Accepted: *28 January 2021*
Published: *08 April 2021*

Citation:
Duong NT, Gan Z and Nishiyama Y (2021) Selective ^{1}H-^{14}N Distance Measurements by ^{14}N Overtone Solid-State NMR Spectroscopy at Fast MAS. Front. Mol. Biosci. 8:645347. doi: 10.3389/fmolb.2021.645347

INTRODUCTION

H-N distance is of importance for deeper insights into the structures and dynamics of chemical and biological systems due to the ubiquity of both proton and nitrogen. Such distance can be obtained by solid-state nuclear magnetic resonance (ssNMR) through the determination of H-N dipolar coupling, which is inversely proportional to the cube of the H-N distance. There are a few reasons why ssNMR has advantages for the H-N measurement over traditional diffraction techniques. First, ssNMR spectroscopy is applicable to various systems no matter their states, i.e., lacking long-range order or even being a disorder, which are intractable by diffraction techniques. Second, it enables the precise location of the H-atom positions, which is poorly determined by X-ray diffraction (XRD) or electron diffraction (Guzmán-Afonso et al., 2019).

Despite such advantages, the H-N distance measurement by NMR has difficulties due to some unfavorable nuclear characteristics of these two elements. For nitrogen, it has two stable and NMR active isotopes, ^{15}N and ^{14}N. The former is preferred in ssNMR because it is a spin-1/2 nucleus; thus, it is easy to manipulate and to obtains high resolution. Many methods have been designed to measure ^{1}H-^{15}N distances (Hohwy et al., 2000; Zhao et al., 2001; Schnell and Saalwächter, 2002; Fu, 2003; Chevelkov et al., 2009; Hou et al., 2011; Schanda et al., 2011; Paluch et al., 2013; Hou et al., 2014; Nishiyama et al., 2016). Nevertheless, the main drawback of ^{15}N isotope is its insensitivity owing to the low natural abundance (0.4%). It makes the measurements lengthy for sufficient signal-to-noise ratio (S/N); otherwise, 1) the isotopic labeling is needed, which is not always simple and cost-effective or 2) dynamic nuclear polarization experiments are required (Zhao et al., 2018). However, the ^{1}H-^{15}N experiments on that work only allowed the determination of the shortest ^{1}H-^{15}N distance due to the dipolar truncation effect. On the other hand, ^{14}N isotope benefits from the high natural abundance (99.6%), but it suffers from the severe quadrupolar broadening and complicated spin dynamics because ^{14}N is an integer quadrupolar nucleus (spin $I = 1$). For protons, the intense ^{1}H–^{1}H homonuclear dipolar couplings in the solid state cause ^{1}H line broadening and shorten the ^{1}H coherence time. Consequently, these unfavorable characteristics of both N isotopes and H nucleus make H-N distance measurement by NMR challenging.

The development of fast magic-angle spinning (MAS, $\nu_R \geq$ 60 kHz) with proton detection has made ^{14}N NMR spectroscopy a routinely used method, overcoming the difficulty associated with quadrupolar interaction (Cavadini et al., 2006; Gan et al., 2007; Cavadini, 2010; Nishiyama et al., 2011; Brown, 2014; Pandey and Nishiyama, 2015; Shen et al., 2015; Pandey et al., 2016; Carnevale et al., 2017; Hung et al., 2019; Jarvis et al., 2019; Rankin et al., 2019; Wijesekara et al., 2020). Furthermore, under fast MAS conditions, the strong ^{1}H–^{1}H dipolar network is largely suppressed (Nishiyama, 2016). These two advantages potentially facilitate the ^{1}H-^{14}N distance measurement. Recently, our group have introduced a combination of phase-modulated (PM) pulse (Nimerovsky et al., 2014; Makrinich et al., 2017; Makrinich et al., 2018), $SR4_1^2$ recoupling (Brinkmann and Kentgens, 2006), and rotational-echo saturation-pulse double-resonance (RESPDOR) (Gan, 2006; Chen et al., 2010a; Chen et al., 2010b; Lu et al., 2011) (PM-S-RESPDOR) that can extract multiple ^{1}H-^{14}N distances with high accuracy at fast MAS of 70 kHz (Duong et al., 2019). Such success mainly comes from the robustness that universal fraction curves can be obtained for the distance measurement under the saturation by the PM pulse for a wide range of ^{14}N quadrupolar coupling constant (C_Q) and ^{1}H-^{14}N dipolar coupling. This broadband characteristics of PM pulse is useful when we work on systems containing a single ^{14}N site, as shown in the previous study. However, complications may arise for systems where multiple ^{14}N sites are present, as shown below.

For example, a 5-spin system, as shown in **Figure 1**, consists of two N and three H atoms. We assume that the ^{14}N and ^{1}H NMR peaks are well resolved for simplicity. The first difficulty associated with this system is the ambiguity of ^{1}H-^{14}N distance measurement. For instance, we can determine the distance of H3-N by PM-S-RESPDOR sequence but cannot know whether such distance is between H3 and N1 or H3 and N2. The second difficulty relates to the complex spin dynamics of H2 nucleus, which is close to both N1 and N2 nuclei. As PM pulse is broadband, it completely saturates both ^{14}N1 and ^{14}N2 nuclei; thus, the PM-S-RESPDOR sequence will give the H2-N fraction curve experiencing the combined effects of H2-N1 and H2-N2 pairs. The H2-N distance from the fraction curve would be shorter than those extracted from H2-N1 or H2-N2 pair; or in other words, no reliable distance is yielded. A solution to overcome this cumulative contribution is to selectively saturate each N nucleus, which can be achieved in the manner of Delays Alternating with Nutation for Tailored Excitation (DANTE) (Vitzthum et al., 2011; Vitzthum et al., 2012; Lu et al., 2013; Pourpoint et al., 2014). This approach can be our future work.

An alternative approach is the ^{14}N overtone (OT) NMR spectroscopy, where the forbidden transitions $|\Delta m| = 2$ are weakly allowed (m is the energy level) (Bloom and LeGros, 1986; Tycko and Opella, 1987; Jayanthi and Ramanathan, 2011; O'Dell and Ratcliffe, 2011; Nishiyama et al., 2013; O'Dell and Brinkmann, 2013; O'Dell et al., 2013; Haies et al., 2015a; Haies et al., 2015b; Shen et al., 2017; Concistré et al., 2018; Gan et al., 2018; Pandey and Nishiyama, 2018). Because it is twice the fundamental frequency, ^{14}N OT frequency is more available to commercial MAS probes since many probes are not designed to

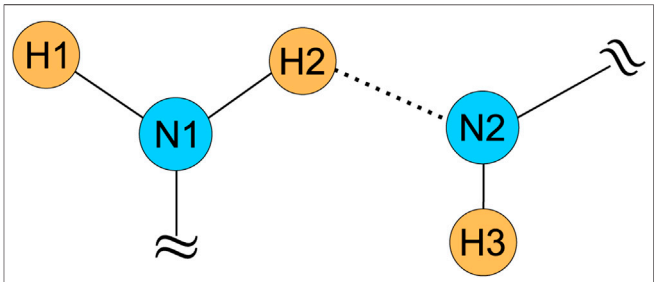


FIGURE 1 | The 5-spin system with two N and three H atoms complicating the ^{1}H-^{14}N distance measurements by PM-S-RESPDOR.

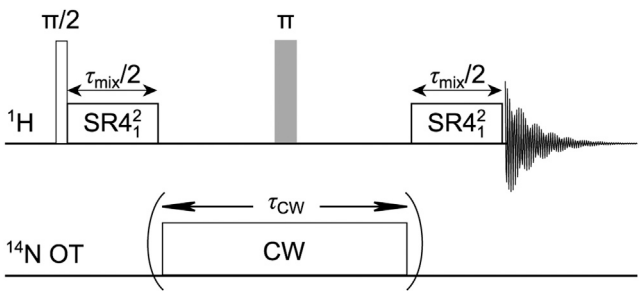


FIGURE 2 | ^{1}H-^{14}N OT-REDOR sequence. On ^{1}H channel, $SR4_1^2$ recoupling is used to reintroduce the ^{1}H-^{14}N dipolar couplings. On ^{14}N OT channel, CW pulse is used to perturb ^{14}N OT populations. For the fraction curve, two sets of data (S_0 and S') are acquired by employing OT-REDOR sequence without and with CW pulse (within the two brackets), respectively.

tune to ^{14}N Larmor frequency. Importantly, the ^{14}N overtone excitation can achieve band-selective observation of ^{14}N (Pandey and Nishiyama, 2018). The narrow bandwidth results from the slow effective ^{14}N OT nutation, which is proportional to C_Q/ν_0, where ν_0 is the ^{14}N Larmor frequency. Besides the selectivity, ^{14}N OT spectra are free from the first-order quadrupolar interaction because of the symmetric transitions, $m = -1 \leftrightarrow m = +1$. Therefore, ^{14}N OT NMR is much narrower than the single-quantum ^{14}N spectra and robust to the misadjustment of the magic angle. Moreover, since the transitions are only between two energy levels involved in OT ($m = \pm 1$), the spin dynamics of ^{14}N OT behaves similarly to that of spin-1/2. Hence, for a ^{1}H-^{14}N OT system, the working conditions under RESPDOR can also be described by rotational-echo double-resonance (REDOR) (Gullion and Schaefer, 1989; Gullion, 2007) depending on whether the population transfer is saturation (RESPDOR) or inversion (REDOR) (Nimerovsky et al., 2017). Since the ideal saturation or inversion by continuous-wave (CW) is impractical, the working condition is an intermediate between REDOR and RESPDOR regimes. In this work, we combine ^{14}N OT and REDOR sequence (^{1}H-^{14}N OT-REDOR) to explore its feasibility for distance measurements. This sequence is firstly demonstrated using two model biological compounds of L-tyrosine.HCl (Tyr) and N-acetyl-L-alanine (AcAla) and then applied to a more complex dipeptide system of L-alanyl-L-alanine (AlaAla) that involves two inequivalent nitrogen sites in a single molecule.

PULSE SEQUENCE AND THE UNIVERSAL EXPRESSION

Figure 2 depicts the ^{1}H-^{14}N OT-REDOR sequence. It is identical to the conventional S-REDOR sequence (Chen et al., 2010b), where $SR4_1^2$ recoupling (lasting for τ_{mix}) is used to recover the ^{1}H-^{14}N dipolar coupling and CW (lasting for τ_{CW}) is used to saturate/invert the ^{14}N OT populations between the two energy levels. We note that since $SR4_1^2$ is not γ-encoded, the interval between the two SR4 blocks should be rotor-synchronized to avoid the spatial modulation of the recoupled ^{1}H-^{14}N dipolar couplings. For distance measurement, we measure two signals, S_0 and S', acquired without and with CW pulse, respectively, for obtaining the fraction curve $\Delta S/S_0 = (S_0 - S')/S_0$ as a function of τ_{mix}.

The fraction curve excludes the signal attenuation from T_2 relaxation, making it dependent solely on the dipolar coupling constant as for the case of REDOR with complete inversion. However, for ^{14}N OT, complete inversion is difficult to achieve. Subsequently, the distance is extracted by fitting the fraction curve to either the numerically exact or universal curves. For the numerically exact curve, the simulation is extremely difficult as it must work in the laboratory frame without high-field approximation, leading to time-consuming calculations (O'Dell and Brinkmann, 2013). Even if this condition is met, various parameters must be known; for instance, the ^{14}N C_Q and the ^{14}N OT effective nutation fields, which are not straightforwardly determined. For the universal curve, it has shown to be an almost identical match to the numerically exact ^{1}H-^{14}N PM-S-RESPDOR curve, allowing simple distance extractions (Duong et al., 2019). In addition, the condition and knowledge required by the numerically exact ^{1}H-^{14}N OT-REDOR curve above are not necessary for the universal curve. Indeed, semiquantitative evaluation for ^{14}N OT transitions only requires the fitting parameter f and the ^{1}H-^{14}N dipolar coupling (shown below). Thus, for objective fitting, we use the universal curve approach, which is derived by following the original work of Gan or later analysis of Chen and coworkers (Gan, 2006; Chen et al., 2010a). The derivation starts with fundamental ^{14}N transitions for verification and then applies to ^{14}N OT.

A general expression for any spin and type of experiment is

$$\frac{\Delta S}{S_0} = 1 - \sum_{i,j} P_i W_{ij} REDOR(|\Delta m|), \quad (1)$$

where $P_i = 1/(2I+1)$ is the population of spin state $m = i$ under high-temperature approximation, W_{ij} is the population transfer probability from $m = i$ to $m = j$ spin state, and $REDOR(|\Delta m|)$ presents the normalized dipolar-dephased signal intensity for classical REDOR. The general expression helps to derive the universal curves mentioned in **Figure 3**. It is worth noting that the natural abundance of a specific isotope should also be considered in **Eq. (1)**. However, the natural abundance of ^{14}N isotope is 99.6%, very close to 100%; hence, we can safely neglect it.

For ^{14}N ($I = 1$) spin, under the Zeeman interaction with the external magnetic field, there are three energy levels of $m = 0$ and ± 1. We assume the population for each level is 1/3. The population transfers among the energy levels are categorized into zero- (ZQ or $\Delta m = 0$), single- (SQ or $|\Delta m| = 1$), and double-quantum (DQ or $|\Delta m| = 2$) transitions, which determine the RESPDOR effect. Both energy levels $m = +1$ and $m = -1$ are involved in all three ZQ, SQ, and DQ transitions, as shown in **Figure 3A**. Under the ideal saturation of ^{14}N spin, $W_{\pm 1j}$ for these transfers are equal; hence, each transition has $W_{\pm 1j}$ of 1/3. On the other hand, the energy level $m = 0$ is only involved in ZQ and SQ transitions, but

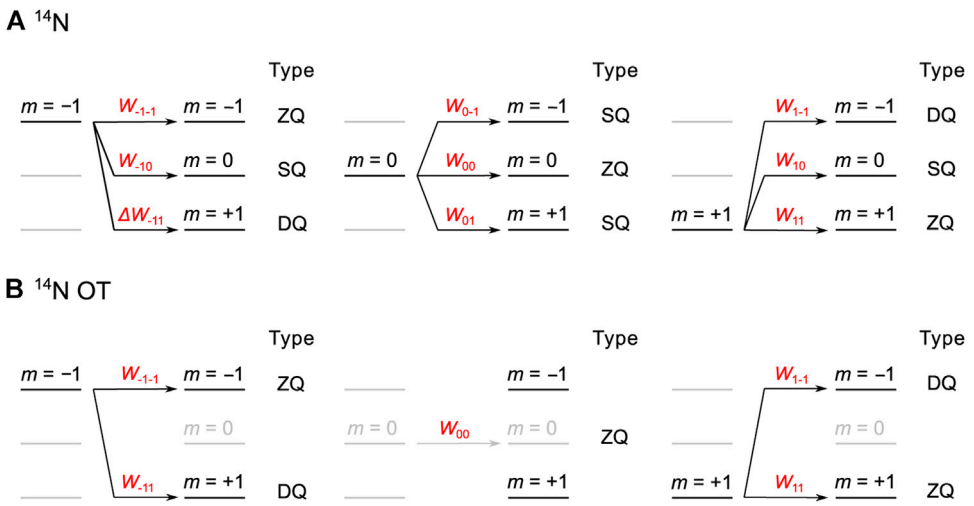


FIGURE 3 | The population transfers and their probabilities (W_{ij}) for ZQ ($\Delta m = 0$), SQ ($|\Delta m| = 1$), and DQ ($|\Delta m| = 2$) transitions of **(A)** ^{14}N and **(B)** ^{14}N OT to derive the universal curves.

there are two SQ transitions of ($m = 0 \rightarrow m = -1$) and ($m = 0 \rightarrow m = +1$). Hence, for $m = 0$, W_{0j} of ZQ, SQ, and DQ transitions are 1/3, 2/3, and 0/3, respectively. Taken together, under the complete saturation of ^{14}N spin, $\Sigma P_i W_{ij}$ for ZQ, SQ, and DQ transitions for $m = (-1, 0, 1)$ are 1/3·3/3, 1/3·4/3, and 1/3·2/3, respectively (see **Figure 3A**). Replacing these P_i and W_{ij} in **Eq. (1)**, the universal expression for ^{14}N is given by

$$\begin{aligned}\frac{\Delta S}{S_0} &= 1 - \frac{3}{9} - \frac{4}{9}.REDOR(|\Delta m| = 1) - \frac{2}{9}.REDOR(|\Delta m| = 2)\\ &= \frac{2}{3} - \frac{\pi\sqrt{2}}{9} J_{1/4}\left(\frac{\pi}{4}(b_{1H-14N}/2\pi)\tau_{mix}\right) J_{-1/4}\left(\frac{\pi}{4}(b_{1H-14N}/2\pi)\tau_{mix}\right)\\ &\quad - \frac{\pi\sqrt{2}}{18} J_{1/4}\left(\frac{2\pi}{4}(b_{1H-14N}/2\pi)\tau_{mix}\right) J_{-1/4}\left(\frac{2\pi}{4}(b_{1H-14N}/2\pi)\tau_{mix}\right),\end{aligned} \tag{2}$$

where $J_{\pm 1/4}$ denotes the ±1/4-order Bessel functions of the first kind and $b_{1H-14N}/(2\pi)$ is the ^{1}H-^{14}N dipolar coupling constant while τ_{mix} is the total mixing time of $SR4_1^2$ recoupling sequence. **Eq. (2)** is identical to the universal curve for ^{1}H-^{14}N RESPDOR in the literature (Gan, 2006; Chen et al., 2010b), verifying our analysis.

Next, we consider the case of ^{14}N OT. Again, three energy levels are present with the population P_i of 1/3 for each level. The energy level $m = 0$ is not involved in OT transitions; thus, it is blurred in **Figure 3B**. P_i of $m = 0$ remains at ZQ transition, meaning that W_{0j} is 1 for $j = 0$ and 0 for $j \neq 0$. Conversely, both energy levels $m = \pm 1$ are involved in the saturation/inversion of ZQ and DQ transitions. However, owing to the slow and orientation-dependent effective nutation of ^{14}N OT, the complete saturation/inversion is difficult. Considering this incompletion, we assume that the DQ $W_{\pm 1j}$ for $m = \pm 1$ are f with $0 \leq f \leq 1$, in which $f = 0.5$ corresponds to complete saturation while $f = 1.0$ corresponds to complete inversion. Although we mentioned that the working condition for ^{1}H-^{14}N OT-REDOR is between REDOR and RESPDOR regimes in the *Introduction* section, this does not mean that f should be between 0.5 and 1.0. Indeed, if the complete saturation is not achieved, parameter f could be smaller than 0.5. With the introduction of f, $W_{\pm 1j}$ for ZQ and DQ transitions are $1-f$ and f, respectively. Combining P_i of each transition for each level and under incomplete saturation/inversion of ^{14}N OT, **Figure 3B** shows that $\Sigma P_i W_{ij}$ for ZQ, SQ, and DQ transitions are $1/3 \cdot (1 + 2(1-f))$, 0, and $1/3 \cdot 2f$, respectively. The universal expression for REDOR/RESPDOR on ^{14}N OT is given by

$$\begin{aligned}\frac{\Delta S}{S_0} &= 1 - \left(1 - \frac{2f}{3}\right) - \frac{2f}{3}.REDOR(|\Delta m| = 2)\\ &= \frac{2f}{3}[1 - REDOR(|\Delta m| = 2)]\\ &= \frac{2f}{3}\left[1 - \frac{\pi\sqrt{2}}{4} J_{1/4}\left(\frac{2\pi}{4}(b_{1H-14N}/2\pi)\tau_{mix}\right) J_{-1/4}\left(\frac{2\pi}{4}(b_{1H-14N}/2\pi)\tau_{mix}\right)\right].\end{aligned} \tag{3}$$

From **Eq. (3)**, the coefficient for REDOR ($|\Delta m| = 2$) is proportional to f, affecting the slope of the fraction curve. However, since f uniformly affects the other elements in the equation, the universal curves derived from **Eq. (3)** would reach the maximum at the same τ_{mix} no matter f value. It is worth noting that the introduction of f makes the fitting among universal curves and experimental fraction curve better, but it makes the extracted b_{1H-14N} inaccurate. Particularly, when the fraction curve has not reached the maximum $\Delta S/S_0$, universal curves generated by different combinations of f and b_{1H-14N} can reproduce the very similar fraction curve, thus giving ambiguous results. To avoid this situation, our fitting strategy consists of two steps. The first is to determine f, which is possible only when the fraction curve of the shortest H-N distance must show the

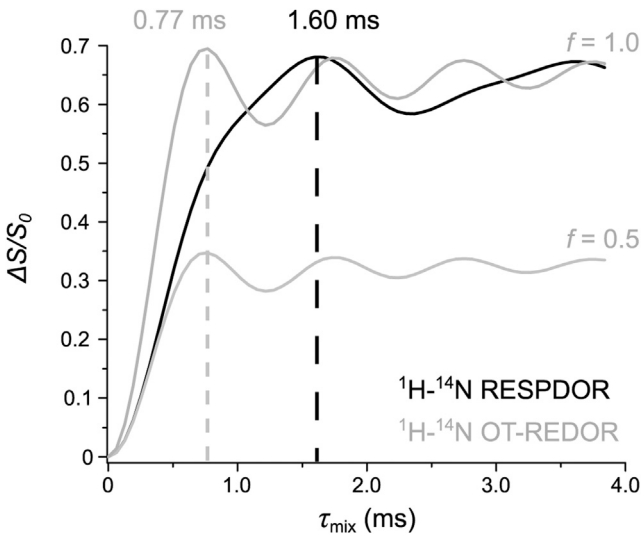


FIGURE 4 | Comparison between the universal curves of ^{1}H-^{14}N RESPDOR (black) and ^{1}H-^{14}N OT-REDOR (grey) with f = 0.5 (complete saturation) and 1.0 (complete inversion) derived from **Eqs 2,3**, respectively, under the same ^{1}H-^{14}N dipolar coupling of 2.0 kHz. The optimum τ_{mix} of both curves are shown and highlighted by the dashed lines.

maximum. Under this condition, the fitting parameter f is determined as the ratio of the experimental and theoretical maxima $\Delta S/S_0$ (2/3 = 0.67). That precisely known f leads to the unambiguous determination of b_{1H-14N}. For longer H-N distance, the REDOR curve may possibly not show the maximum while the oscillation is damped, making the fitting difficult. Under this situation, accurate distance determinations are still possible by the second step. It is to use this observed f from the first step for measuring longer ^{1}H-^{14}N distances of the same ^{14}N site. This strategy is a disadvantage of OT-REDOR compared to PM-S-RESPDOR. The latter does not require the prior knowledge of the fitting parameter f owing to the complete saturation of all ^{14}N crystallites by the PM pulse, thus enabling the reliable fitting even when the maximum is not observed.

It is of practical use to clarify the differences between the universal expressions for ^{1}H-^{14}N RESPDOR and ^{1}H-^{14}N OT-REDOR. This can be done by comparing the universal curves resulting from **Eqs. 2, 3** under the same b_{1H-14N}. **Figure 4** compares the three universal curves, one from **Eq. (2)** and two from **Eq. (3)** with f = 0.5 (complete saturation) and f = 1.0 (complete inversion). These two f values are chosen because they correspond to ideal RESPDOR (f = 0.5) and REDOR (f = 1.0) conditions and our working condition is an intermediate between these two, as mentioned in the *Introduction*. The two curves from **Eq. (3)** are identical except for the intensity (a factor of 2), which is in agreement with the discussion above (see **Figure 4**). A notable difference between the universal curves of ^{1}H-^{14}N OT and that of ^{1}H-^{14}N is that the dephasing rate of the former curves is about two times faster than that of the latter (0.77 ms compared to 1.60 ms, respectively). This is because, for ^{1}H-^{14}N OT, the REDOR effect is determined by the DQ transitions, whereas, for ^{1}H-^{14}N, the RESPDOR effect is determined by both the SQ and DQ transitions. Such a faster dephasing rate associated with the multiple quantum transitions has been known in the literature (Pruski et al., 1999). This potentially allows ^{1}H-^{14}N OT-REDOR to probe long ^{1}H-^{14}N distance better than ^{1}H-^{14}N RESPDOR as it is less affected by the poor sensitivity and uncertainty of $\Delta S/S_0$ at long τ_{mix}.

RESULTS AND DISCUSSIONS

In this section, we firstly explore the feasibility of ^{1}H-^{14}N OT-REDOR on two model biological compounds of Tyr and AcAla. These two compounds only consist of a single ^{14}N site and were well characterized by ^{1}H-^{14}N PM-S-RESPDOR in the previous study. Upon the feasibility test, we apply this technique to a more complex dipeptide system of AlaAla where two inequivalent ^{14}N sites are present.

L-Tyrosine.HCl

In order to obtain an efficient ^{1}H-^{14}N OT-REDOR fraction curve, experimental optimizations are required. Such optimizations require the knowledge of ^{14}N OT resonance frequency as it significantly affects the sensitivity of OT experiments due to the narrow bandwidth. In this current work, the ^{14}N OT frequency is indirectly determined by the two-dimensional (2D) ^{1}H-{^{14}N OT} *D*- or T-HMQC (sequences in **Supplementary Figures S1A,B** respectively) at the second OT spinning sideband (n = -2 SSB) for the highest S/N. Such n value depends on the sense of rotation with respect to the magnetic field (Gan et al., 2018). For our configuration, the

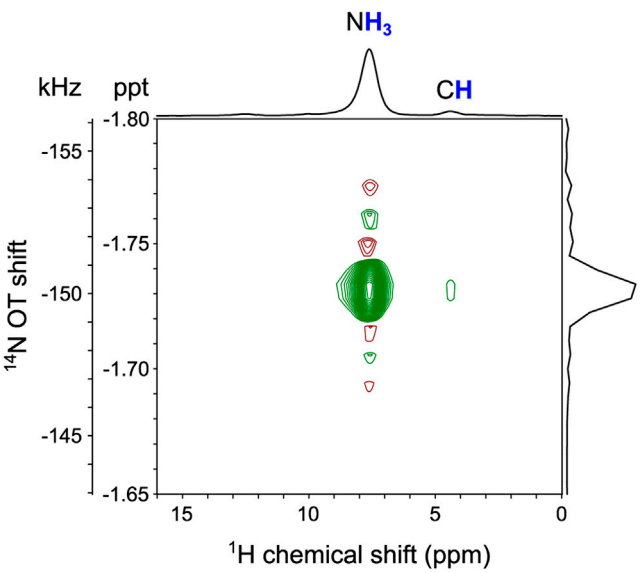


FIGURE 5 | The 2D ^{1}H-{^{14}N OT} *D*-HMQC experiment of Tyr at the magnetic field (B_0) of 14.1 T and ν_R of 62.5 kHz. The experiment was performed at the second ^{14}N OT SSB (n = -2) for the highest S/N. Further details are given in the *Experiments* section.

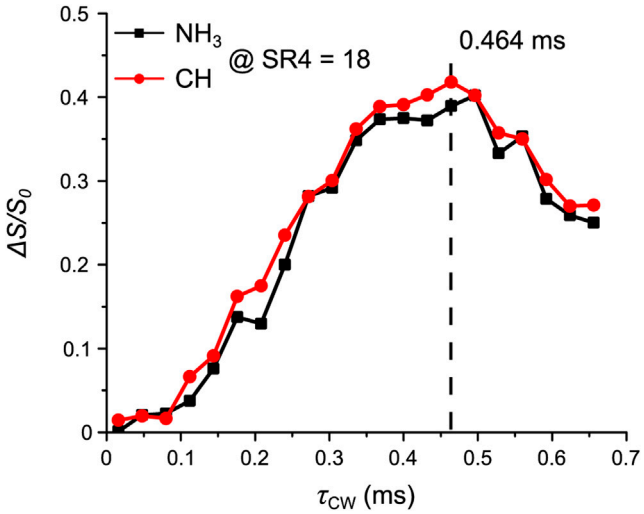


FIGURE 6 | Tyr: the signal fraction $\Delta S/S_0$ of N**H₃** (black squares) and C**H** (red circles) as a function of τ_{CW}. Experiments were performed at τ_{mix} of 1.15 ms. The optimum τ_{CW} is shown and highlighted by the dashed line. Further experimental details are given in the *Experiments* section.

magnetic field is toward the top of the magnet and the spinning rotation is clockwise looking from the top. Taking benefits of optimum S/N, all the remaining experiments were also performed at the second ^{14}N OT SSB (n = −2). **Figure 5** clearly indicates the ^{14}N OT frequency and its correlations to both proton sites of N**H₃** and C**H** by the *D*-HMQC experiment. The smaller correlation of N to C**H** under τ_{mix} of 0.51 ms is explained due to the longer ^{1}H-^{14}N distance compared to the directly bonded H-N distance of the NH_3 group. This 2D spectrum is in agreement with the ^{1}H-{^{14}N} *D*-HMQC spectrum in the previous study (Duong et al., 2019).

After the ^{14}N OT frequency has been determined, the next parameter for optimization is τ_{CW} so that as many as possible ^{14}N crystallites can be saturated/inverted. The ^{14}N OT rf-field was 120 kHz, the highest technically possible value. It was calibrated by the use of the Bloch–Siegert shift of the proton approach (Hung et al., 2020). However, it is noted that the ^{14}N OT nutation frequency is scaled on C_Q, the magnetic field, the powder distribution, making it much weaker than ^{14}N OT rf-field. **Figure 6** shows the signal fraction $\Delta S/S_0$ of N**H₃** and C**H** at a fixed τ_{mix} of 1.15 ms (or 18 loops of $SR4_1^2$ recoupling blocks) under varying τ_{CW} values.

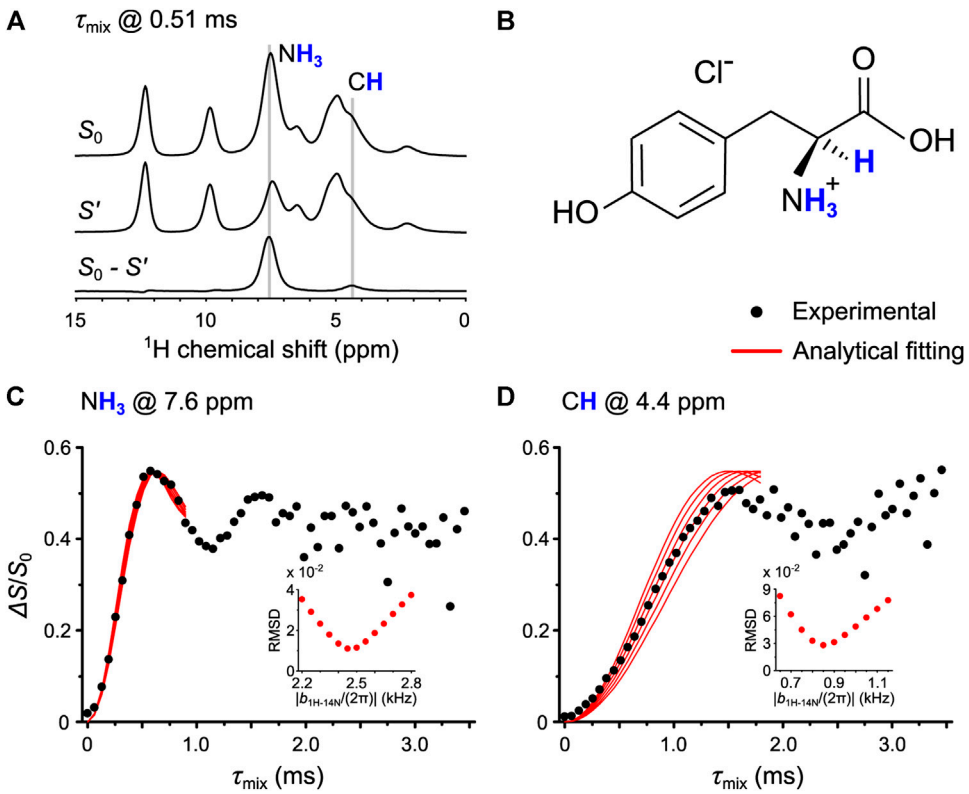


FIGURE 7 | (A) The 1D ^{1}H spin echo (S_0), the dephased (S'), and the difference (S_0–S') spectra of Tyr. The former two were extracted from ^{1}H-^{14}N OT-REDOR with τ_{mix} of 0.51 ms, whereas the latter is calculated from the former two. **(B)** The molecular structure of Tyr where N**H₃** and C**H** are presented. (**C,D)** The fitting of experimental ^{1}H-^{14}N OT-REDOR fraction curves (black circles) by the universal curves (red lines) for N**H₃** at 7.6 ppm in **(C)** and C**H** at 4.4 ppm in **(D)**. The fitting parameter *f* for universal curves is 0.82. The RMSD analyses (inset) were calculated for the best fitting ^{1}H-^{14}N dipolar couplings.

TABLE 1 | ^{1}H-^{14}N distances measured by OT-REDOR, PM-S-RESPDOR, and ND on Tyr.

	OT-REDOR	PM-S-RESPDOR	ND
NH$_3$	1.05 ± 0.03 Å	1.16 Å	1.01 Å
CH	2.21 ± 0.10 Å	2.24 Å	2.10 Å

Once the parameters were optimized, we performed the ^{1}H-^{14}N OT-REDOR experiment on Tyr. **Figure 7A** shows the one-dimensional (1D) ^{1}H spin echo (S_0), the dephased (S'), and the difference (S_0–S') spectra, in which the former two were extracted from the ^{1}H-^{14}N OT-REDOR experiment with τ_{mix} of 0.51 ms, whereas the latter is calculated from the former two. The N**H$_3$** and C**H** sites are evidently assigned in **Figure 7A** and are located in the molecular structure of Tyr (**Figure 7B**). The experimental fraction curves (black circles) of N**H$_3$** and C**H** are shown in **Figures 7C,D**, respectively. There are two important points to mention for the fraction curve of N**H$_3$** in **Figure 7C**. First, the oscillation is observed and the maximum $\Delta S/S_0$ is reached at τ_{mix} of 0.6 ms. In our previous study, the ^{1}H-^{14}N PM-S-RESPDOR fraction curve shows the maximum at τ_{mix} of ~1.4 ms (Duong et al., 2019), which is about two times slower than that of ^{1}H-^{14}N OT-REDOR fraction curve. This result is in qualitative agreement with the analysis in *Pulse Sequence and the Universal Expression* and **Figure 4**. Second, the fraction curve in **Figure 7C** shows the experimental maximum $\Delta S/S_0$ of 0.55, smaller than the theoretical maximum of 0.67 by the universal curve in **Eq. (3)** for complete inversion. Thus, the fitting parameter f of 0.55/0.67~0.82 is required for the reliable fittings. In addition, the value of 0.82 is close to $f = 1$ in the case of complete inversion, revealing the dominance of the inversion process under the application of ^{14}N OT CW pulse. Once f is determined, the only unknown remaining parameter is b_{1H-14N}. Moreover, according to the fitting strategy in *Pulse Sequence and the Universal Expression*, this fitting parameter f can also be used in **Figure 7D**. This is because f only depends on ^{14}N C_Q, τ_{CW}, and ^{14}N OT nutation frequency (which are the same as fraction curves in **Figures 7C,D** are from the single experiment) and thus should remain the same for other ^{1}H-^{14}N pairs from the same ^{14}N site. The ^{1}H-^{14}N dipolar couplings, thus distances, can be extracted by fitting the scaled universal curves (red solid lines) to the experimental fraction curves (black circles) presented in **Figures 7C,D**. Although the fraction curve in **Figure 7C** shows the oscillation up to τ_{mix} of ~1.8 ms, the fitting by the universal curves is only up to τ_{mix} of ~1.0 ms, owing to the poor agreement between the experimental and universal curves for NH_3 at τ_{mix} > 1.0 ms (**Supplementary Figure S2A**). The deviation is mainly caused by the fact that each crystallite experiences different ^{14}N OT saturation/inversion extent depending on its relative orientation between quadrupolar tensor to the rotor-fixed frame, whereas, for the universal approach, the behaviors of the entire crystallites are considered uniform. The root-mean-square deviation (RMSD) in **Figures 7C,D** was calculated for the best fit of ^{1}H-^{14}N dipolar couplings. It is noted that, for N**H$_3$** (**Figure 7C**), a scaling factor $P_2(\cos(\theta))$ (θ, the angle between H-N and C-N, is 109.5°) is used for the dynamic average of the N-H dipolar coupling due to the threefold rotation. The ^{1}H-^{14}N distances by OT-REDOR are shown in **Table 1** along with those by PM-S-RESPDOR and neutron diffraction (ND). The distances are in good agreement with each other, which demonstrates the feasibility of OT-REDOR for obtaining accurate ^{1}H-^{14}N distances. It is worth noting that the longer distances by ssNMR than those from neutron result from the different vibrational averages of the internuclear distances of the two techniques (Ishii et al., 1997).

N-Acetyl-L-alanine

To further demonstrate the feasibility of OT-REDOR for a system with a larger ^{14}N C_Q, we apply it to AcAla. A similar experimental procedure as described for Tyr was applied, including the 1) determination of ^{14}N OT resonance frequency, 2) optimization of τ_{CW}, and 3) implementation of OT-REDOR. These experiments were all performed at the second ^{14}N OT SSB for the highest S/N. For step 1, the ^{1}H-{^{14}N OT} T-HMQC was performed (**Supplementary Figure S3**). Again, for the efficient OT-REDOR fraction curve, τ_{CW} must be optimized. For step 2, such optimization for N**H** (black squares) at τ_{mix} of 0.19 ms (or three loops of $SR4_1^2$ recoupling blocks) and C**H** (red circles) at τ_{mix} of 0.96 ms (or 15 loops of $SR4_1^2$ recoupling blocks) under identical τ_{CW} range is shown in **Figure 8**. The reason for different τ_{mix} is due to the large difference between ^{1}H-^{14}N distances for these proton sites. The optimum τ_{CW} of 0.192 ms for **^{14}N**H in AcAla is shorter than τ_{CW} of 0.464 ms for **^{14}N**H_3 in Tyr. This result is expected since the ^{14}N site of NH has a larger quadrupolar interaction, thus resulting in a larger ^{14}N OT nutation field and shorter pulse length for efficient saturation/inversion.

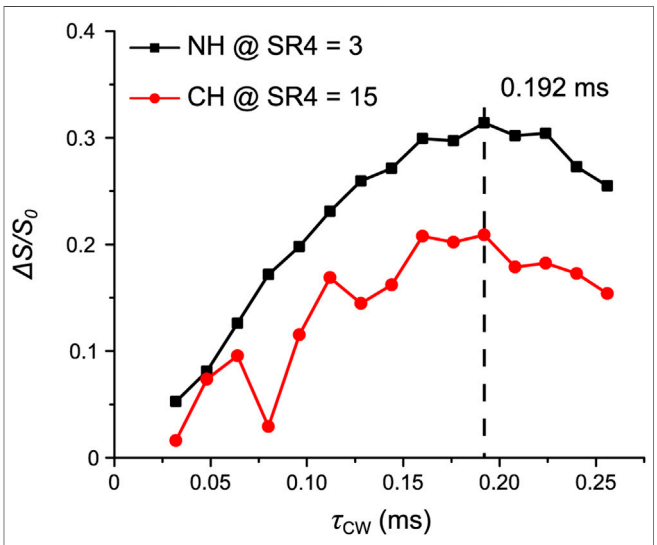


FIGURE 8 | AcAla: the signal fraction $\Delta S/S_0$ of N**H** (black squares) at τ_{mix} of 0.19 ms and C**H** (red circles) at τ_{mix} of 0.96 ms as a function of τ_{CW}. The optimum τ_{CW} is shown and highlighted by the dashed line. Further experimental details are given in the *Experiments* section.

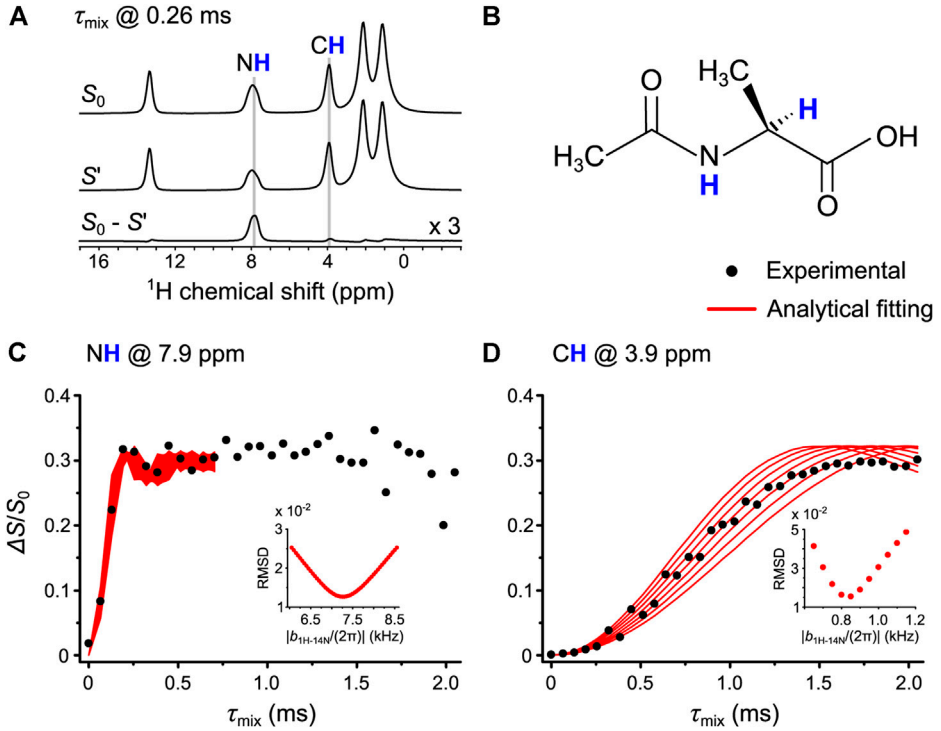


FIGURE 9 | (A) The 1D ^{1}H spin echo (S_0), the dephased (S'), and the difference (S_0–S') spectra of AcAla. These spectra result from the ^{1}H-^{14}N OT-REDOR experiment with τ_{mix} of 0.26 ms. For readability, the intensity of the difference spectrum is triple. **(B)** The molecular structure of AcAla where N**H** and C**H** are located. **(C,D)** The fitting of experimental ^{1}H-^{14}N OT-REDOR fraction curves (black circles) by the universal curves (red lines) for N**H** at 7.9 ppm in **(C)** and C**H** at 3.9 ppm in **(D)**. The fitting parameter f for universal curves is 0.50. The RMSD analyses (inset) were calculated for the best fitting ^{1}H-^{14}N dipolar couplings.

TABLE 2 | ^{1}H-^{14}N distances measured by OT-REDOR, PM-S-RESPDOR, and XRD on AcAla.

	OT-REDOR	PM-S-RESPDOR	XRD
NH	1.07 ± 0.06 Å	1.06 Å	0.78 Å
CH	2.21 ± 0.12 Å	2.11 Å	2.09 Å

For step 3, these parameters were used for the OT-REDOR experiment on AcAla. **Figure 9A** shows the one-dimensional (1D) ^{1}H spin echo (S_0), the dephased (S'), and the difference (S_0–S') spectra. The former two were extracted from the ^{1}H-^{14}N OT-REDOR experiment with τ_{mix} of 0.26 ms, whereas the latter results from the difference of the former two. For readability, the intensity of the difference spectrum is triple, showing the N**H** and C**H** sites which experience the REDOR effect. These two sites are also located in the molecular structure of AcAla (**Figure 9B**). **Figures 9C,D** show the fittings between the universal curves (red lines) and the experimental fraction curves (black circles) for N**H** and C**H**, respectively. For reliable fittings, the fitting parameter f must be known. From **Figure 9C**, the experimental maximum $\Delta S/S_0$ intensity of 0.33 results in f of 0.33/0.67 = 0.50. While Tyr shows the dominance of inversion (f = 0.82), AcAla experiences the saturation of overall magnetization (f = 0.50). The difference may arise from the large frequency linewidth (up to 8.6 kHz in **Supplementary Figure S3**) of ^{14}N OT spectrum of NH of AcAla relative to the weak ^{14}N OT nutation frequency. Since both fraction curves in **Figures 9C,D** were obtained from the single experiment where ^{14}N C_Q, τ_{CW}, and ^{14}N OT nutation frequency are the same, the identical f value can be used in **Figure 9D**. The extracted ^{1}H-^{14}N distances are summarized in **Table 2** along with those from PM-S-RESPDOR and XRD. The distances are in agreement with each other. We note that the deviation between the distance of N–H by XRD and those by OT-REDOR and PM-S-RESPDOR is due to the poor capability of XRD to locate H position, resulting from the limited scattering power of hydrogen and the vibrational effect mentioned in *L-Tyrosine.HCl*. In short, the applicability of OT-REDOR on Tyr and AcAla for obtaining accurate ^{1}H-^{14}N distances has been validated.

L-Alanyl-L-alanine

In the previous section, OT-REDOR experiments have been successfully applied to Tyr and AcAla, both containing a single ^{14}N site. As the main usage of OT-REDOR is for systems where multiple ^{14}N sites are present, here, we apply this sequence to AlaAla. Besides the 1D ^{1}H spin echo (S_0) at the top, **Figure 10A** also shows the two difference (S_0–S') spectra where 1) N**H$_3$** and 2) N**H** sites are saturated/inverted during ^{1}H-^{14}N OT-REDOR experiments with τ_{mix} of 0.51 ms and 0.19 ms, respectively. The N**H**, N**H$_3$**, C**H(1)**, and C**H(2)** sites are unambiguously assigned (**Figure 10A**) and located in the molecular structure of AlaAla (**Figure 10B**). This compound

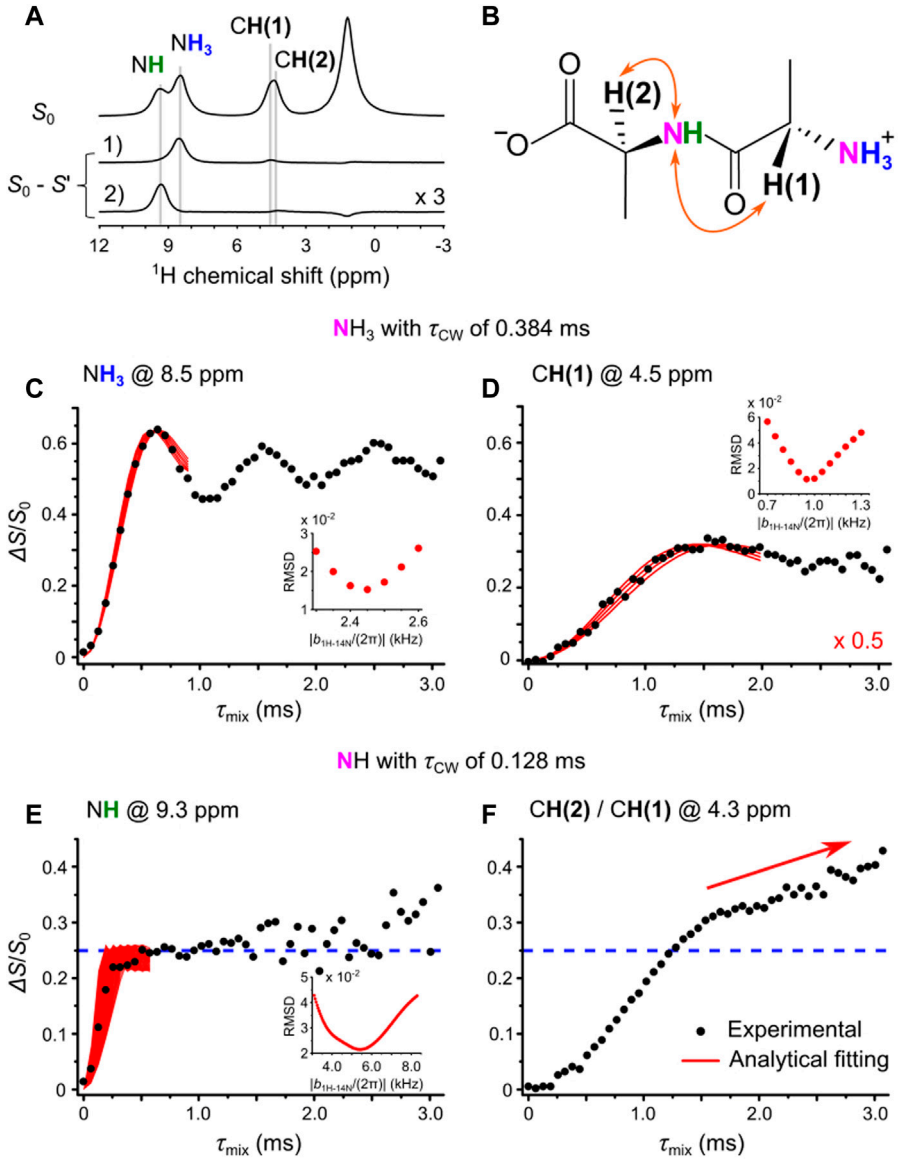


FIGURE 10 | (A) The 1D ^{1}H spin echo and the two difference spectra where 1) N**H$_3$** and 2) N**H** sites are saturated/inverted during ^{1}H-^{14}N OT-REDOR experiments with τ_{mix} of 0.51 ms and 0.19 ms, respectively. For readability, the intensity of spectrum 2) is triple. **(B)** The N**H**, N**H$_3$**, C**H(1)**, and C**H(2)** sites, assigned in **(A)**, are located in the molecular structure of AlaAla. **(C–F)** The fitting of experimental ^{1}H-^{14}N OT-REDOR fraction curves (black circles) by the universal curves (red lines) for N**H$_3$** at 8.5 ppm in **(C)** and C**H(1)** at 4.5 ppm in **(D)** when **^{14}N**H$_3$ OT was saturated/inverted by CW pulse with τ_{CW} of 0.384 ms and N**H** at 9.3 ppm in **(E)** when **^{14}N**H OT was saturated/inverted by CW pulse with τ_{CW} of 0.128 ms. The fitting parameter *f* for universal curves is (**C,D**) 0.94 and (**E**) 0.37. For **(D)**, the universal curves are halved, highlighted by a factor of 0.5 in red. For **(E)**, the maximum signal fraction $\Delta S/S_0$ of 0.25 for N**H** is presented by the horizontal dash line. For **(F)**, the signal fraction $\Delta S/S_0$ of C**H(2)**/C**H(1)** at 4.3 ppm with **^{14}N**H OT irradiation is over this line and is still increasing with τ_{mix} (highlighted by the arrow). The RMSD analyses (inset) were calculated for the best fitting ^{1}H-^{14}N dipolar couplings. The two arrows show the proximity of both C**H(2)**/C**H(1)** to **^{14}N**H site.

consists of two ^{14}N sites of **N**H$_3$ and **N**H; hence, it is similar to the combination of ^{14}N sites of Tyr and AcAla.

Similar experimental procedures were applied. *D*-HMQC experiments were performed again at the second ^{14}N OT SSB (n = -2) for the highest S/N to determine ^{14}N OT frequencies (**Supplementary Figures S4A,B**). Then, τ_{CW} for **^{14}N**H$_3$ and **^{14}N**H of AlaAla were optimized, which were 0.384 and 0.128 ms, respectively (**Supplementary Figures S4C,D**). These optimized values, similar to those of Tyr and AcAla, were used to obtain ^{1}H-^{14}N OT-REDOR fraction curves.

We discuss first the case where **^{14}N**H$_3$ is saturated/inverted by CW pulse with τ_{CW} of 0.384 ms. The experimental fraction curves of N**H$_3$** and C**H(1)** (black circles) are shown in **Figures 10C,D**, respectively. For N**H$_3$**, the experimental curve shows the oscillation and the maximum $\Delta S/S_0$ is 0.63 (see **Figure 10C**). Thus, the fitting parameter f of 0.63/0.67~0.94 for universal

TABLE 3 | ^{1}H-^{14}N distances measured by OT-REDOR and XRD on AlaAla.

	Saturation/inversion on ^{14}N(H_3)		Saturation/inversion on ^{14}N(H)
	NH_3	CH(1)	NH
OT-REDOR	1.05 ± 0.03 Å	2.09 ± 0.05 Å	1.19 ± 0.20 Å
XRD	0.93 Å	1.97 Å	0.96 Å

curves is required for reliable fittings. This indicates that an almost complete inversion of ^{14}N OT is achieved ($f = 1$), similar to the case of NH_3 of Tyr. Another similarity to Tyr is that significant deviation is observed, especially at long τ_{mix} for the full-scale fitting between the fraction curve of **Figure 10C** and the universal curves (see **Supplementary Figure S2B**). For C**H(1)**, the fraction curve reaches the plateau $\Delta S/S_0$ of 0.35 at τ_{mix} of 1.50 ms (**Supplementary Figure S5**). The reason for the lower signal fraction $\Delta S/S_0$ is the overlapping of ^{1}H signals of C**H(1)** and C**H(2)** sites, in which only C**H(1)** is close to ^{14}**N**H_3. For a good match with the experimental fraction curve, we need to halve the universal curves in **Figure 10D**. The reason for using a factor of 0.5 is explained in the **Supplementary Eq. S3**. Without scaling, the obtained ^{1}H-^{14}N distance by NMR is in poor agreement with that reported from XRD (see **Supplementary Figure S5**). The fitting of the scaled universal curves to the ^{1}H-^{14}N OT-REDOR fraction curves gives the ^{1}H-^{14}N distances of 1.05 ± 0.03 Å for NH_3 (after a modulation factor of $P_2(\cos(\theta))$ as described in *L-Tyrosine.HCl* and *Pulse Sequence and the Universal Expression* and 2.09 ± 0.05 Å for CH(1) (after halving the universal curves). Both measured distances are in good agreement with those reported from XRD (see **Table 3**).

We then consider the case where ^{14}**N**H is saturated/inverted by CW pulse with τ_{CW} of 0.128 ms. The fraction curves of N**H** and C**H(2)**/C**H(1)** (black circles) are shown in **Figures 10E,F**, respectively. For N**H** in **Figure 10E**, the experimental curve shows the plateau $\Delta S/S_0$ of 0.25 at τ_{mix} of 0.50 ms and then large fluctuations of $\Delta S/S_0$ at τ_{mix} larger than 1.50 ms. The origin of such fluctuations may be due to t_1 noise from spinning frequency fluctuation (Nishiyama et al., 2020). As the maximum $\Delta S/S_0$ is smaller than the theoretical maximum of 0.67 of the universal curve, a fitting parameter f of 0.25/0.67~0.37 is required. Based on the fittings of the universal curves (red solid lines) to the ^{1}H-^{14}N OT-REDOR fraction curve, the ^{1}H-^{14}N distance is measured to be 1.19 ± 0.20 Å. This value is in excellent agreement with ^{1}H-^{15}N distance by inverse cross-polarization with variable contact (CPVC) (Nishiyama et al., 2016) (**Supplementary Figure S6**) and in close agreement with the distance of 0.96 Å by XRD (see **Table 3**). An advantage of ^{1}H-^{14}N OT-REDOR compared to ^{1}H-^{15}N inverse CPVC is that multiple H-N distances can be simultaneously determined by the former, whereas only directly bonded distance is determined by the latter due to the dipolar truncation effect. Indeed, ^{1}H-^{15}N inverse CPVC only provides the distance of directly bonded H-N for the NH site (see **Supplementary Figure S6**). For C**H(2)**/C**H(1)** fraction curve, its fraction signal $\Delta S/S_0$ is larger than 0.25 and continues to grow at long τ_{mix} (see **Figure 10F**). Although this curve is the combination of two curves because there are two C**H** groups that are close to ^{14}**N**H site and their ^{1}H chemical shifts are overlapped, $\Delta S/S_0$ is larger than 0.25 may result from the intermolecular couplings. Because of this complexity, we did not fit this with the universal curves.

In conclusion, for AlaAla, the ^{1}H-^{14}N OT-REDOR experiment can be used for the accurate measurement of ^{1}H-^{14}N distances for the bonded H–N distances of each nitrogen. However, extracting distances for nonbonded H-N pairs is still difficult, especially when the chemical shifts of these ^{1}H sites are overlapped, as shown in **Figure 10F**. Such problem will be solved by the multidimensional NMR experiments, for example, with an addition of the ^{13}C dimension.

CONCLUSION

In summary, we have presented the feasibility of ^{1}H-^{14}N OT-REDOR with proton detection at fast MAS to extract ^{1}H-^{14}N distances for Tyr, AcAla, and AlaAla. Owing to the selective characteristics of ^{14}N OT spectroscopy, this sequence is useful for systems with multiple ^{14}N sites. Other advantages of ^{14}N OT are the availability of commercial MAS probes, the robustness of misadjustment of the magic angle, and the fast dephasing rate. The final advantage is that it allows probing longer ^{1}H-^{14}N distances better than ^{1}H-^{14}N RESPDOR experiment. For efficient ^{1}H-^{14}N OT-REDOR fraction curve, the ^{14}N OT resonance frequency, in this work, must be determined with *D*- or T-HMQC experiments and the CW pulse length must be optimized. For reliable ^{1}H-^{14}N distances, the fitting parameter f is a prerequisite; otherwise, distances cannot be accurately determined. The knowledge of f value also enables the evaluation of saturation/inversion degree of ^{14}N OT by the CW pulse. For Tyr and AcAla compounds, the extracted distances from OT-REDOR are in good agreement with PM-S-RESPDOR and the diffraction techniques. For AlaAla, the extracted ^{1}H-^{14}N distances from directly bonded N–H well agree with those reported from XRD and ^{1}H-^{15}N inverse CPVC. However, this is not the case for nonbonded N–H pairs since distance deviations from those reported by XRD are observed. The reason for such deviation is the overlapping of ^{1}H signals. This issue can be overcome by performing multidimensional NMR experiments. In conclusion, we believe that the ^{1}H-^{14}N OT-REDOR has the potential of selectively measuring ^{1}H-^{14}N distances on systems containing multiple ^{14}N sites, giving deep insights into structural studies of biological, chemical, and pharmaceutical compounds. It is worth noting that ^{14}N selective saturation can also be achieved in the manner of DANTE. It is promising to perform ^{1}H-^{14}N DANTE-RESPDOR experiments in future studies.

EXPERIMENTS

L-tyrosine.HCl (Tyr), N-acetyl-L-alanine (AcAla), and L-alanyl-L-alanine were purchased from Sigma-Aldrich and used as received. The samples were separately packed into 1.0 mm zirconia rotors and then inserted into 1 mm ^{1}H/X

double-resonance probe. The rotors were spun at a MAS frequency of 62.5 kHz, except for ^{1}H-^{15}N inverse CPVC at 70 kHz.

All NMR experiments were recorded at a room temperature of 25 °C on JNM-ECZ600R (JEOL RESONANCE Inc.) at 14.1 T solid-state NMR spectrometers. The ^{1}H and ^{14}N OT Larmor frequencies are 600.0 and 86.8 MHz, respectively. For the highest S/N, the ^{14}N OT frequency was set at the second SSB (n = −2). The ^{14}N and ^{14}N OT shifts are referenced to CH_3NO_2, whose ^{14}N and ^{14}N OT shifts are equal to 0 ppm or 0 kHz. The ^{1}H rf-field was 328 kHz for π/2 and π pulses and 140 kHz for the $SR4_1^2$ recoupling sequence. The ^{14}N OT rf-field was 120 kHz.

For Tyr, the 2D ^{1}H-{^{14}N} *D*-HMQC spectrum in **Figure 5** was recorded using the sequence shown in **Supplementary Figure S1A** with 8 scans, 32 t_1 points, and rotor-synchronized t_1 increment of 16.0 μs. τ_p, τ_{mix}, and recycling delay (RD) were 200 μs, 512 μs, and 4 s, respectively. The experimental time was about 0.6 h. The States-TPPI method was employed for quadrature detection along the indirect dimension. For **Figure 6**, τ_{CW} was optimized within the range from 16 μs to 656 μs with a step of 32 μs; the ^{14}N OT frequency was −1.73 ppt (parts per thousand), the τ_{mix} was fixed at 1152 μs, the number of scans (NS) was 18, and RD was 5.0 s. The experimental time was 1.1 h. For **Figure 7**, the ^{1}H-^{14}N OT-REDOR was performed at τ_{CW} of 464 μs, ^{14}N OT frequency of −1.73 ppt, NS of 72, RD of 6.5 s, and τ_{mix} from 0 to 3456 μs with a step of 64 μs. The experimental time was 14.3 h.

For AcAla, the τ_{CW} optimization in **Figure 8** was implemented within the range from 32 μs to 256 μs with a step of 16 μs; the ^{14}N OT frequency was −1.267 ppt, NS was 18 and RD was 10.0 s, and τ_{mix} was fixed at 192 μs for N**H** at 7.9 ppm and 960 μs for C**H** at 3.9 ppm. The experimental times for both experiments were 1.5 h. For **Figure 9**, the ^{1}H-^{14}N OT-REDOR was performed at τ_{CW} of 192 μs, ^{14}N OT frequency of −1.267 ppt, NS of 108, RD of 10 s, and τ_{mix} from 0 to 2048 μs with a step of 64 μs. The experimental time was 19.8 h.

For AlaAla, the ^{1}H-^{14}N OT-REDOR experiments in **Figure 10** were performed at NS of 144, RD of 2.5 s, τ_{mix} from 0 to 3072 μs with a step of 64 μs, and τ_{CW}, ^{14}N OT frequencies were of 384 μs, −1.72 ppt and 128 μs, −1.26 ppt for ^{14}**N**H_3 and ^{14}**N**H, respectively. The experimental times for both experiments were 9.8 h.

NMR data are available upon request.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**; further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

ND was responsible for methodology, NMR measurements, and analysis; ZG was responsible for conceptualization, and methodology; YN was responsible for conceptualization, methodology, and supervision. All authors were responsible for writing the manuscript.

FUNDING

This work was supported by JSPS KAKENHI Grant Number 20K05483 to YN. ZG acknowledges the support from the US National Science Foundation through NSF/DMR-1644779 and the National High Magnetic Field Laboratory at Florida.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.645347/full#supplementary-material.

REFERENCES

Bloom, M., and LeGros, M. A. (1986). Direct detection of two-quantum coherence. *Can. J. Phys.* 64, 1522–1528. doi:10.1139/p86-271

Brinkmann, A, and Kentgens, A. P. (2006). Proton-selective 17O-1H distance measurements in fast magic-angle-spinning solid-state NMR spectroscopy for the determination of hydrogen bond lengths. *J. Am. Chem. Soc.* 128, 14758–14759. doi:10.1021/ja065415k

Brown, S. P. (2014). Nitrogen-proton correlation experiments of organic solids at natural isotopic abundance. *eMagRes.* 3, 243–254. doi:10.1002/9780470034590.emrstm1323

Carnevale, D., Ji, X., and Bodenhausen, G. (2017). Double cross polarization for the indirect detection of nitrogen-14 nuclei in magic angle spinning NMR spectroscopy. *J. Chem. Phys.* 147, 184201. doi:10.1063/1.5000689

Cavadini, S. (2010). Indirect detection of nitrogen-14 in solid-state NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* 56, 46–77. doi:10.1016/j.pnmrs.2009.08.001

Cavadini, S., Antonijevic, S., Lupulescu, A., and Bodenhausen, G. (2006). Indirect detection of nitrogen-14 in solids via protons by nuclear magnetic resonance spectroscopy. *J. Magn. Reson.* 182, 168–172. doi:10.1016/j.jmr.2006.06.003

Chen, L., Lu, X., Wang, Q., Lafon, O., Trébosc, J., Deng, F., et al. (2010a). Distance measurement between a spin-1/2 and a half-integer quadrupolar nuclei by solid-state NMR using exact analytical expressions. *J. Magn. Reson.* 206, 269–273. doi:10.1016/j.jmr.2010.07.009

Chen, L., Wang, Q., Hu, B., Lafon, O., Trébosc, J., Deng, F., et al. (2010b). Measurement of hetero-nuclear distances using a symmetry-based pulse sequence in solid-state NMR. *Phys. Chem. Chem. Phys.* 12, 9395–9405. doi:10.1039/b926546e

Chevelkov, V., Fink, U., and Reif, B. (2009). Accurate determination of order parameters from 1H,15N dipolar couplings in MAS solid-state NMR experiments. *J. Am. Chem. Soc.* 131, 14018–14022. doi:10.1021/ja902649u

Concistré, M., Kuprov, I., Haies, I. M., Williams, P. T. F., and Carravetta, M. (2018). ^{14}N overtone NMR under MAS: signal enhancement using cross-polarization methods. *J. Magn. Reson.* 298, 1–5. doi:10.1016/j.jmr.2018.10.017

Duong, N. T., Rossi, F., Makrinich, M., Goldbourt, A., Chierotti, M. R., Gobetto, R., et al. (2019). Accurate 1H-14N distance measurements by phase-modulated RESPDOR at ultra-fast MAS. *J. Magn. Reson.* 308, 106559. doi:10.1016/j.jmr.2019.07.046

Fu, R. (2003). Measurement of 15N-1H bond lengths by rotational-echo double-resonance NMR spectroscopy. *Chem. Phys. Lett.* 376, 62–67. doi:10.1016/S0009-2614(03)00951-5

Gan, Z. (2006). Measuring multiple carbon-nitrogen distances in natural abundant solids using R-RESPDOR NMR. *Chem Commun (Camb).* 7, 4712–4714. doi:10.1039/b611447d

Gan, Z., Amoureux, J. P., and Trébosc, J. (2007). Proton-detected ^{14}N MAS NMR using homonuclear decoupled rotary resonance. *Chem. Phys. Lett.* 435, 163–169. doi:10.1016/j.cplett.2006.12.066

Gan, Z., Hung, I., Nishiyama, Y., Amoureux, J. P., Lafon, O., Nagashima, H., et al. (2018). 14N overtone nuclear magnetic resonance of rotating solids. *J. Chem. Phys.* 149, 064201. doi:10.1063/1.5044653

Gullion, T. (2007). Rotational-Echo, Double-Resonance NMR. *MRI.* 1, 713–718. doi:10.1007/1-4020-3910-7_89

Gullion, T., and Schaefer, J. (1989). Rotational-echo double-resonance NMR. *J. Magn. Reson. Imaging.* 81, 196–200. doi:10.1016/0022-2364(89)90280-1

Guzmán-Afonso, C., Hong, Y. L., Colaux, H., Iijima, H., Saitow, A., Fukumura, T., et al. (2019). Understanding hydrogen-bonding structures of molecular crystals via electron and NMR nanocrystallography. *Nat. Commun.* 10, 3537. doi:10.1038/s41467-019-11469-2

Haies, IM, Jarvis, JA, Bentley, H, Heinmaa, I, Kuprov, I, Williamson, PT, et al. (2015a). (14)N overtone NMR under MAS: signal enhancement using symmetry-based sequences and novel simulation strategies. *Phys. Chem. Chem. Phys.* 17, 6577–87. doi:10.1039/C4CP03994G

Haies, I. M., Jarvis, J. A., Brown, L. J., Kuprov, I., Williamson, P. T., and Carravetta, M. (2015b). (14)N overtone transition in double rotation solid-state NMR. *Phys. Chem. Chem. Phys.* 17, 23748–23753. doi:10.1039/C5CP03266K

Hohwy, M., Jaroniec, C. P., Reif, B., Rienstra, C. M., and Griffin, R. G. (2000). Local structure and relaxation in solid-state NMR: accurate measurement of amide N–H bond lengths and H–N–H bond angles. *J. Am. Chem. Soc.* 122, 3218–3219. doi:10.1021/ja9913737

Hou, G., Byeon, I. J. L., Ahn, J., Gronenborn, A. M., and Polenova, T. (2011). ^{1}H-^{13}C/^{1}H-^{15}N heteronuclear dipolar recoupling by R-symmetry sequences under fast magic angle spinning for dynamics analysis of biological and organic solids. *J. Am. Chem. Soc.* 133, 18646–18655. doi:10.1021/ja203771a

Hou, G., Lu, X., Vega, A. J., and Polenova, T. (2014). Accurate measurement of heteronuclear dipolar couplings by phase-alternating R-symmetry (PARS) sequences in magic angle spinning NMR spectroscopy. *J. Chem. Phys.* 141, 104202. doi:10.1063/1.4894226

Hung, I., Gor'kov, P., and Gan, Z. (2019). Efficient and sideband-free ^{1}H-detected ^{14}N magic-angle spinning NMR. *J. Chem. Phys.* 151, 154202. doi:10.1063/1.5126599

Hung, I., Gor'kov, P., and Gan, Z. (2020). Using the heteronuclear Bloch-Siegert shift of protons for B1 calibration of insensitive nuclei not present in the sample. *J. Magn. Reson.* 310, 106636. doi:10.1016/j.jmr.2019.106636

Ishii, Y., Terao, T., and Hayashi, S. (1997). Theory and simulation of vibrational effects on structural measurements by solid-state nuclear magnetic resonance. *J. Chem. Phys.* 107, 2760–2774. doi:10.1063/1.474633

Jarvis, J. A., Concistre, M., Haies, I. M., Bounds, R. W., Kuprov, I., Carravetta, M., et al. (2019). Quantitative analysis of 14N quadrupolar coupling using 1H detected 14N solid-state NMR. *Phys. Chem. Chem. Phys.* 21, 5941–5949. doi:10.1039/c8cp06276e

Jayanthi, S., and Ramanathan, K. V. (2011). Excitation and correlation of ^{14}N overtone transitions and measurement of heteronuclear dipolar coupling using DAPT. *Chem. Phys. Lett.* 502, 121–125. doi:doi:10.1016/j.cplett.2010.12.007

Lu, X., Lafon, O., Trébosc, J., and Amoureux, J. P. (2011). Detailed analysis of the S-RESPDOR solid-state NMR method for inter-nuclear distance measurement between spin-1/2 and quadrupolar nuclei. *J. Magn. Reson.* 215, 34–49. doi:10.1016/j.jmr.2011.12.009

Lu, X., Trébosc, J., Lafon, O., Carnevale, D., Ulzega, S., Bodenhausen, G., et al. (2013). Broadband excitation in solid-state NMR using interleaved DANTE pulse trains with N pulses per rotor period. *J. Magn. Reson.* 236, 105–116. doi:10.1016/j.jmr.2013.09.003

Makrinich, M., Gupta, R., Polenova, T., and Goldbourt, A (2017). Saturation capability of short phase modulated pulses facilitates the measurement of longitudinal relaxation times of quadrupolar nuclei. *Solid State Nucl. Magn. Reson.* 84, 196–203. doi:10.1016/j.ssnmr.2017.04.003

Makrinich, M., Nimerovsky, E., and Goldbourt, A. (2018). Pushing the limit of NMR-based distance measurements – retrieving dipolar couplings to spins with extensively large quadrupolar frequencies. *Solid State Nucl. Magn. Reson.* 92, 19–24. doi:10.1016/j.ssnmr.2018.04.001

Nimerovsky, E., Gupta, R., Yehl, J., Li, M., Polenova, T., and Goldbourt, A. (2014). Phase-modulated LA-REDOR: A robust, accurate and efficient solid-state NMR technique for distance measurements between a spin-1/2 and a quadrupole spin. *J. Magn. Reson.* 244, 107–113. doi:10.1016/j.jmr.2014.03.003

Nimerovsky, E, Makrinich, M, and Goldbourt, A (2017). Analysis of large-anisotropy-spin recoupling pulses for distance measurement under magic-angle spinning NMR shows the superiority and robustness of a phase modulated saturation pulse. *J. Chem. Phys.* 146, 124202. doi:10.1063/1.4978472

Nishiyama, Y. (2016). Fast magic-angle sample spinning solid-state NMR at 60-100kHz for natural abundance samples. *Solid State Nucl. Magn. Reson.* 78, 24–36. doi:10.1016/j.ssnmr.2016.06.002

Nishiyama, Y., Agarwal, V., and Zhang, R. (2020). t_1-Noise Suppression by γ-Free Recoupling Sequences in Solid-State NMR for Structural Characterization of Fully Protonated Molecules at Fast MAS. *J. Phys. Chem. C.* 128, 26332–26343. doi:10.1021/acs.jpcc.0c08828

Nishiyama, Y, Endo, Y, Nemoto, T, Utsumi, H, Yamauchi, K, Hioka, K, et al. (2011). Very fast magic angle spinning (1)H-(14)N 2D solid-state NMR: sub-micro-liter sample data collection in a few minutes. *J. Magn. Reson.* 208, 44–48. doi:10.1016/j.jmr.2010.10.001

Nishiyama, Y., Malon, M., Gan, Z., Endo, Y., and Nemoto, T. (2013). Proton-nitrogen-14 overtone two-dimensional correlation NMR spectroscopy of solid-sample at very fast magic angle sample spinning. *J. Magn. Reson.* 230, 160–164. doi:10.1016/j.jmr.2013.02.015

Nishiyama, Y., Malon, M., Potrzebowski, M. J., Paluch, P., and Amoureux, J. P. (2016). Accurate NMR determination of C-H or N-H distances for unlabeled molecules. *Solid State Nucl. Magn. Reson.* 73, 15–21. doi:10.1016/j.ssnmr.2015.06.005

O'Dell, L. A., and Brinkmann, A. (2013). ^{14}N overtone NMR spectra under magic angle spinning: Experiments and numerically exact simulations. *J. Chem. Phys.* 138, 064201. doi:10.1063/1.4775592

O'Dell, L. A., He, R., and Pandohee, J. (2013). Identifying H–N proximities in solid-state NMR using ^{14}N overtone irradiation under fast MAS. *CrystEngComm.* 15, 8657–8667. doi:10.1039/c3ce40967h

O'Dell, L. A., and Ratcliffe, C. I. (2011). ^{14}N magic angle spinning overtone NMR spectra. *J. Chem. Phys.* 514, 168–173. doi:10.1016/j.cplett.2011.08.030

Paluch, P., Pawlak, T., Amoureux, J. P., and Potrzebowski, M. J. (2013). Simple and accurate determination of X-H distances under ultra-fast MAS NMR. *J. Magn. Reson.* 233, 56–63. doi:10.1016/j.jmr.2013.05.005

Pandey, M. K., Amoureux, J. P., Asakura, T., and Nishiyama, Y. (2016). Sensitivity enhanced (14)N/(14)N correlations to probe inter-beta-sheet interactions using fast magic angle spinning solid-state NMR in biological solids. *Phys. Chem. Chem. Phys.* 18, 22583–22589. doi:10.1039/c6cp03848d

Pandey, M. K., and Nishiyama, Y. (2015). Proton-detected 3D ^{14}N/^{14}N/^{1}H isotropic shift correlation experiment mediated through ^{1}H-^{1}H RFDR mixing on a natural abundant sample under ultrafast MAS. *J. Magn. Reson.* 258, 96–101. doi:10.1016/j.jmr.2015.06.012

Pandey, M. K., and Nishiyama, Y. (2018). A one-dimensional solid-state NMR approach for 14NH/14NH overtone correlation through 1H/1H mixing under fast MAS. *Phys. Chem. Chem. Phys.* 20, 25849–25853. doi:10.1039/c8cp05000g

Pourpoint, F., Trébosc, J., Bonhomme, C., Durupthy, O., Steunou, N., Lafon, O., et al. (2014). Quantitative analysis of the proximities of OH ligands and vanadium sites in a polyoxovanadate cluster using frequency-selective ^{1}H-^{51}V solid-state NMR spectroscopy. *J. Phys. Chem. C.* 118, 18580–18588. doi:10.1021/jp505194q

Pruski, M., Bailly, A., Lang, D. P., Amoureux, J. P., and Fernandez, C. (1999). Studies of heteronuclear dipolar interactions between spin-1/2 and quadrupolar nuclei by using REDOR during multiple quantum evolution. *Chem. Phys. Lett.* 307, 35–40. doi:10.1016/S0009-2614(99)00490-X

Rankin, A. G. M., Trébosc, J., Paluch, P., Lafon, O., and Amoureux, J. P. (2019). Evaluation of excitation schemes for indirect detection of 14N via solid-state HMQC NMR experiments. *J Magn Reson.* 303, 28–41. doi:10.1016/j.jmr.2019.04.004

Schanda, P., Meier, B. H., and Ernst, M. (2011). Accurate measurement of one-bond H-X heteronuclear dipolar couplings in MAS solid-state NMR. *J. Magn. Reson.* 210, 246–259. doi:10.1016/j.jmr.2011.03.015

Schnell, I., and Saalwächter, K. (2002). (15)N-(1)H bond length determination in natural abundance by inverse detection in fast-MAS solid-state NMR spectroscopy. *J. Am. Chem. Soc.* 124, 10938–10939. doi:10.1021/ja026657x

Shen, M., Chen, Q., and Hu, B. (2017). Composite pulses in directly and indirectly detected 14N MAS overtone spectroscopy. *Solid State Nucl. Magn. Reson.* 84, 132–136. doi:10.1016/j.ssnmr.2017.01.007

Shen, M., Trébosc, J., O'Dell, L. A., Lafon, O., Pourpoint, F., Hu, B., et al. (2015). Comparison of various NMR methods for the indirect detection of nitrogen-14 nuclei via protons in solids. *J. Magn. Reson.* 258, 86–95. doi:10.1016/j.jmr.2015.06.008

Tycko, R., and Opella, S. J. (1987). Overtone NMR spectroscopy. *J. Chem. Phys.* 86, 1761. doi:10.1063/1.452176

Vitzthum, V., Caporini, M. A., Ulzega, S., and Bodenhausen, G. (2011). Broadband excitation and indirect detection of nitrogen-14 in rotating solids using Delays Alternating with Nutation (DANTE). *J. Magn. Reson.* 212, 234–239. doi:10.1016/j.jmr.2011.06.013

Vitzthum, V., Caporini, M. A., Ulzega, S., Trébosc, J., Lafon, O., Amoureux, J. P., et al. (2012). Uniform broadband excitation of crystallites in rotating solids using interleaved sequences of delays alternating with nutation. *J. Magn. Reson.* 223, 228–236. doi:10.1016/j.jmr.2012.05.024

Wijesekara, A. V., Venkatesh, A., Lampkin, B. J., VanVeller, B., Lubach, J. W., Nagapudi, K., et al. (2020). Fast acquisition of proton-detected HETCOR solid-state NMR spectra of quadrupolar nuclei and rapid measurement of NH bond lengths by frequency selective HMQC and RESPDOR pulse sequences. *Chemist.* 26, 7881–7888. doi:10.1002/chem.202000390

Zhao, L., Hanrahan, M. P., Chakravarty, P., Dipasquale, A. G., Sirois, L. E., Nagapudi, K., et al. (2018). Characterization of pharmaceutical cocrystals and salts by dynamic nuclear polarization-enhanced solid-state NMR spectroscopy. *Cryst. Growth Des.* 18, 2588–2601. doi:10.1021/acs.cgd.8b00203

Zhao, X., Sudmeier, J. L., Bachovchin, W. W., and Levitt, M. H. (2001). Measurement of NH bond lengths by fast magic-angle spinning solid-state NMR spectroscopy: A new method for the quantification of hydrogen bonds. *J. Am. Chem. Soc.* 123, 11097–11098. doi:10.1021/ja016328p

Conflict of Interest: YN is employed by JEOL RESONANCE Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a past co-authorship with one of the authors YN.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 27 July 2021
doi: 10.3389/fmolb.2021.705572

Deuteron Chemical Exchange Saturation Transfer for the Detection of Slow Motions in Rotating Solids

Liliya Vugmeyster[1], Dmitry Ostrovsky[2], Alexander Greenwood[3] and Riqiang Fu[4]*

[1]Department of Chemistry, University of Colorado Denver, Denver, CO, United States, [2]Department of Mathematics, University of Colorado Denver, Denver, CO, United States, [3]Department of Chemistry, University of Cincinnati, Cincinnati, OH, United States, [4]National High Field Magnetic Laboratory, Tallahassee, FL, United States

We utilized the ^{2}H Chemical Exchange Saturation Transfer (CEST) technique under magic angle spinning (MAS) conditions to demonstrate the feasibility of the method for studies of slow motions in the solid state. For the quadrupolar anisotropic interaction, the essence of CEST is to scan the saturation pattern over a range of offsets corresponding to the entire spectral region(s) for all conformational states involved, which translates into a range of −60–+ 60 kHz for methyl groups. Rotary resonances occur when the offsets are at half-and full-integer of the MAS rates. The choice of the optimal MAS rate is governed by the condition to reduce the number of rotary resonances in the CEST profile patterns and retain a sufficiently large quadrupolar interaction active under MAS to maintain sensitivity to motions. As examples, we applied this technique to a well-known model compound dimethyl-sulfone (DMS) as well as amyloid-β fibrils selectively deuterated at a single methyl group of A2 belonging to the disordered domain. It is demonstrated that the obtained exchange rate between the two rotameric states of DMS at elevated temperatures fell within known ranges and the fitted model parameters for the fibrils agree well with the previously obtained value using static ^{2}H NMR techniques. Additionally, for the fibrils we have observed characteristic broadening of rotary resonances in the presence of conformational exchange, which provides implications for model selection and refinement. This work sets the stage for future potential extensions of the ^{2}H CEST under MAS technique to multiple-labeled samples in small molecules and proteins.

Keywords: CEST (chemical exchange dependent saturation transfer), deuterium, solids-state NMR, biomolecular dynamics, conformational exchange

OPEN ACCESS

Edited by:
Józef Romuald Lewandowski, University of Warwick, United Kingdom

Reviewed by:
Rachel Wagner Martin, University of California, United States
Fabien Ferrage, UMR7203 Laboratoire des Biomolécules, France
Petra Rovo, Ludwig Maximilian University of Munich, Germany

****Correspondence:***
Liliya Vugmeyster
liliya.vugmeyster@ucdenver.edu

Specialty section:
This article was submitted to Structural Biology, a section of the journal Frontiers in Molecular Biosciences

Received: *05 May 2021*
Accepted: *12 July 2021*
Published: *27 July 2021*

Citation:
Vugmeyster L, Ostrovsky D, Greenwood A and Fu R (2021) Deuteron Chemical Exchange Saturation Transfer for the Detection of Slow Motions in Rotating Solids. Front. Mol. Biosci. 8:705572. doi: 10.3389/fmolb.2021.705572

INTRODUCTION

Chemical Exchange Saturation Transfer (CEST) experiments provide insights into the molecular dynamics in solution and solid-state NMR studies (Siemer et al., 2010; Bouvignies and Kay, 2012; Vallurupalli et al., 2012; Palmer, 2014; Rovó and Linser, 2018; Palmer and Koss, 2019). They employ weak RF fields for the saturation of selected frequencies as a function of resonance offsets. In most cases, the fluctuations of the isotropic chemical shift interaction is probed when the conformational states have inequivalent chemical shifts. However, anisotropic interactions can also be targeted in the solid state. These measurements are expected to be useful for probing molecular dynamics in a variety of biological systems, including protein fibrils, aggregates and microcrystals. They can elucidate the presence of minor conformational states exchanging with the major state at a slow

timescale with the rate constant in the $5{\cdot}10^2$–$5{\cdot}10^6$ s^{-1} range, with the highest sensitivity around $1{\cdot}10^4$–$5{\cdot}10^5$ s^{-1}. Local motional modes of protein side chains, such as rotameric exchange of methyl-bearing side chains, as well as aromatic ring flips can also be probed with the use of this technique. In addition, backbone motions of C_α deuterons can be elucidated for mobile sites such as loop regions.

We recently demonstrated the effectiveness of the technique for fluctuations of the anisotropic quadrupolar tensor of ^{2}H nuclei under static conditions in the solid state (Vugmeyster et al., 2020). The goal of this work is to expand the methodology for magic angle spinning (MAS) conditions. MAS has proven to be indispensable for dynamic studies of many biomolecular samples with multiple labels (Krushelnitsky et al., 2014; van der Wel, 2017; Rovó, 2020). In this work, we use a single-labeled sample to demonstrate the effectiveness of the ^{2}H CEST experiment, which encourages follow-up studies employing polarization transfer approaches to achieve site-specific resolution (Grey et al., 1993; Bjerring et al., 2012; Akbey et al., 2014; Jain et al., 2014; Matlahov and van der Wel, 2018).

In particular, we apply the experiment to the model compound dimethyl-sulfone (DMS) deuterated at its two methyl groups that undergo 2-site rotameric exchange (Frydman et al., 1994; Brown et al., 1996; Gerardy-Montouillout et al., 1996; Favre et al., 1998; Quinn and McDermott, 2012) and to amyloid-β fibrils ($A\beta_{1\text{-}40}$) with the deuterium label at a mobile methyl group of the A2 residue belonging to the disordered N-terminal domain, for which we previously determined the dynamics using static ^{2}H solid-state NMR techniques (Au et al., 2019; Vugmeyster et al., 2019; Vugmeyster et al., 2020). The experimental work is complemented with theoretical considerations using the Liouvillian formalism (Bain and Berno, 2011) and insights into the main features of the CEST profiles resulting from simple 2-site exchange simulations. Our combined experimental and theoretical/modeling analysis allows us to outline consideration for optimization of the technique and define the ranges of its sensitivity to motions.

EXPERIMENTAL

Materials

DMS-D_6 and hexamethyl-benzene-D_{18} were purchased from Cambridge Isotope Laboratories, Inc. (MA) and packed as a powder into rotors. The $A\beta_{1\text{-}40}$ fibrils labeled at the A2-CD_3 site were prepared as previously described in the 3-fold symmetric toxic polymorph (Au et al., 2019; Vugmeyster et al., 2020). The monomeric sequence of the $A\beta_{1\text{-}40}$ peptide is D(**A-CD_3**)
EFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV.
The lyophilized powder was hydrated to 200% by weight with deuterium-depleted water using direct pipetting and equilibrating at room temperature for 5 days. The hydrated sample was then packed into a 2.5 mm rotor.

NMR Spectroscopy

The measurements for 10 and 25 kHz MAS frequency were performed at 17.6 T Bruker Avance I spectrometer equipped with a Bruker 2.5 mm HXY probe. The measurements at 60 kHz MAS were performed at 14.1 T Bruker neo console spectrometer equipped with a Bruker 1.3 mm HXY probe. The high-power 90° RF pulses corresponded to 2 μs The number of scans for the acquisition was between 32 and 64 for DMS and between 2048 and 3072 for the protein sample. The inter-scan delay was set between 0.5 and 2 s. Because one potential source of systematic error in the ^{2}H CEST intensities is probe detuning, data collection is optimized when the order of the offsets is randomized.

DMS longitudinal relaxation times (T_1) are very sensitive to temperatures in around 40–85°C range, and thus we have used it as an internal calibrations standard to obtain the actual temperature in the samples (Vugmeyster and Ostrovsky, 2019). For the 60 kHz MAS rate, the effect of MAS on T_1 for a given methyl 3-site jump rate was simulated. We also confirmed that the weak-amplitude RF field does not contribute to heating with the chosen relaxation delay value. The longitudinal relaxation times were measured using the inversion recovery experiment, which included a heat compensation block to match the temperature conditions of the sample in the CEST measurements.

Modeling

The simulations were performed on a cluster comprising six x86_64 computer nodes**.** Each node had 16 Intel Xenon Silver dual core CPUs and 16 GB of memory. The procedures closely followed those developed for static conditions in prior work (Vugmeyster et al., 2020). Here, we focus on the details pertinent to MAS conditions.

The evolution of the coherences under MAS was modeled by the direct numerical integration of the Liouville–von Neumann equation (**Supplementary Equation. S1**; all the notations used are elaborated in the Theory section of the **Supplementary Material SI1**). The coherent time-dependent values of the quadrupolar frequency $\omega_Q(t)$ for each site are given by **Eq. 4**. The numerical integration for each saturation time delay was separated into two blocks. The first block comprised the calculation of the evolution matrix for a single MAS rotation $T\exp\left(\int_0^{2\pi/\omega_{MAS}} L dt\right)$, where T stands for the time-ordered exponential function and L is the Liouvillian operator of **Supplementary Equation. S1**. This integration was performed by numerical quadrature with 20 time steps along a single MAS rotation period $2\pi/\omega_{MAS}$. For an individual step, the exponentiation was conducted with fixed $\omega_Q(t)$ values using the internal MATLAB function (Higham, 2005; Al-Mohy and Higham, 2010). We did not use the approximation involving separate integration steps due to the coherent evolution and exchange processes (Saalwächter and Fischbach, 2002). The sufficiency of 20 time steps was confirmed by comparing the results with selected trials with 100 steps. The high consistency of the results holds down to values of $\omega_{MAS}/2\pi$ as low as 1 kHz. The second part of the calculation of the saturation period evolution involved taking the appropriate powers of the evolution matrix for a single MAS rotation as well as the additional multiplicative factor involving the fractional part of the rotation calculated in a similar manner. The equilibrium component for the Zeeman

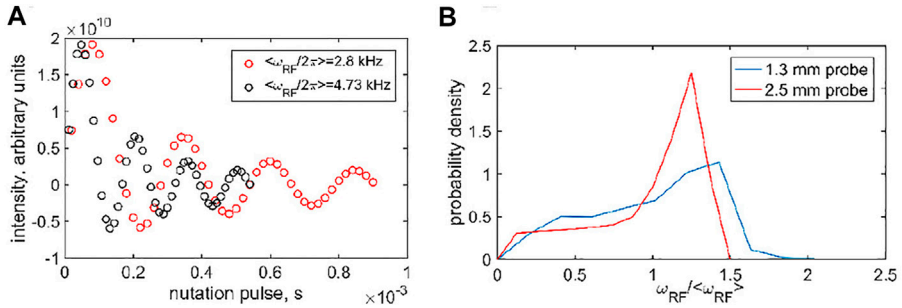


FIGURE 1 | (A) Examples of the ^{2}H nutation measurements used to determine RF inhomogeneity, shown for the 1.3 mm Bruker HXY probe using the DMS-D_6 central band intensity at the 60 kHz MAS rate at 14.1 T. **(B)** Resulting inhomogeneity profiles, obtained as the Fourier analysis of the nutation data, as a function of $\omega_{RF}/<\omega_{RF}>$ obtained for a $<\omega_{RF}>$ range of 2.8–20 kHz for the 1.3 mm Bruker HX probe using DMS-D_6 (at 14.1 T) and 5–31 kHz for the 2.5 mm Bruker HXY probe using HMB-D_{18} (at 17.6 T).

order coherence was introduced phenomenologically as an additional term in the density matrix similar to the Bloch–McConnell treatment of the z-component of magnetization (Vallurupalli et al., 2012).

The detection block was performed starting with the approach described for the static case. As usual, the value of S_z for each site was rotated onto the transverse plane and the simulation of the evolution during the acquisition period followed a similar outline, but involved only the transverse coherences. The time step of calculations was selected as the smaller of 1/20 of the MAS period and FID dwell time.

For the fibrils, longitudinal relaxation was taken into account phenomenologically (with T_1 = 50 ms) by including an additional term in the Liouvillian evolution matrix, which was identical for all eight coherences. This approach was tested for DMS for which the inclusion of the 3-site jumps mode explicitly yielded the same results. To model the effects of the RF inhomogeneity on the CEST profiles, we included five values of $\omega_{RF}/2\pi$ in the ±0.5 kHz range from its central value and averaged the simulated profiles.

The RF inhomogeneity profiles (**Figure 1B**) were discretized by selecting a grid of either six points (for resonance offset values outside the −2–2 kHz region) or 30 points (in the −2–2 kHz region). A larger number of points for the central region was needed due to the enhanced coherent oscillations. The six grid points of the RF inhomogeneity profiles corresponded to the RF field values at 0.25, 0.5, 0.75, 1, 1.25, and 1.5 multiples of the average frequency with the respective relative weights (0.084, 0.143, 0.126, 0.176, 0.285, 0.187 and 0.065, 0.097, 0.103, 0.236, 0.499, 0) for the 1.3 and 2.5 mm probes, respectively. The 30-point grids were obtained by the interpolation of the six-point grids.

RESULTS AND DISCUSSION

Details on the Systems and Known Motional Models

DMS has been widely used as a model system for solid-state NMR technique development and, in particular, for deuteron NMR. Its methyl group undergoes a 2-site rotameric exchange with an angle of rotation of 180° around the C_2 axis of the molecule (Frydman et al., 1994; Brown et al., 1996; Gerardy-Montouillout et al., 1996; Favre et al., 1998; Quinn and McDermott, 2012). These motions are the most pronounced above around 45°C. We have previously utilized this system to develop a deuteron CEST measurement under static conditions and extend it in this work to MAS conditions (Vugmeyster and Ostrovsky, 2019). The ^{2}H spectra under static and MAS conditions (at spinning rates of 10, 25, and 60 kHz) are shown in **Figure 2**. Although there are some spectral distortions due to the motions in the regime in which the flip rate is of the order of the effective value of the quadrupolar constant, the overall width of the pattern remains largely unchanged by the motions. The quadrupolar coupling constant is 55 kHz after averaging over fast methyl rotations.

Our second system is designed to test the applicability of the methods to protein samples with much lower sensitivities than small molecule compounds and with complex motional models. In particular, we employ amyloid fibril systems with monomers consisting of A$\beta_{1\text{-}40}$ labeled at a single methyl group: the CD_3 side chain of the A2 residue located in the beginning of the flexible N-terminal domain (spanning residues 1–16). We have previously characterized the motions of this domain at the A2-CD_3 site using static solid-state NMR techniques (Au et al., 2019; Vugmeyster et al., 2019; Vugmeyster et al., 2020). In the hydrated state, the μs-ms motions at this site can be described by two essential processes. The main state (labeled as "free" in **Figure 3B**) is characterized by the pronounced large-scale fluctuations of the domain, approximated as isotropic diffusion with the diffusion coefficient D. They dramatically narrow the static linewidth (**Figure 2**) with an effective quadrupolar coupling constant of around 3 kHz. This value should be compared with the 53–55 kHz quadrupolar coupling constant expected for the methyl group without large-scale motions (Vold and Vold, 1991). There is also a minor state of the domain at around 8% of the population for the A2-CD_3 site, in which this diffusive motion is quenched. The two states are in the conformational exchange process, with the rate constant (k_{ex}) ranges as determined previously. The presence of the chemical exchange process was particularly evident from the dispersion pattern of ^{2}H $R_{1\rho}$ profiles under static conditions. (Au et al.,

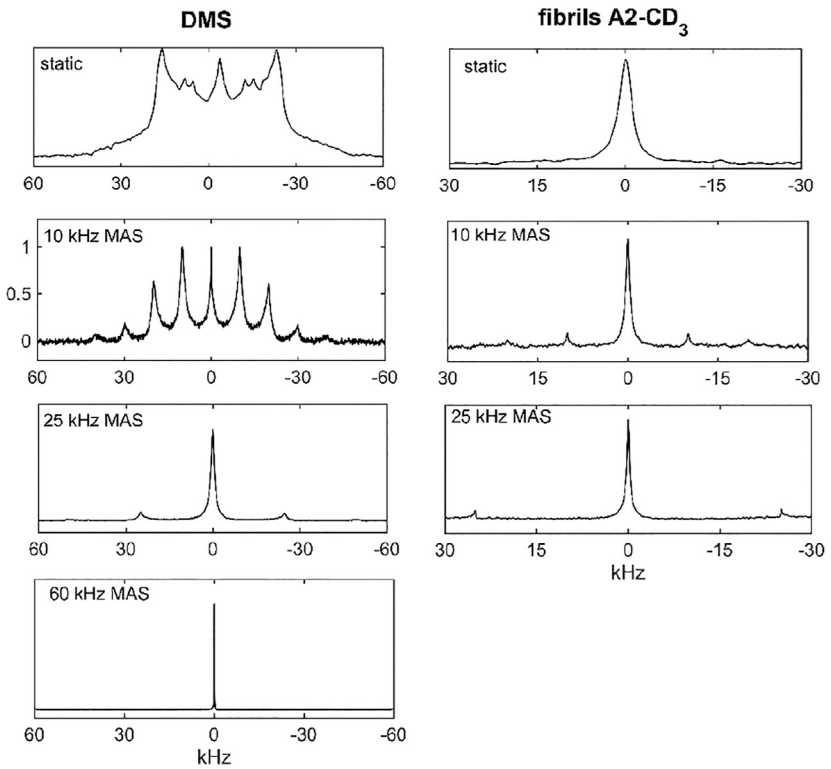


FIGURE 2 | Spectra for DMS-D_6 (left column) and hydrated A$\beta_{1\text{-}40}$ fibrils in the 3-fold symmetric polymorph labeled at the A2-CD_3 site (right column). The following conditions applied: static data–14.1 T and 76°C for DMS and 37°C for the fibrils; 10 and 25 kHz MAS rates–17.6 T and 76°C for DMS and 37°C for the fibrils; 60 kHz MAS rate for DMS only–14.1 T and 55°C. The number of scans and processing parameters are listed in **Supplementary Table S1**.

2019). The combined analysis of the ^{2}H static rotating frame relaxation rates $R_{1\rho}$, quadrupolar CPMG, and CEST calls for a more complicated model in which there is an ensemble of free states with a range of diffusion coefficients that are in conformational exchange with a single rigid bound state. Our strategy to extend the experiment to MAS conditions is to employ the simplest 2-state model of **Figure 3B** and assess if the fitted values of D and k_{ex} fall within the boundaries found by previous ^{2}H static NMR techniques.

Quadrupolar Chemical Exchange Saturation Transfer Experiment and RF Field Strength Calibrations

A simple quadrupolar CEST pulse sequence (**Supplementary Figure S1**) consists of the low-amplitude saturation pulse ω_{RF} (T, Ω) that acts on the longitudinal magnetization, followed by a non-selective 90° pulse with the same phase that brings the magnetization to the transverse plane for detection. The essence of the quadrupolar CEST is to scan the resonance offset values Ω corresponding to the entire spectral region(s) for all conformational states involved, which often falls into the −60–+60 kHz range for methyl groups. The saturation time T is chosen to optimize the efficiency of the conformational exchange and competing longitudinal relaxation. For deuteron in the methyl groups, typical values of T are expected to be between 1 and 40 ms. To determine the motional parameters, the RF field strength should be lower than the typical value of the quadrupolar frequencies (defined in **Eq. 4** in the Theory section) in the two exchanging states. For methyl groups, the 1–5 kHz range is likely to represent the optimal conditions for most samples.

Precise calibrations of the RF field strength using the nutation experiment is complicated by two factors: the evolution of quadrupolar coupling during the nutation pulse and presence of RF inhomogeneity. The evolution of quadrupolar coupling is stronger for larger quadrupolar coupling constants and lower MAS rates (**Supplementary Figure S2**). The effective width of the DMS quadrupolar tensor after averaging over methyl rotations is around 55 kHz. Thus, for the 60 kHz MAS rate, the nutation experiment can be performed on the DMS central band directly with a slight correction for quadrupolar evolution. However, for the 10 and 25 kHz MAS rates, quadrupolar evolution during the nutation pulse is too pronounced and nutation can instead be performed on compounds with naturally narrower tensors. One option is to use liquid D_2O under static conditions, utilizing the same probe as used for the compound of interest. If a solid powder sample is desirable, a good choice is hexamethyl-benzene-D_{18}, whose six methyl groups participate in fast methyl jumps and 6-site jumps about the ring axis, (Vold, 1994; Gupta et al., 2015), leading to an effective quadrupolar coupling constant of about 23 kHz with an asymmetry parameter of 0.07 (Vold et al., 2009). In principle, it is also possible to

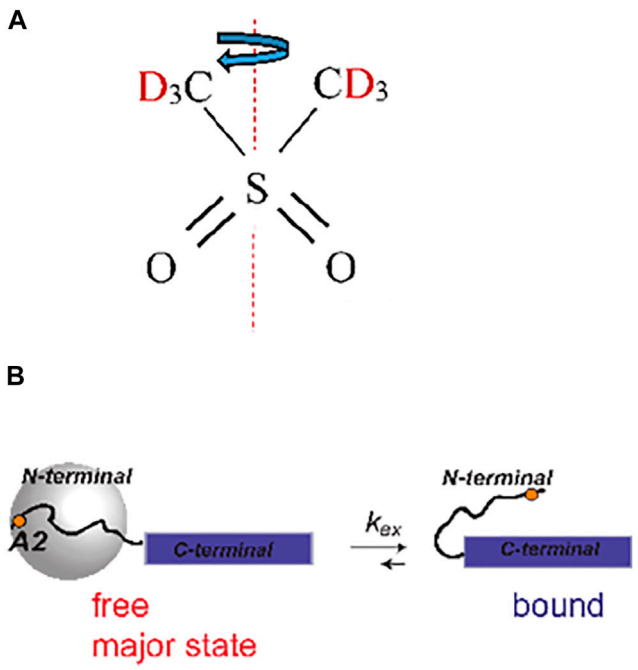


FIGURE 3 | (A) Structure of DMS indicating the 180° flipping motion around the C_2 axis. **(B)** Slow time-scale motional modes of the A2-CD_3 methyl group of the N-terminal domain of $A\beta_{1\text{-}40}$ fibrils (Au et al., 2019). The disordered N-terminal domain (residues 1–16, curved line) transiently interacts with the structured C-terminal domain (blue rectangle). In the free state, the N-terminal domain undergoes large-scale diffusive motion, as represented by the gray sphere, which is absent in the bound state. The position of the A2 residue is shown as an orange dot.

perform nutation measurements directly on the fibrils sample labeled at the A2-CD_3 site due to the narrow effective tensor in the dominant state of the fibrils at the A2 site (**Figure 2B**).

The RF inhomogeneity profiles can also be assessed using the nutation experiment (**Figure 1A**) (Gupta et al., 2015) For the 2.5 mm Bruker probe used for the 10 and 25 kHz MAS rates, powder HMB-D_{18} is the sample of choice with the sample in the rotor having a comparable length to that of DMS in the same probe. DMS is used for the 1.3 mm Bruker probe at the 60 kHz MAS rate. The RF inhomogeneity profiles (i.e., the shape of the distribution of the RF frequencies detected by the nutation experiment) are approximately proportional to the average RF frequency for a number of nominal applied RF powers. This allows us to construct a combined profile as a function of $\omega_{RF}/<\omega_{RF}>$ (**Figure 1B**), in which $<\omega_{RF}>$ is the weighted average over the distribution. We report $<\omega_{RF}>$ as the RF field strength for the CEST measurements. The inhomogeneity is rather significant and roughly comparable with the profiles reported by Gupta et al. for the 2.5 mm probe focusing on ^{13}C frequency (Gupta et al., 2015). The inhomogeneity can be expected to affect the CEST measurements. Thus, the modeling procedures for the simulations of the dynamics also need to include these distributions.

Insights From Theory and Simulations

The following matrices (plus the identity matrix) constitute a basis of the density matrix for the spin-1 system, as well as operators acting in this space: (Grey et al., 1993)

$$\hat{S}_x = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \hat{S}_y = \frac{1}{2}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \hat{J}_x = \frac{1}{2}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & i \\ 0 & -i & 0 \end{pmatrix}, \hat{J}_y = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$
$$\hat{J}_z = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \hat{K} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \hat{S}_z = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \hat{Q} = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{1}$$

The first row represents the single-quantum coherences, followed by the two double-quantum coherences, $\hat{K}$ and $\hat{J}_z$, $\hat{S}_z$ and $\hat{Q}$ stand for the Zeeman and quadrupolar order.

During the saturation period and in the frame rotating with the Larmor frequency, the secular part of the Hamiltonian is given by

$$\hat{H} = \sqrt{\frac{2}{3}}\omega_Q\hat{Q} + \sqrt{2}\omega_{RF}\left(\hat{S}_x \cos\Omega t - \hat{S}_y \sin\Omega t\right) \tag{2}$$

where ω_{RF} is the RF field strength and Ω is its off-resonance offset. ω_Q is the frequency of the secular part of the quadrupole interaction with the angles (θ, ϕ) representing the rotation of the principal-axis system of the quadrupole interaction with respect to the laboratory frame.

$$\omega_Q = \frac{3\pi}{2}C_q\left(\frac{3\cos^2\theta - 1}{2} + \frac{\eta}{2}\sin^2\theta\cos 2\phi\right) \tag{3}$$

The quadrupolar coupling constant is given by $C_q = \frac{e^2qQ}{h}$, and $\eta = \frac{q_{xx}-q_{yy}}{q_{zz}}$ represents the asymmetry of the tensor, defined in the interval $0 \le \eta \le 1$ with $|q_{zz}| \ge |q_{yy}| \ge |q_{xx}|$. eQ is the electric quadrupole moment of the nucleus and eq is the largest component of the electric field gradient.

Under MAS rotation and with $\eta = 0$, ω_Q becomes

$$\omega_Q(t) = \frac{3\pi}{4}C_q\left(\sqrt{2}\sin 2\beta \sin(\omega_{MAS}t + \alpha) - \sin^2\beta\cos(2\omega_{MAS}t + 2\alpha)\right) \tag{4}$$

where β and α are the polar and azimuthal angles with respect to the axis of rotation.

In the frame with an additional rotation with frequency Ω around the z-axis, the Hamiltonian of **Eq. 2** can be transformed into the tilted frame:

$$\hat{H}_{sec} = \sqrt{\frac{2}{3}}\omega_Q\hat{Q} + \sqrt{2}\omega_{RF}\hat{S}_x + \sqrt{2}\Omega\hat{S}_z \tag{5}$$

Analogous to the off-resonance rotating frame relaxation case considered in detail for homonuclear interactions, (Rovó and Linser, 2017; Krushelnitsky et al., 2018; Rovó et al., 2019), the effect of the last two terms of **Eq. 5** can be considered as an action of the effective field given by

$$\omega_e = \sqrt{\omega_{RF}^2 + \Omega^2} \tag{6}$$

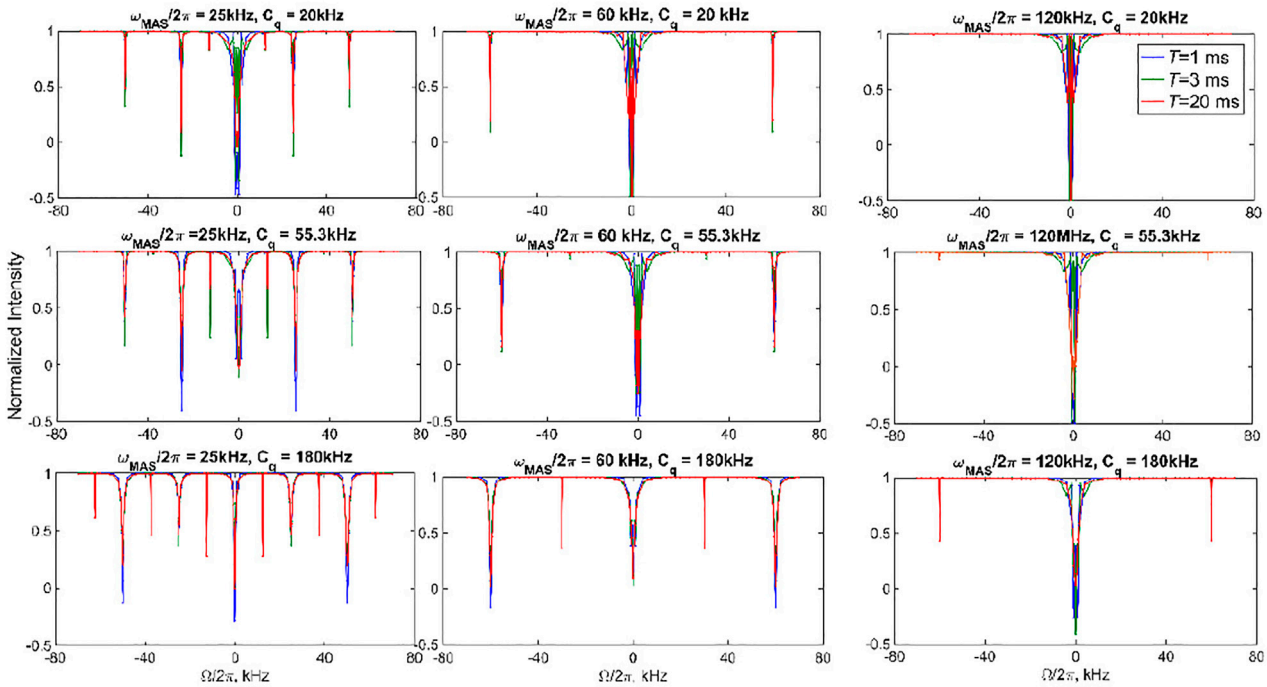


FIGURE 4 | Simulated ^{2}H CEST profiles in the absence of motions. The integrated intensity of all the spectral side bands normalized to intensity at T = 0 versus offsets $\Omega/2\pi$ for $\omega_{RF}/2\pi$ = 1.3 kHz and three chosen values of the saturation times T. Values of C_q and $\omega_{MAS}/2\pi$ are shown directly on the panels. η =0 in all cases.

For the small values of $\omega_{RF} \ll |\Omega|$ employed in the CEST experiment, one expects the occurrence of rotary resonances at $|\Omega| = \frac{n}{2}\omega_{MAS}$, in which n is an integer. The condition for the half-integer MAS rate is again analogous to homonuclear dipolar recoupling, (Rovó and Linser, 2017; Krushelnitsky et al., 2018; Rovó et al., 2019), in which it is referred to as the HORROR condition (Nielsen et al., 1994). SI1-A provides a theoretical description of the rotary resonances' positions and relative widths based on second-order perturbation theory. The insight rendered by this theoretical description can also be demonstrated using simulations in which the Liouvillian equation (**Supplementary Equation S1**) is solved explicitly without any approximation. Because of the large magnitude of the quadrupolar tensor interactions with the C_q values comparable to the effective fields employed, the transformation into the tilted frame of the effective field does not lead to any simplification or render additional qualitative insights.

Figure 4 demonstrates several examples of ^{2}H CEST profiles corresponding to coherent contributions in the absence of motions for axially symmetric tensors with three values of C_q (20, 55, and 180 kHz), MAS rates of 25, 60, and 120 kHz, and ω_{RF} = 1.3 kHz. The rotary resonances are evident at the values of the offsets equal to integer and half-integer values of the MAS frequency. Their intensities are modulated by the interplay between the MAS rates and C_q. The intensity of the resonances depends on the spectral intensity at $\Omega/2\pi$ frequencies. The half-integer resonances are much narrower (**Supplementary Figure S3A**) and often not as deep as the integer ones, as predicted by simple perturbation theory considerations. (**Supplementary Information S1**). We also explore the coherent behavior of the individual coherences of **Eq. 1** for single crystallites, which demonstrates the extent of the coherent oscillations for the single and double quantum coherences and confirms the qualitative insights from perturbation theory. **Supplementary Figure S3B** shows an example for a single crystallite oriented at 30° to the MAS axis.

As usual, motions are introduced into the Liouvillian equation by expanding the density matrix ρ into a direct product of the coherences (**Eq. 1**) and the sites corresponding to either different intra-molecular orientations of the quadrupolar tensor or changes in the value of C_q or η (Vold and Vold, 1991). In this direct product, the coherent evolution acts on the coherences confined to the same site, but with the site-dependent value of ω_Q. The motions are introduced through the matrix elements between the same coherences belonging to different sites, thus encoding the model of Markovian jumps between sites. Because the rate constants of the jumps do not depend on the individual coherences, they can be represented as elements of an exchange matrix. The extended description applicable to the ^{2}H CEST experiment is given in previous work (Vugmeyster et al., 2020).

The motions induce the relaxation of the coherences. To gain insights into the effect of motions on the ^{2}H CEST profiles and interplay between the values of the rate constants and MAS rate, we consider the relaxation behavior according to a simple 2-site exchange model with two axially symmetric tensors. The geometry of the motions is chosen as in **Figure 3A** (i.e., a jump angle of 106°) and three C_q values are considered (**Figure 5**). We select several resonance offset values, several

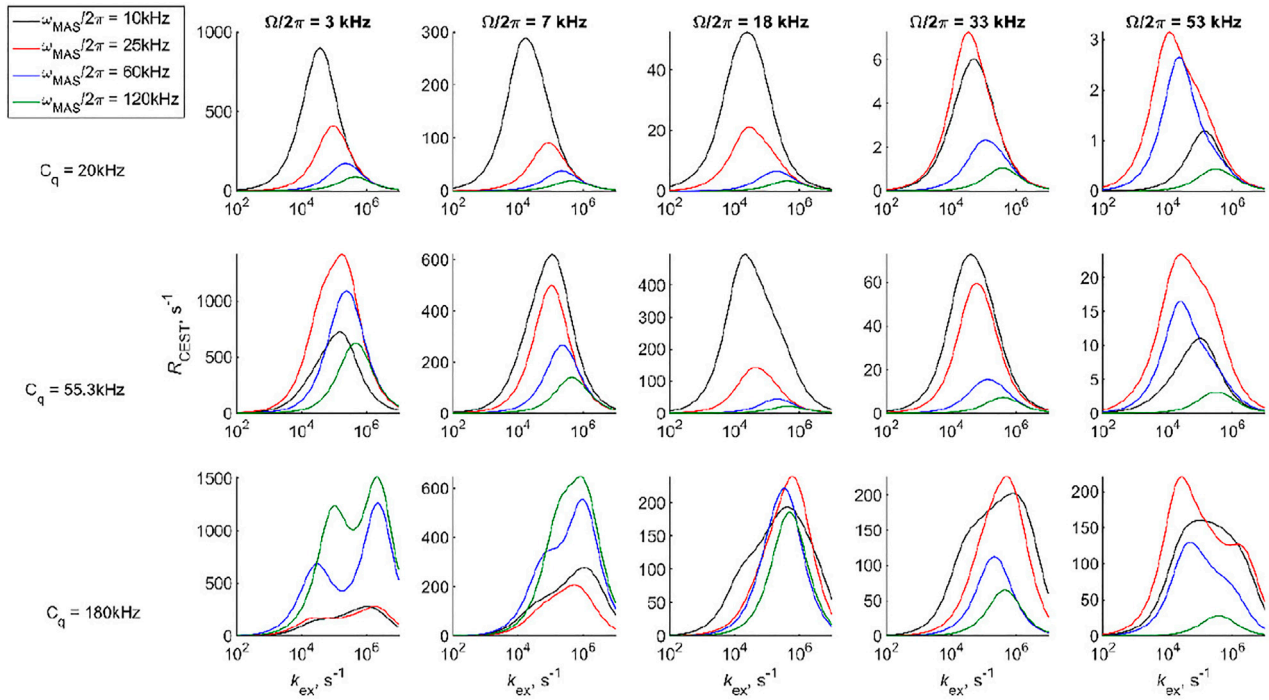


FIGURE 5 | Simulated values of R_{CEST} versus k_{ex} shown for several selected offsets for the 2-site jump model with equivalent axially symmetric tensors, an equal population of sites, and a jump angle of 106°. The R_{CEST} values were obtained for integrated spectral band intensities with a single exponential approximation to the decay curves for T in the 0.25–128 ms range. $\omega_{RF}/2\pi$ was fixed at 1.3 kHz and the four MAS rates of 10, 25, 60, and 120 kHz were selected. The three values of C_q of 20, 55.3, and 180 kHz were chosen.

MAS rates in the 10–120 kHz range, and saturation times T in 0.25–128 ms range for $\omega_{RF}/2\pi$ = 1.3 kHz. For these ranges, the magnetization decays of integrated spectral band intensities can be approximated as single exponential. The effectiveness of the relaxation, given by the relaxation rate R_{CEST}, is an interplay between several factors. First, the most effective relaxation is in a broad region of rate constants k_{ex} between 10^4 and 10^6 s^{-1}. Second, the condition $|\Omega|/2\pi < C_q$ is necessary for effective relaxation because it ensures significant spectral intensity at the $\Omega/2\pi$ frequency. Third, the ω_{MAS} dependence of the relaxation rate is heterogeneous, as it depends on the values of both C_q and Ω. For low values of C_q, the relaxation rate decreases as ω_{MAS} increases, while for high C_q values, this trend is observed only for relatively high values of Ω, but is reversed for low values of Ω. **Supplementary Information S1B** provides qualitative insights into the origin of these trends. Additional mechanisms affecting the $\hat{S}_z$ coherence, such as fast time-scale motions causing longitudinal relaxation, typically lead to strongly non-exponential magnetization decay curves and preclude the R_{CEST} type analysis outlined here. Instead, we focus on the CEST profiles over the range of Ω values but for selected saturation times T.

The sensitivity to k_{ex} can be seen directly from the simulated CEST profiles of the 2-site exchange model for two axially symmetric tensors, which display characteristic line broadening when the time scale of the exchange processes falls within the range CEST sensitivity (**Figure 6**). The calculations in **Figure 6** are performed with the DMS tensor parameters and geometry of **Figure 3A** (i.e., C_q values of 55.3 kHz at both sites and a jump angle of 106°) and a fixed MAS rate of 30 kHz. The overall line broadening of the profiles falls into the 10^4–10^6 s^{-1} rate constant range, as expected from the analysis in **Figure 5**. While ample broadening is observed in the center of this range for all the resonance offset values, at the edges of the sensitivity ranges, the center region for which $|\Omega|/2\pi \ll C_q$ is differentially broadened, especially for short saturation times. Thus, to assess the time scales of the motions, it is critical to measure different values of the saturation fields and saturation times to capture the pattern of the entire profile. Another important feature is the broadening of the coherent resonances in the presence of slow motions and consequent differential changes in intensities between the half- and full-integer rotary resonance conditions. The latter can also be useful in a qualitative assessment of whether the system falls closer to the fast or slow ends of the sensitivity range. For example, the panels corresponding to $k_{ex} = 3{\cdot}10^3$ s^{-1} and $k_{ex} = 1{\cdot}10^6$ s^{-1} are qualitatively similar in the saturation patterns, except for the first half-integer rotary resonance behavior, which is much more broadened in the $k_{ex} = 1{\cdot}10^6$ s^{-1} case. In general, these broadening patterns of rotary resonances are expected to be sensitive to both the values of the rate constants and the choice of the motional model, similar to the Near Rotary Resonance Relaxation Dispersion effects in rotating frame relaxation experiments (Kurauskas et al., 2017; Krushelnitsky et al., 2018; Rovó et al., 2019). Additional 2-site

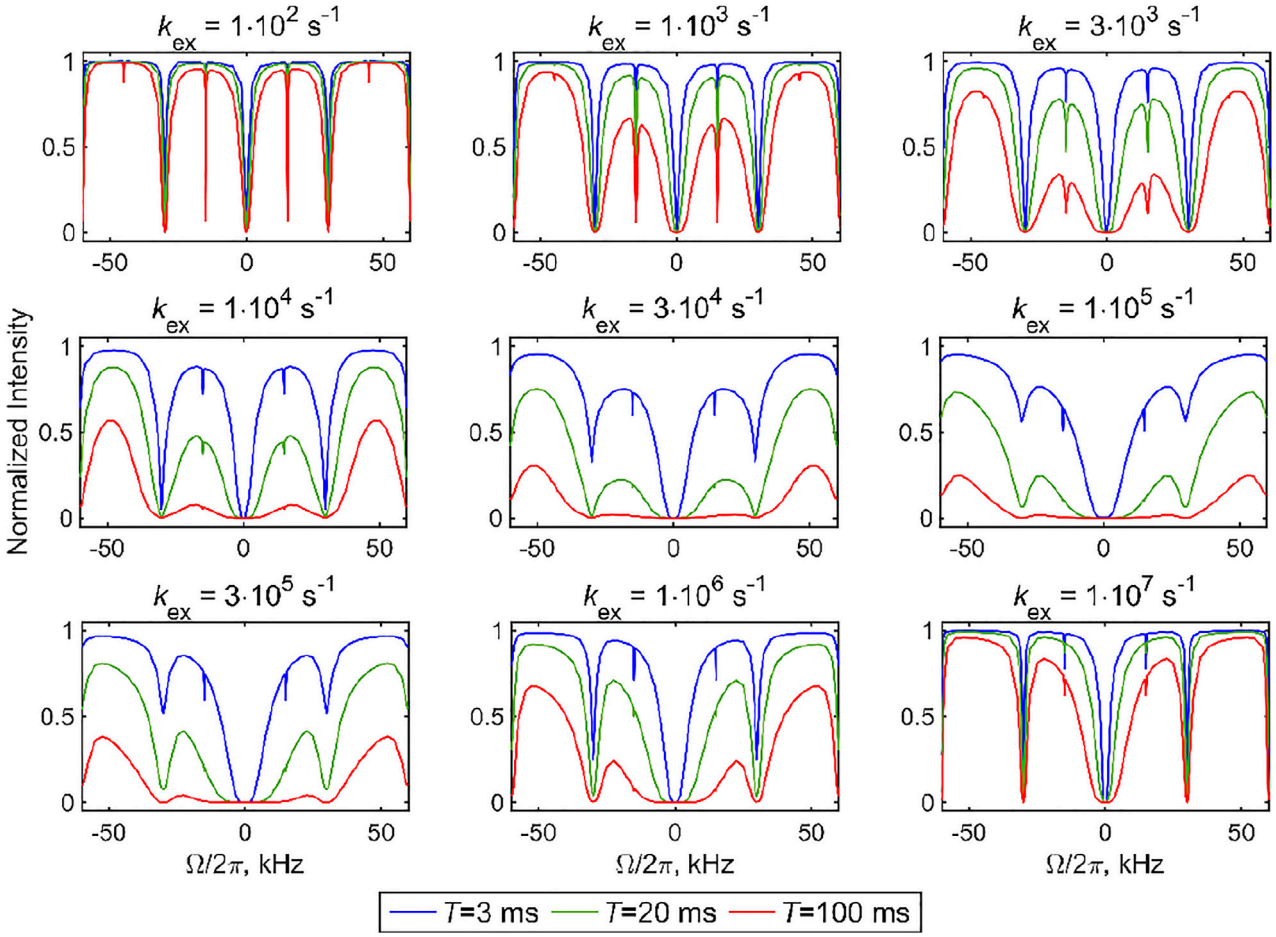


FIGURE 6 | Simulated ^{2}H CEST profiles for the 2-site exchange model using two axially symmetric tensors with the quadrupolar coupling constants of C_q = 55.3 kHz in both sites, equal populations, a jump angle of 106°, and values of the k_{ex} rate constant between 1·10^2 and 1·10^7 s^{-1}. The integrated intensity of all the spectral side bands normalized to intensity at T = 0 versus offsets $\Omega/2\pi$. The simulations were performed with $\omega_{MAS}/2\pi$ = 30 kHz, $\omega_{RF}/2\pi$ = 1.3 kHz, 3000 powder crystallite orientations, and three saturation times T. The offset sampling schedule was 2.5 kHz outside the rotary resonances and 250–500 Hz in the vicinity of the rotary resonances.

exchange examples are shown in **Supplementary Figure S4**, which includes the dependence on the MAS rate, C_q values, and unequal populations.

Dimethyl-Sulfone Results

Before evaluating the effect of the motions on the CEST profiles of DMS-D$_6$ at high temperatures at which the motions are most pronounced, we first performed the measurements at a low temperature at which the flip motion is essentially frozen. These measurements were done to confirm the effect of coherent contributions, i.e., the presence of resonances at $\pm\,\Omega = \frac{n}{2}\omega_{MAS}$. **Figure 7** shows the ^{2}H CEST profiles of DMS at 270 K with an MAS rate of either 10 or 25 kHz and using a saturation field of 1.3 kHz and a saturation time of 3 ms. The presence of integer rotary resonances is evident throughout the profiles, and the first half-integer resonance (n = 1) can also be seen. The width of the half-integer resonances is significantly narrower than that of the whole ones (see the theoretical considerations in SI1-A). Thus, to observe them, a dense sampling throughout the offsets is needed. We focused the dense sampling schedule on the n = 1 condition, i.e. $\pm\,\Omega = \frac{1}{2}\omega_{MAS}$, to demonstrate the principle.

The presence of multiple resonances (i.e., the coherent contributions shown in **Figures 4**, **7**) in the CEST profiles at low MAS rates precludes the quantitative interpretation of the motional contributions of these profiles. In most cases, a compromise needs to be found between an MAS rate high enough not to render extensive resonance patterns and low enough to retain a sufficient magnitude of the unaveraged quadrupolar interaction. For DMS, we collect high temperature data at the 25 kHz MAS rate (at 76°C, 17.6 T, 2.5 mm probe) and 60 kHz MAS rate (at 55°C, 14.1 T, 1.3 mm probe) to analyze the sensitivity of the profiles to the flipping motion, which is characterized by the rate constant k_{flip} (**Figure 3**). The overall strategy is to fit the experimental data to simulations as a function of k_{flip} to assess whether the resulting fitted values fall within the range determined by other NMR techniques as well as evaluate the general sensitivity of the technique.

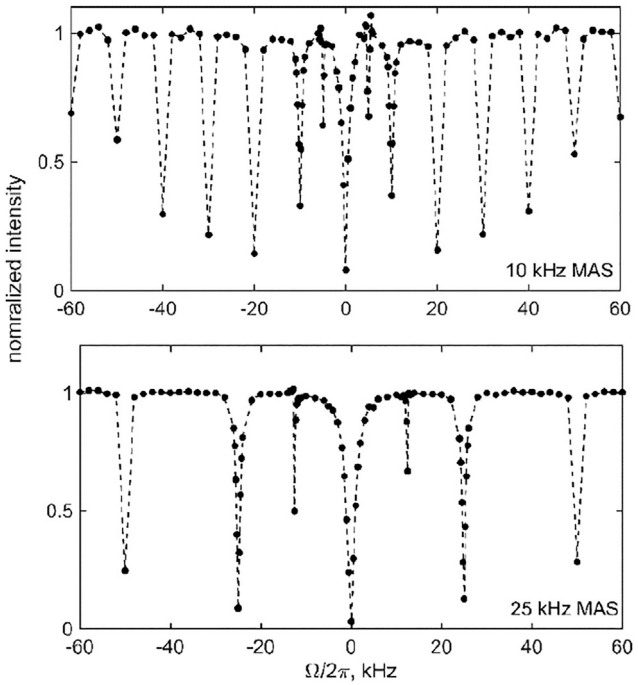


FIGURE 7 | Experimental ^{2}H MAS CEST profiles for DMS-D6 at 270 K and 17.6 T for the 1.3 kHz RF field strength and T = 3 ms. Normalized integrated central band intensity versus resonance offset $\Omega/2\pi$ for MAS rates of 10 kHz (upper panel) and 25 kHz (lower panel). The dotted lines are shown to guide the eye.

As follows from the theoretical discussion of the 2-site jump model results (**Figure 6**), it is desirable to obtain the data at more than one combination of RF field strength and saturation time to cross-validate the fits and models. With high sensitivity samples such as DMS, this task is relatively easy to accomplish. We collect the data at the 25 kHz MAS rate at RF field strengths of 1.3 and 2.5 kHz and saturation times of 3 and 20 ms, with the experimental ranges of the spin-locking fields and saturation times mimicking those used to develop the CEST technique under static conditions. For the faster 60 kHz MAS rate, we utilize RF fields in the range of 1.5–4.7 kHz and saturation times between 3 and 20 ms. **Figures 8**, **9** present the experimental results corresponding to the central band. There is a negligible difference in the profiles when the central band results are compared with the sum of all the bands (**Supplementary Figure S5**).

The data were then fitted with the two-mode motional model: the slow 106° flip mode between two equally populated rotamers with the corresponding rate constant k_{flip} (**Figure 3A**) and fast time-scale methyl 3-site jumps with the rate constant k_3. The k_{flip} rate was varied and k_3 value was fixed from the fits to the longitudinal relaxation times T_1. The T_1 times were measured with the inversion recovery pulse sequence and were 41 ms for 76°C and the 25 kHz MAS rate and 26 ms for 55°C and the 60 kHz MAS rate, corresponding to 3-site jump rate constants of $4.9{\cdot}10^9\ \text{s}^{-1}$ and $3.2{\cdot}10^9\ \text{s}^{-1}$, respectively. The best-fit k_{flip} value for the 76°C data was between 9,000 and 10,000 s^{-1}, whereas it was 2,100 s^{-1} for the 55°C data (shown in **Figure 8** as solid lines). These are in the range found by other techniques, particularly static ^{2}H CEST (Au et al., 2019; Vugmeyster et al., 2019; Vugmeyster et al., 2020). All the simulations of the profiles included the effect of the RF inhomogeneity of the coil with the profiles of **Figure 1B** the CEST profiles simulated without the effect of RF inhomogeneity overestimated the saturation for offsets for which the saturation extent was significant, thus also affecting the overall shape of the profile, not only the resulting fitted rate constant. An example of one such profile simulated without inhomogeneity for $\omega_{RF}/2\pi$ = 2.5 kHz is shown by the dotted line in **Figure 8**.

The sensitivity of the fits to k_{flip} is shown in **Supplementary Figures S6–S8**, which allow us to assess the quality of the fits using the mean absolute difference between the experimental and simulated profiles. They demonstrate that a careful choice of RF field and saturation time delay is needed to determine the motional rate constant precisely. This is especially evident for the 60 kHz MAS data for which a wider range or RF fields are considered. At $\omega_{RF}/2\pi$ = 1.55 kHz, the 3 ms saturation time appears to be too weak to cause any significant motion dependence, while at $\omega_{RF}/2\pi$ = 2.8 and 4.7 kHz, the 5 and 20 ms saturation times both yield the desired sensitivity to the motional parameters. The effective tensor narrowing due to the fast MAS rate leads to the necessity of larger saturation times to observe the motional effects.

The data also confirm the effect of the motions on broadening the rotary resonances: the width of the full-integer rotary resonance widths is consistent between the experiment and simulations at both MAS rates. The half-integer resonances at the high temperature of 76°C at which the flipping motions are most pronounced are completely broadened for the 25 kHz MAS rate results, in accordance with the theory. The simulations show residual first half-integer rotary resonance peaks but these are too small to detect in the experiment. At the lower temperature of 55°C and high MAS rate, the first half-integer resonance is clearly visible in the data. In general, RF inhomogeneity can affect the apparent width of the rotary resonances.

Additional insights can be obtained by focusing on the intensities of selected offsets for several values of ω_{RF} (**Figure 10**). This type of analysis can provide further confirmation of the model as well as point to the limits of validity of the approximations used to model RF inhomogeneity. In the case of DMS, at the 60 kHz MAS rate, for offset values below 3–5 kHz at which the saturation of intensities due to motions is most pronounced, the current approximation used for the simulations of inhomogeneity is likely to be somewhat imprecise. In general, however, **Figure 10** demonstrates the good agreement between the modeled and experimental RF field strength dependence when the RF inhomogeneity profile of the probe is taken into account. The inhomogeneity effect is more pronounced for higher values of ω_{RF} as expected. The dynamic radial RF inhomogeneities induced by sample rotation might become relevant, (Tošner et al., 2017). They can be simulated by introducing an additional fluctuating term along $\hat{S}_x$ as a function of the phase of the MAS rotation. For the conditions of our experiment it

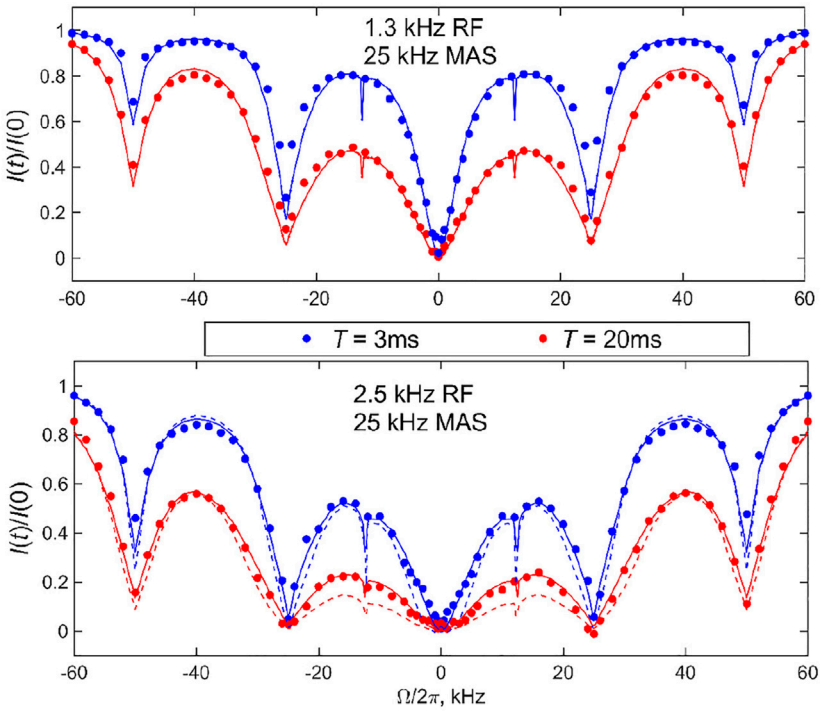


FIGURE 8 | ^{2}H CEST center band profiles for DMS at the 25 kHz MAS rate. Experimental normalized integrated central band intensities $I(t)/I$ (0) versus resonance offsets $\Omega/2\pi$ for saturation fields of 1.3 kHz (upper panel) and 2.5 kHz (lower panel) and saturation times of T =3 ms (blue circles) and T =20 ms (red circles) at 17.6 T and 76°C. The solid lines represent the best fit to the data, corresponding to k_{flip} = 9,000 s^{-1} for $\omega_{RF}/2\pi$ = 1.3 kHz and k_{flip} = 10,000 s^{-1} for $\omega_{RF}/2\pi$ = 2.5 kHz. The k_3 value was fixed at 4.9·10^{9} s^{-1}. The effect of RF inhomogeneity with the inhomogeneity profiles of **Figure 1B** was included as described in the text. The dotted lines in the bottom panel show the simulations in the absence of RF inhomogeneity.

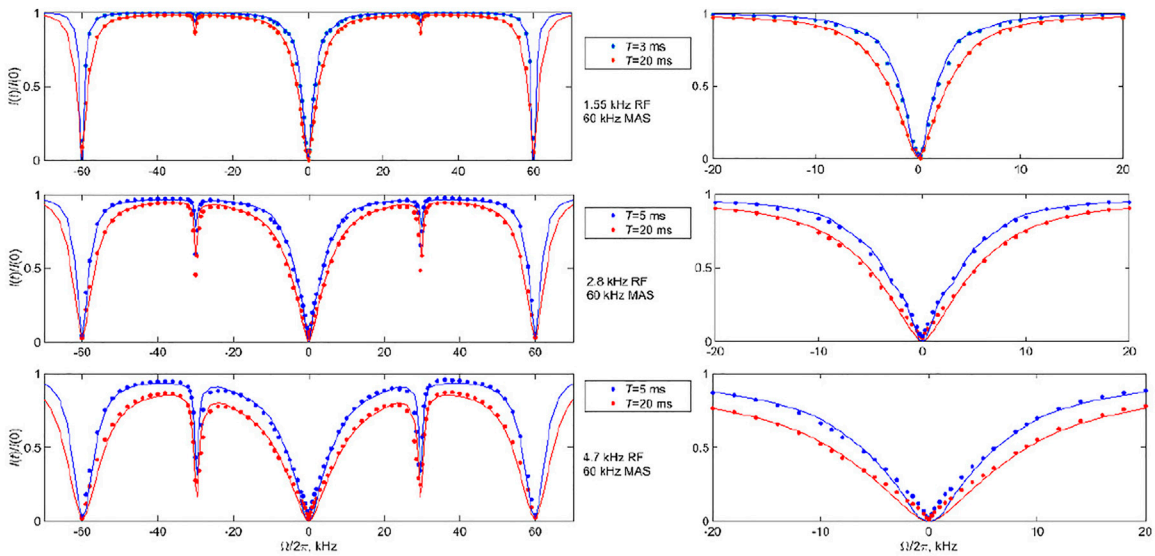


FIGURE 9 | ^{2}H CEST central band profiles for DMS at the 60 kHz MAS rate. Experimental normalized integrated central band intensities $I(t)/I$ (0) versus resonance offsets $\Omega/2\pi$ for the saturation fields of 1.55, 2.8, and 4.7 kHz and saturation times indicated on the panels at 14.1 T and 55°C. The right panels show the expansion of the −20–+20 kHz offset region. The solid lines represent the best fit to the data to the model of **Figure 3A** with a k_{flip} rate constant of 2,100 s^{-1}. The k_3 value was fixed at 3.2·10^{9} s^{-1}. The effect of RF inhomogeneity with the inhomogeneity profiles of **Figure 1B** was included as described in the text.

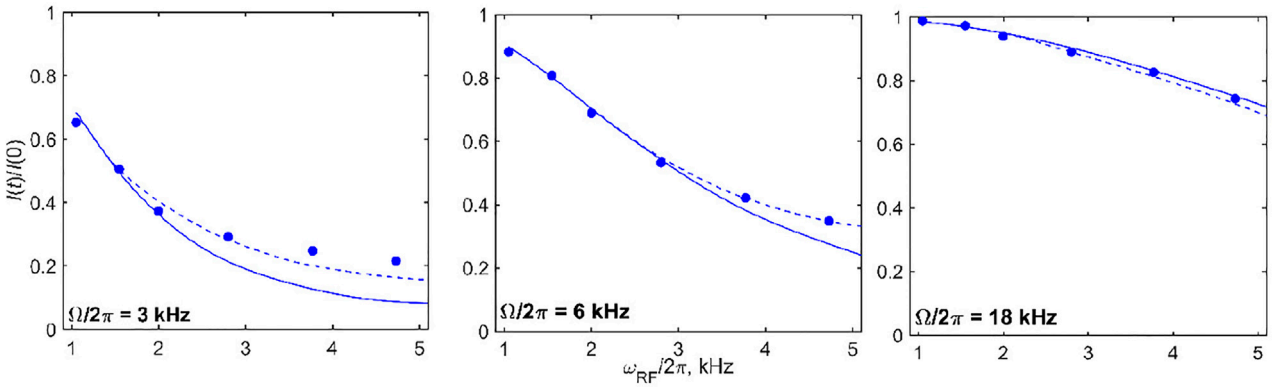


FIGURE 10 | ^{2}H CEST DMS central band intensities $I(t)/I$ (0) for selected values of offsets (shown directly on the panels) as a function of $\omega_{RF}/2\pi$ for T= 20 ms at the 60 kHz MAS rate and at 14.1 T and 55°C. The solid lines represent the simulations according to the model of **Figure 3A** with k_{flip} =2,100 s^{-1} and k_3 = 3.2·10^9 s^{-1} without the inclusion of RF inhomogeneity. The dotted lines stand for the same simulations with the inclusion of the RF inhomogeneity profiles of **Figure 1B**. Experimental errors are within the sizes of the symbols and were obtained as standard errors from three repeated measurements.

turned out to be minor and definitely within the experimental errors. The radial inhomogeneity effects can become more important for the spin-locking of magnetization around the *x*-axis.

Overall, the analysis of the DMS profiles demonstrates that when the appropriate saturation conditions are satisfied, ^{2}H CEST under MAS is a sensitive technique for the detection of slow time-scale motions. In comparison to methyl groups, for aromatic rings sites and backbone C_α sites, which generally correspond to C_q = 180 kHz, the condition for optimal CEST sensitivity range may be shifted to higher MAS rates (see **Figure 5**), as well as toward potentially higher values of the saturation field strength. The choice of the best experimental conditions will be ultimately governed by the tensor magnitude, the time scales of motions, and tolerance of the sample toward the RF-induced heating.

Aβ Fibril Results

For the Aβ fibrils labeled at the A2-CD_3 site, due to the significant narrowing of the static linewidth in the free state of the fibrils (see the spectra in **Figure 2**), we choose the 10 and 25 kHz MAS conditions. The measurements were performed at a 17.6 T field strength using a 2.5 mm diameter probe and at 37°C. With significantly longer data acquisition times than for DMS, the single RF field strength of 1.3 kHz and two saturation times of 3 and 20 ms suffice (**Figure 11**). The overall data collection time was 5.5 days. The profiles clearly display the presence of coherent rotary resonances. The half-integer resonances are difficult to observe due to the need to implement the detailed sampling schedules necessary to catch these relatively narrow dips. We include enough offsets to observe the n = 1 half-integer resonances at $\Omega/2\pi$ = ±12.5 kHz for the 25 kHz MAS condition to explicitly confirm their existence. The profiles are clearly sensitive to the choice of saturation time (3 or 20 ms). The T_1 relaxation time of the A2 methyl group is 51 ms. For the 10 kHz MAS condition the width of the profile is somewhat dependent on whether the central band or the sum of all the bands is used. This difference is not observed for the 25 kHz MAS condition. In the discussion of the modeling and fitted parameters we will focus on the analysis for the sum of the intensities of all the side-bands and return to the potential origin of the slight differences in the profiles at the end of the section.

The modeling was performed according to the 2-state model of **Figure 3B**. The diffusive motion of the N-terminal domain in the free state is incorporated *via* a matrix of 192 neighboring sites on the surface of a sphere, (Au et al., 2019), with one additional site representing the bound state. To optimize the simulation time for the system with many exchanging sites, the additional mode of fast methyl 3-site jumps can be included as a phenomenological R_1 term, (Vugmeyster et al., 2020), rather than introducing an additional explicit motional frame that would triple the total number of sites. Including 20 steps in each MAS period and using the model of **Figure 3B** are computationally demanding tasks (the details are listed in the Modeling section). The simulations took 96 h with our computational system for each MAS rate condition and a single set of the D and k_{ex} values with the inclusion of the RF inhomogeneity effect of **Figure 1B**.

Thus, rather than calculating a comprehensive (D, k_{ex}) grid, we retained within the main range found previously using the static CEST and $R_{1\rho}$ methods: the value of D varied from 1·10^6 to 6·10^6 rad^2/s, while k_{ex} varied from 1·10^4 to 2·10^5 s^{-1}. Further, instead of calculating all the profiles for the best-fit analysis, we focused on a range of characteristic offset values that can capture the widths of the pattern and resonance positions. The values were (±2, ±4, ±6, ±8, ±14, ±16, ±24, ±26 and ±10, ±20, ±30 kHz) for the 10 kHz MAS rate and (±2, ±4, ±6, ±8, ±10, ±16, ±18, ±20, and ±12.5, ±25, ±50 kHz) for the25 kHz MAS rate.

The mean absolute differences of these searches are shown in **Supplementary Figure S9**. There are shallow minima around the best-fit parameters as follows: for the 10 kHz MAS rate, the values are D= 1.7·10^6 rad^2/s and k_{ex} = 5·10^4 s^{-1} for T = 3 ms and D = 1.7·10^6 rad^2/s and k_{ex} = 1·10^5 s^{-1} for T = 20 ms, while for the 25 kHz MAS rate, the values are D = 1.0·10^6 rad^2/s and k_{ex} = 2·10^4 s^{-1} for T = 3 ms

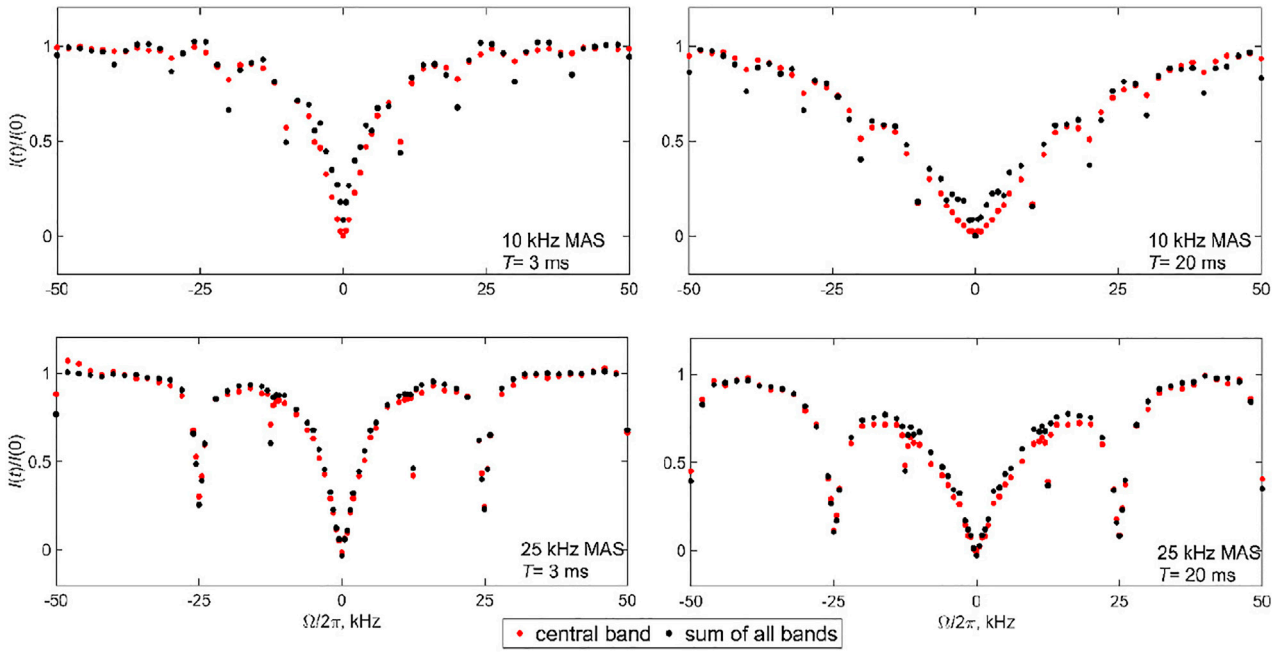


FIGURE 11 | Experimental ^{2}H CEST profiles of the A$\beta_{1\text{-}40}$ fibrils labeled at the A2-CD_3 side chain. The normalized integrated intensities $I(t)/I$ (0) of the central band (red circles) are compared with the intensities of all the spectra bands shown in the spectra of **Figure 2** (black circles) and are plotted as a function of the offsets $\Omega/2\pi$. The saturation times and MAS rates are shown directly on the panels. The intensities $I(t)/I$ (0) are plotted as a function of the offsets $\Omega/2\pi$. Data were collected at 17.6 T and 37°C with $\omega_{RF}/2\pi$ = 1.3 kHz and an MAS rate of either 10 or 25 kHz.

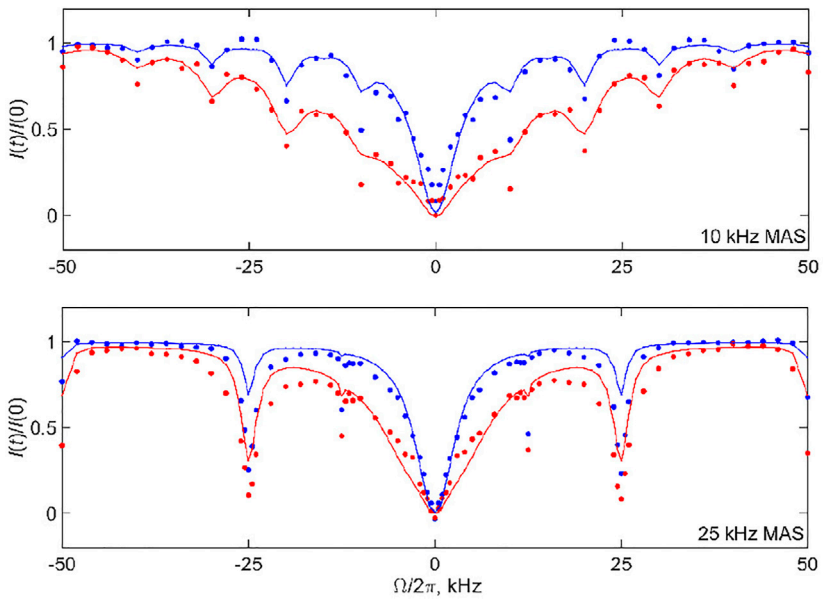


FIGURE 12 | ^{2}H CEST profiles of the A$\beta_{1\text{-}40}$ fibrils labeled at the A2-CD_3 side chain corresponding to the intensities of all the bands (shown in the spectra of **Figure 2**) at the 10 and 25 kHz MAS rates. The experimental normalized integrated summed intensities of all the bands $I(t)/I$ (0) versus resonance offsets $\Omega/2\pi$ for $\omega_{RF}/2\pi$ = 1.3 kHz and T =3 ms (blue circles) and T =20 ms (red circles). The lines represent the best fit to the data according to the model of **Figure 3B** with the following parameters: 10 kHz MAS rate (D = 1.7·10^{6} rad^{2}/s, k_{ex} = 5·10^{4} s^{-1}) for T = 3 ms and (D = 1.7·10^{6} rad^{2}/s, k_{ex} = 1·10^{5} s^{-1}) for T = 20 ms; 25 kHz MAS rate (D = 1.0·10^{6} rad^{2}/s, k_{ex} = 2·10^{4} s^{-1}) for T = 3 ms and (D = 1.7·10^{6} rad^{2}/s, k_{ex} = 3·10^{4} s^{-1}) for T = 20 ms. Data were collected at 17.6 T and 37°C. The effect of RF inhomogeneity with the inhomogeneity profiles of **Figure 1B** was included as described in the text.

and $D = 1.7{\cdot}10^6$ rad^2/s and $k_{ex} = 3{\cdot}10^4$ s^{-1} for $T = 20$ ms. These fits are demonstrated in **Figure 12** by the solid lines. There is a positive correlation between the D and k_{ex} values, which can be rationalized by the fact that fast diffusion narrows the overall CEST pattern, whereas relatively slow conformational exchange widens it. This correlation, which was also noted and analyzed in more detail for the static case, (Vugmeyster et al., 2020), leads to the whole subset of relatively comparable (D, k_{ex}) pairs in terms of the quality of the fits. The shallow minima chosen for the profiles in **Figure 12** are the result of the compromise between matching the overall width of the pattern across all the offsets and the intensities at the coherent resonances. We have performed fits for individual data sets rather than the combined fit in order to determine the ranges of acceptable parameters within the limitation of the model and correlations between the fitted value of D and k_{ex}. If the global fit is performed (**Supplementary Material S10**), the best-fit parameters are $D = 1.7{\cdot}10^6$ rad^2/s, $k_{ex} = 3{\cdot}10^4$ s^{-1}.

The presence of rotary resonances in the data is not trivial, as it confirms from another angle the existence of the slow conformational exchange between the free and bound states of the fibrils. As demonstrated in **Supplementray Figure S11**, the diffusion mode alone can match the width of the narrow central region in the −6–+6 kHz range. However, it not only misses the overall outer width and shape of the pattern [which was also observed for the static data, see **Supplementray Figure S5** of prior work (Vugmeyster et al., 2020)], but also broadens any traces of the coherent resonances. **Supplementray Figure S12** demonstrates the alternative fits for the 10 kHz MAS rate, T= 20 ms profile using three (D, k_{ex}) pairs, including the best-fit profile of **Figure 12**. The parameters ($D = 6.0{\cdot}10^6$ rad^2/s, $k_{ex} = 2{\cdot}10^5$ s^{-1}) lead to a comparable quality of the overall fit judging by the grid search. However, the dips in intensities at the resonance positions are underestimated compared with the best fit results of **Figure 12**. By contrast, the parameters ($D = 1.0{\cdot}10^6$ rad^2/s, $k_{ex} = 5{\cdot}10^4$ s^{-1}) capture the intensities at the resonances somewhat better at the expense of matching the overall width of the pattern. Thus, the optimal values of k_{ex} for observing the existence of coherent resonances fall into the 3–5·10^4 s^{-1} range. For the 25 kHz MAS rate and T= 20 ms profile, **Supplementray Figure S12B** shows how the intensity at the first resonance varies for different pairs of (D, k_{ex}). While none of them capture the experimental intensity directly, the closest match is seen for k_{ex} in the 2–5·10^4 s^{-1} range, while there is relatively weak dependence on the value of D. As half-integer rotary resonances could be difficult to observe, for most cases of low sensitivity samples the observation of the first full-integer resonance's intensity and width will be sufficient for constraining the model and its parameters.

Our fits clarify that the model is far from perfect in catching the exact intensities at the resonances. The first half-integer resonances evident in the experimental data are completely absent for both MAS conditions and the first integer resonance is underestimated. This implies that the exchange processes in the fibrils are likely to be more complex than in the simplified model of **Figure 3B**. Nonetheless, within the precision of the analysis, we confirm the qualitative presence of the conformational exchange and its time scale. Another hint that our model is an oversimplification is suggested by the slight discrepancy between the CEST profiles of the central band and profiles for the sum of all the bands in the 10 kHz MAS data of **Figure 6**. These discrepancies may reflect the existence of more complex ensembles of free and bound states of the N-terminal subdomain of the fibrils. A previous combined analysis of static ^{2}H NMR CEST, rotating frame $R_{1\rho}$ rates, and QCPMG data of the N-terminal subdomain indicated a complex conformational space, corresponding to an ensemble of conformations for the free state in exchange with a single bound state (Au et al., 2019; Vugmeyster et al., 2019; Vugmeyster et al., 2020). The ensemble is characterized by clusters of D values around 1–3·10^6 rad^2/s, and 1·10^8 rad^2/s, with corresponding k_{ex} values clustered in 0.1–1·10^5 and 1–3·10^6 s^{-1}. The values from the new MAS measurements are in line with these previously determined ranges.

The overall strategies presented for DMS and Aβ fibrils will hold for a variety of biological systems including protein aggregates, complexes, and crystals. Simpler NMR measurements such as line shape analysis can serve as a complementary tool to narrow some of the expected time scale ranges. The determination of the models for this complex systems should start with the simplest scenarios of limited number of exchanging sites, and increase in the level of complexity when governed by the experimental data. Future improvements in the speed of computations are expected to greatly benefit model selection procedures.

CONCLUSION

The analyses of the results of DMS-D$_6$ with a simple 2-site rotameric flip model with known parameters as well as the Aβ_{1-40} fibril sample with a complex model previously assessed by other techniques indicated that the ^{2}H CEST experiment can quantify the slow motional modes in rotating solids. For the best precision and motional model development, it is desirable to perform the measurements for more than one combination of the saturation fields and saturation times. Special attention should be paid to the examination of the experimental and simulated intensities at rotary resonance positions, as they can pinpoint to details of motional regimes and mechanisms. This is an additional strength of rotating versus static approach. Assessing probe RF inhomogeneity can be important for improving the accuracy of the results. For low sensitivity protein samples with complex models, MAS rates, saturation field strength, and saturation times must be selected carefully to optimize data collection strategies. Further, the explicit modeling procedures for complex models have to be computationally optimized to render them friendly for model and parameter selection. Once these strategies are in place, the ^{2}H CEST technique can be a powerful tool for studies of protein dynamics.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

LV conceived experiments, designed experiments and models, samples, pulse sequence, wrote the manuscript; DO wrote computer modeling program, conceived strategies and performed computational modeling, participated in data analysis and draft writing; AG assisted with data collection at 17.6T spectrometer; RF assisted with data collection at 14.1T spectrometer, participated in draft editing.

FUNDING

This work was supported by a National Institutes of Health Grant 1R15-GM111681. Several experiments were performed at the National High Magnetic Field Laboratory, which is supported by NSF Cooperative Agreement NSF/DMR-1644779, the State of Florida and the U.S. Department of Energy.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.705572/full#supplementary-material.

REFERENCES

Akbey, Ü., Nieuwkoop, A. J., Wegner, S., Voreck, A., Kunert, B., Bandara, P., et al. (2014). Quadruple-resonance Magic-Angle Spinning NMR Spectroscopy of Deuterated Solid Proteins. *Angew. Chem. Int. Ed.* 53, 2438–2442. doi:10.1002/anie.201308927

Al-Mohy, A. H., and Higham, N. J. (2010). A New Scaling and Squaring Algorithm for the Matrix Exponential. *SIAM J. Matrix Anal. Appl.* 31, 970–989. doi:10.1137/09074721x

Au, D. F., Ostrovsky, D., Fu, R., and Vugmeyster, L. (2019). Solid-state NMR Reveals a Comprehensive View of the Dynamics of the Flexible, Disordered N-Terminal Domain of Amyloid-β Fibrils. *J. Biol. Chem.* 294, 5840–5853. doi:10.1074/jbc.ra118.006559

Bain, A. D., and Berno, B. (2011). Liouvillians in NMR: the Direct Method Revisited. *Prog. Nucl. Magn. Reson. Spectrosc.* 59, 223–244. doi:10.1016/j.pnmrs.2010.12.002

Bjerring, M., Paaske, B., Oschkinat, H., Akbey, Ü., and Nielsen, N. C. (2012). Rapid Solid-State NMR of Deuterated Proteins by Interleaved Cross-Polarization from 1H and 2H Nuclei. *J. Magn. Reson.* 214, 324–328. doi:10.1016/j.jmr.2011.10.020

Bouvignies, G., and Kay, L. E. (2012). Measurement of Proton Chemical Shifts in Invisible States of Slowly Exchanging Protein Systems by Chemical Exchange Saturation Transfer. *J. Phys. Chem. B* 116, 14311–14317. doi:10.1021/jp311109u

Brown, M. J., Void, R. L., and Hoatson, G. L. (1996). Selective Inversion Investigations of Slow Molecular Motion in Solid State Deuteron NMR Spectroscopy. *Solid State. Nucl. Magn. Reson.* 6, 167–185. doi:10.1016/0926-2040(95)01213-3

Favre, D. E., Schaefer, D. J., and Chmelka, B. F. (1998). Direct Determination of Motional Correlation Times by 1D MAS and 2D Exchange NMR Techniques. *J. Magn. Reson.* 134, 261–279. doi:10.1006/jmre.1998.1506

Frydman, L., Vallabhaneni, S., Lee, Y. K., and Emsley, L. (1994). Solid-State Dynamic Processes in Complex Systems Analyzed by Two-dimensional Isotropic-Anisotropic Correlation Nuclear Magnetic Resonance. *J. Chem. Phys.* 101, 111–117. doi:10.1063/1.468185

Gérardy-Montouillout, V., Malveau, C., Tekely, P., Olender, Z., and Luz, Z. (1996). ODESSA, a New 1D NMR Exchange experiment for Chemically Equivalent Nuclei in Rotating Solids. *J. Magn. Reson. Ser. A* 123, 7–15. doi:10.1006/jmra.1996.0208

Grey, C. P., Veeman, W. S. W., and Vega, A. J. A. (1993). Rotational echo14N/13C/1H Triple Resonance Solid-state Nuclear Magnetic Resonance: A Probe of13C-14N Internuclear Distances. *J. Chem. Phys.* 98, 7711–7724. doi:10.1063/1.464579

Gupta, R., Hou, G., Polenova, T., and Vega, A. J. (2015). RF Inhomogeneity and How it Controls CPMAS. *Solid State. Nucl. Magn. Reson.* 72, 17–26. doi:10.1016/j.ssnmr.2015.09.005

Higham, N. J. (2005). The Scaling and Squaring Method for the Matrix Exponential Revisited. *SIAM J. Matrix Anal. Appl.* 26, 1179–1193. doi:10.1137/04061101x

Jain, S. K., Nielsen, A. B., Hiller, M., Handel, L., Ernst, M., Oschkinat, H., et al. (2014). Low-power Polarization Transfer Between Deuterons and Spin-1/2 Nuclei Using Adiabatic RESPIRATIONCP in Solid-State NMR. *Phys. Chem. Chem. Phys.* 16, 2827–2830. doi:10.1039/c3cp54419b

Krushelnitsky, A., Zinkevich, T., Reif, B., and Saalwächter, K. (2014). Slow Motions in Microcrystalline Proteins as Observed by MAS-Dependent 15N Rotating-Frame NMR Relaxation. *J. Magn. Reson.* 248, 8–12. doi:10.1016/j.jmr.2014.09.007

Krushelnitsky, A., Gauto, D., Rodriguez Camargo, D. C., Schanda, P., and Saalwächter, K. (2018). Microsecond Motions Probed by Near-Rotary-Resonance R1ρ 15N MAS NMR Experiments: the Model Case of Protein Overall-Rocking in Crystals. *J. Biomol. NMR* 71, 53–67. doi:10.1007/s10858-018-0191-4

Kurauskas, V., Izmailov, S. A., Rogacheva, O. N., Hessel, A., Ayala, I., Woodhouse, J., et al. (2017). Slow Conformational Exchange and Overall Rocking Motion in Ubiquitin Protein Crystals. *Nat. Commun.* 8, 145. doi:10.1038/s41467-017-00165-8

Matlahov, I., and Van der Wel, P. C. A. (2018). Hidden Motions and Motion-Induced Invisibility: Dynamics-Based Spectral Editing in Solid-State NMR. *Methods* 148, 123–135. doi:10.1016/j.ymeth.2018.04.015

Nielsen, N. C., Bildso/E, H., Jakobsen, H. J., and Levitt, M. H. (1994). Double-quantum Homonuclear Rotary Resonance: Efficient Dipolar Recovery in Magic-Angle Spinning Nuclear Magnetic Resonance. *J. Chem. Phys.* 101, 1805–1812. doi:10.1063/1.467759

Palmer, A. G., and Koss, H. (2019). "Chemical Exchange," in *Methods in Enzymology*. Editor A. J. Wand (Academic Press), 177–236. doi:10.1016/bs.mie.2018.09.028

Palmer, A. G., 3rd (2014). Chemical Exchange in Biomacromolecules: Past, Present, and Future. *J. Magn. Reson.* 241, 3–17. doi:10.1016/j.jmr.2014.01.008

Quinn, C. M., and McDermott, A. E. (2012). Quantifying Conformational Dynamics Using Solid-State R1ρ Experiments. *J. Magn. Reson.* 222, 1–7. doi:10.1016/j.jmr.2012.05.014

Rovó, P. (2020). Recent Advances in Solid-State Relaxation Dispersion Techniques. *Solid State. Nucl. Magn. Reson.* 108, 101665. doi:10.1016/j.ssnmr.2020.101665

Rovó, P., and Linser, R. (2017). Microsecond Time Scale Proton Rotating-Frame Relaxation under Magic Angle Spinning. *J. Phys. Chem. B* 121, 6117–6130. doi:10.1021/acs.jpcb.7b03333

Rovó, P., and Linser, R. (2018). Microsecond Timescale Protein Dynamics: a Combined Solid-State NMR Approach. *ChemPhysChem* 19, 34–39. doi:10.1002/cphc.201701238

Rovó, P., Smith, C. A., Gauto, D., De Groot, B. L., Schanda, P., and Linser, R. (2019). Mechanistic Insights into Microsecond Time-Scale Motion of Solid Proteins Using Complementary 15N and 1H Relaxation Dispersion Techniques. *J. Am. Chem. Soc.* 141, 858–869. doi:10.1021/jacs.8b09258

Saalwächter, K., and Fischbach, I. (2002). The Application of MAS Recoupling Methods in the Intermediate Motional Regime. *J. Magn. Reson.* 157, 17–30. doi:10.1006/jmre.2002.2552

Siemer, A. B., Huang, K.-Y., and Mcdermott, A. E. (2010). Protein-Ice Interaction of an Antifreeze Protein Observed With Solid-State NMR. *Proc. Natl. Acad. Sci. USA* 107, 17580–17585. doi:10.1073/pnas.1009369107

Tošner, Z., Purea, A., Struppe, J. O., Wegner, S., Engelke, F., Glaser, S. J., et al. (2017). Radiofrequency Fields in MAS Solid State NMR Probes. *J. Magn. Reson.* 284, 20–32. doi:10.1016/j.jmr.2017.09.002

Vallurupalli, P., Bouvignies, G., and Kay, L. E. (2012). Studying "invisible" Excited Protein States in Slow Exchange With a Major State Conformation. *J. Am. Chem. Soc.* 134, 8148–8161. doi:10.1021/ja3001419

Van der Wel, P. C. A. (2017). Insights into Protein Misfolding and Aggregation Enabled by Solid-State NMR Spectroscopy. *Solid State. Nucl. Magn. Reson.* 88, 1–14. doi:10.1016/j.ssnmr.2017.10.001

Vold, R. R., and Vold, R. L. (1991). "Deuterium Relaxation in Molecular Solids," in *Advances in Magnetic and Optical Resonance*. Editor W. Warren (San Diego: Acadenic Press), 85–171. doi:10.1016/b978-0-12-025516-0.50006-1

Vold, R. L., Hoatson, G. L., Vugmeyster, L., Ostrovsky, D., and De Castro, P. J. (2009). Solid State Deuteron Relaxation Time Anisotropy Measured with Multiple echo Acquisition. *Phys. Chem. Chem. Phys.* 11, 7008–7012. doi:10.1039/b907343d

Vold, R. R. (1994). "Deuterium NMR Studies of Dynamics in Solids and Liquid Crystals," in *Nuclear Magnetic Resonance Probes of Molecular Dynamics*. Editor R. Tycko (Dordrecht: Kluwer academic Publishers), 27–112. doi:10.1007/978-94-011-1410-3_2

Vugmeyster, L., and Ostrovsky, D. (2019). Deuterium Rotating Frame NMR Relaxation Measurements in the Solid State under Static Conditions for Quantification of Dynamics. *Chemphyschem* 20, 333–342. doi:10.1002/cphc.201800454

Vugmeyster, L., Au, D. F., Ostrovsky, D., and Fu, R. (2019). Deuteron Solid-State NMR Relaxation Measurements Reveal Two Distinct Conformational Exchange Processes in the Disordered N-Terminal Domain of Amyloid-β Fibrils. *ChemPhysChem* 20, 1680–1689. doi:10.1002/cphc.201900363

Vugmeyster, L., Ostrovsky, D., and Fu, R. (2020). Deuteron Quadrupolar Chemical Exchange Saturation Transfer (Q-CEST) Solid-State NMR for Static Powder Samples: Approach and Applications to Amyloid-β Fibrils. *ChemPhysChem* 21, 220–231. doi:10.1002/cphc.201901053

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 25 August 2021
doi: 10.3389/fmolb.2021.727053

Structural Polymorphism of Chitin and Chitosan in Fungal Cell Walls From Solid-State NMR and Principal Component Analysis

Liyanage D. Fernando[1‡], Malitha C. Dickwella Widanage[1‡], Jackson Penfield[2], Andrew S. Lipton[3], Nancy Washton[3], Jean-Paul Latgé[4†], Ping Wang[5], Liqun Zhang[2] and Tuo Wang[1]*

[1]Department of Chemistry, Louisiana State University, Baton Rouge, LA, United States, [2]Department of Chemical Engineering, Tennessee Technological University, Cookeville, TN, United States, [3]Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, United States, [4]Unité des Aspergillus, Département de Mycologie, Institut Pasteur, Paris, France, [5]Department of Microbiology, Immunology and Parasitology, Louisiana State University Health Sciences Center, New Orleans, LA, United States

Chitin is a major carbohydrate component of the fungal cell wall and a promising target for novel antifungal agents. However, it is technically challenging to characterize the structure of this polymer in native cell walls. Here, we recorded and compared ^{13}C chemical shifts of chitin using isotopically enriched cells of six *Aspergillus*, *Rhizopus*, and *Candida* strains, with data interpretation assisted by principal component analysis (PCA) and linear discriminant analysis (LDA) methods. The structure of chitin is found to be intrinsically heterogeneous, with peak multiplicity detected in each sample and distinct fingerprints observed across fungal species. Fungal chitin exhibits partial similarity to the model structures of α- and γ-allomorphs; therefore, chitin structure is not significantly affected by interactions with other cell wall components. Addition of antifungal drugs and salts did not significantly perturb the chemical shifts, revealing the structural resistance of chitin to external stress. In addition, the structure of the deacetylated form, chitosan, was found to resemble a relaxed two-fold helix conformation. This study provides high-resolution information on the structure of chitin and chitosan in their cellular contexts. The method is applicable to the analysis of other complex carbohydrates and polymer composites.

Keywords: chitin, chitosan, solid-state NMR, fungi, cell wall, *Aspergillus*, *Candida*, principal component analysis

OPEN ACCESS

Edited by:
Józef Romuald Lewandowski, University of Warwick, United Kingdom

Reviewed by:
Ray Dupree, University of Warwick, United Kingdom
Valerie Booth, Memorial University of Newfoundland, Canada

****Correspondence:***
Tuo Wang
tuowang@lsu.edu

†Present address:
Institute of Molecular biology and Biotechnology, University of Crete, Heraklion, Greece

‡These authors have contributed equally to this work

Specialty section:
This article was submitted to Structural Biology, a section of the journal Frontiers in Molecular Biosciences

Received: *17 June 2021*
Accepted: *10 August 2021*
Published: *25 August 2021*

Citation:
Fernando LD, Dickwella Widanage MC, Penfield J, Lipton AS, Washton N, Latgé J-P, Wang P, Zhang L and Wang T (2021) Structural Polymorphism of Chitin and Chitosan in Fungal Cell Walls From Solid-State NMR and Principal Component Analysis. Front. Mol. Biosci. 8:727053. doi: 10.3389/fmolb.2021.727053

INTRODUCTION

Chitin is the second-most abundant biopolymer in nature, only behind cellulose. Widely distributed in different organisms, chitin is often found as a supportive and protective component of the body armor (namely the exoskeleton) in arthropods and the cell walls of fungi and some algal species (Pillai et al., 2009; Rinaudo, 2006). The structures of chitin and its largely deacetylated form called chitosan have similarity to the organization of cellulose (Heux et al., 2000; Jarvis, 2003; Okuyama et al., 2000; Rinaudo, 2006; Saito et al., 1987). All these three polysaccharides are linear polymers of β-1,4-linked glucoses or their amide derivatives. Structurally, the hydroxyl group at position C-2 of a glucopyranose unit is replaced by an acetamido or an amino group, changing to the

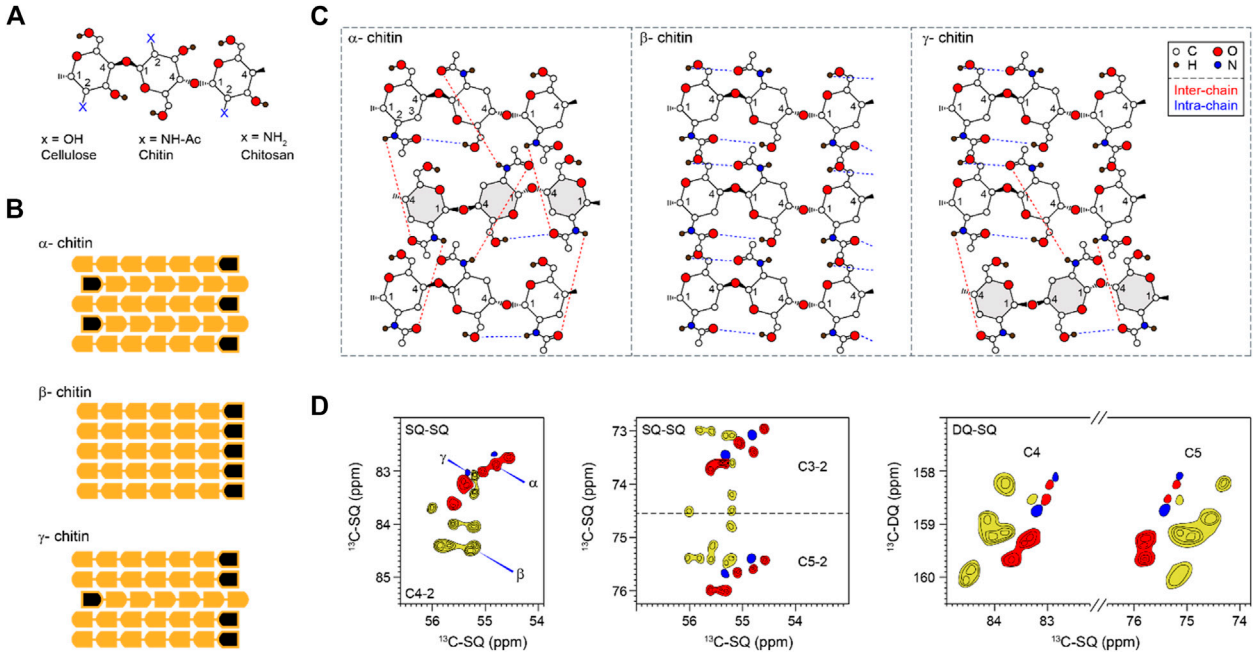


FIGURE 1 | Representative structures and NMR signals of chitin. **(A)** Substitutions at the C2 position for chitin and chitosan. **(B)** Polymorphic types (α, β, and γ) of chitin showing different chain orientations. Black marks denote the non-reducing ends of chains. **(C)** Hydrogen-bonding patterns of different chitin allomorphs. Blue and red dash lines indicate intra-chain and inter-chain hydrogen bonds, respectively. The antiparallel chains in α and γ chitins are in grey. The hydroxyl at C3 is not shown to make the structure less complex. The structural schemes are adapted from (Rudall, 1963; Kameda et al., 2005; Sawada et al., 2012a). **(D)** 2D ^{13}C-^{13}C correlation spectra simulated using literature-reported chemical shifts on model samples (**Supplementary Table S2**). Representative C4-2 and C3/5-2 regions were shown for single-quantum (SQ)-SQ correlation spectra. The C4 and C5 region is also shown for a double-quantum (DQ)-SQ correlation spectrum. α, β, and γ are represented in red, yellow, and blue respectively. Contour lines represent the number of data sets used (and the number of overlapped peaks).

N-acetylglucosamine (GlcNAc) unit in chitin and the glucosamine (GlcN) residue in chitosan (**Figure 1A**). Chitin and chitosan, especially the latter, have also drawn tremendous attention due to their promising applications as polymer scaffolds for tissue engineering, wound dressing, drug delivery, and pharmaceuticals (Jayakumar et al., 2010).

The amide and carbonyl groups in chitins drive the formation of hydrogen bonds and crystalline fibrils. X-ray crystallography has reported three chitin allomorphs, with substantial variation in the chain orientation and the hydrogen-bonding pattern (Sikorski et al., 2009; Yui et al., 2007). Adjacent chains are packed in an antiparallel or parallel way in the α- and β-forms, respectively (**Figure 1B**). The third type of structure, γ-chitin, can be considered as a mixture of parallel and antiparallel packings, but sometimes it is treated simply as a variant of the α-allomorph (Rinaudo, 2006). The structure of α-chitin is stabilized simultaneously by intra-chain O-H. . .O and inter-chain N-H. . .O hydrogen bonding (**Figure 1C**) (Kameda et al., 2005; Deringer et al., 2016). The former is a hydrogen bond consistently observed in all three allomorphs. The latter is relatively rare in the γ-form and is absent in the β-chitin (Kameda et al., 2004; Sawada et al., 2012a; Sawada et al., 2012b). The coexistence of inter- and intra-chain hydrogen bonds has made α-chitin the most stable, ordered, and tightly packed structure, widely found in arthropods, Porifera, Bryozoa, and fungi (Ehrlich et al., 2007; Ehrlich et al., 2017). β- and γ-allomorphs are less common: the former can be found in diatoms and cephalopods, while the latter was reported in beetles and loligo species (Brunner et al., 2009; Kaya et al., 2017). The currently available information on chitin structure was obtained using highly crystalline materials isolated and purified mainly from marine sources. Although chitin is also a major fungal polysaccharide (Erwig and Gow, 2016; Gow et al., 2017), our understanding of its structural characteristics in the fungal cell wall remains inadequate.

Biochemical assays have revealed that chitin, β-glucan, and mannan are held together by covalent linkages in the human pathogen *Aspergillus fumigatus*, forming the core of the cell wall (Latge, 2007; Latgé and Chamilos, 2020). This structural module is resistant to alkali treatment and therefore has been proposed as the central scaffold of fungal cell walls (Latgé et al., 2017). Recently, we have employed high-resolution solid-state NMR methods to investigate the structure of biomolecules in the intact cells of *A. fumigatus* (Kang et al., 2018; Zhao et al., 2020). Unexpectedly, we identified three major types (and in total eleven subtypes) of GlcNAc units, as resolved from their distinct ^{13}C and ^{15}N chemical shifts, which are indicators of structural variations (Kang et al., 2018). These chitin forms were found to be extensively associated with each other inside chitin microfibrils as shown by their strong inter-residue interactions. These findings have unveiled the surprisingly high structural polymorphism of chitin in its cellular environment and raised

three unresolved questions related to the chitin structure. First, is the structure of chitin in the fungal cell wall similar to the crystallographic structures determined using standard samples? Second, is there any dependence between the chitin structure and the fungal type? Third, is chitin structure modulated by external stresses such as antifungal drugs and hypersaline environments?

To answer these questions, we compared the ^{13}C chemical shifts of chitins identified in the cells prepared from three *Aspergillus* species (*Aspergillus fumigatus, A. nidulans,* and *A. sydowii*), *Rhizopus delemar*, and two *Candida* pathogens (*C. albicans* and *C. auris*), following exposure to various antifungal drugs and salt concentrations. All these fungal species investigated here are significant human pathogens causing life-threatening infections in immunodeficient individuals and known to display different chitin composition in their cell walls (Brown et al., 2012; Latge and Calderone, 2006). Root mean square deviation (RMSD) heatmap, principal component analysis (PCA), and linear discriminant analysis (LDA) of chemical shifts were performed for the comparison of 62 chitin forms. Most fungal chitins align well with literature-reported α- and γ-allomorphs but deviate substantially from the β-form. The structure of chitin proved robust, remaining unaffected even under high salinity or in the presence of antifungal drugs, caspofungin and amphotericin B (AmB). In addition, chitosan was also identified in *R. delemar* and *A. sydowii*. Comparison of the literature-reported and our observed chemical shifts showed that most chitosan molecules are closely related to the Type-II salt model compound that has a relaxed two-fold conformational structure. This study presents a widely applicable research strategy for evaluating the structure of cellular carbohydrates and provides the structural basis for developing chitin-targeting antifungal agents.

MATERIALS AND METHODS

Preparation of ^{13}C, ^{15}N-Labeled Fungal Cells

In total, nine ^{13}C,^{15}N-labeled samples were prepared for six fungal species including *A. fumigatus, A. nidulans, A. sydowii, C. albicans*, *C. auris,* and *R. delemar* following a recently established protocol (Kirui et al., 2019). To examine the potential effect of antifungal drugs on chitin structure, three parallel batches were prepared for *A. fumigatus:* without drug, with caspofungin (2.5 μg/ml: above the minimum inhibitory concentration), and with AmB (2.5 μg/ml). To examine if salt concentration and osmotic pressure affect chitin structure, two batches of materials were prepared for the seawater inhabitant *A. sydowii*, with 0.5 and 2.0 M NaCl to represent optimal and high salinity conditions, respectively (Perez-LIano et al., 2020). Briefly, uniformly ^{13}C,^{15}N-labeled materials were obtained by culturing the fungi in modified minimum liquid media containing ^{13}C-glucose as the sole carbon source. The nitrogen sources are different for various fungal species, with ^{15}N-sodium nitrate for *A. fumigatus* and *A. nidulans*, ^{15}N-ammonium nitrate for *A. sydowii*, and ^{15}N-ammonium sulfate for *C. albicans, C. auris,* and *R. delemar*. All these species are able to grow on inorganic nitrogen sources and were cultivated alternatively on ammonium or nitrate salts. The cultures were incubated at the optimum temperatures of 25–31°C for respective fungal species. The culture duration was 3 days for *A. fumigatus, A. nidulans*, *R. delemar*, *C. albicans*, and *C. auris*, and 7 days for *A. sydowii*. Fungal materials were then collected by centrifugation at 7,000 × g for 20 min. The harvested fungal pellets were washed thoroughly using phosphate buffered saline (pH 7.4) to remove small molecules and reduce the ion concentration. For each sample, approximately 30 mg of the hydrated whole-cell material was packed into a 3.2 mm magic-angle spinning (MAS) rotor for solid-state NMR characterization.

Solid-State NMR Experiments

All the high-resolution solid-state NMR data were collected on a Bruker 800 MHz (18.8 Tesla) Bruker Avance III HD spectrometer at the National High Magnetic Field Laboratory (Tallahassee, FL) and a Varian VNMRS 850 MHz (19.9 Tesla) spectrometer at the Environmental Molecular Sciences Laboratory (EMSL; Richland, WA). The experiments were conducted in 3.2 mm MAS HCN probes under 12–13.5 kHz MAS at 290–293 K. The ^{13}C chemical shifts were externally referenced to the adamantane CH_2 signal at 38.48 ppm on the tetramethylsilane scale. The ^{15}N chemical shifts were referred externally through the methionine nitrogen peak (127.88 ppm) in the model peptide formyl-Met-Leu-Phe (MLF). Typical ^{1}H radiofrequency field strengths 50–83 kHz and 50–62.5 kHz for ^{13}C. The ^{13}C chemical shifts were recorded using the 2D Dipolar-Assisted Rotational Resonance (DARR) experiment with a 100-ms mixing time and the 2D ^{13}C-^{13}C COmbined R2$^{\nu}_{n}$ -Driven (CORD) sequence with a 53-ms mixing time (Hou et al., 2013). 2D ^{15}N-^{13}C N(CA)CX heteronuclear correlation spectra were measured to detect chitin amide signals (Pauli et al., 2001). The N(CA)CX spectrum was recorded using a 0.6-ms ^{1}H-^{15}N cross polarization (CP), a 5-ms ^{15}N-^{13}C CP contact, and a 100-ms DARR mixing time. The experimental and processing parameters used for 2D ^{13}C-^{13}C and ^{13}C-^{15}N spectra are summarized in **Supplementary Table S1**. Resonance assignment was facilitated by comparison with previously reported chemical shifts indexed in a carbohydrate database (Kang et al., 2020), which distinguish chitin from glucans and other nitrogenated polysaccharides. To compare the chemical shift differences in different chitin forms observed in fungi and from different model samples, a heat map was constructed from the root-mean-square deviation (RMSD) values calculated using the comparison of the literature-reported and observed chemical shifts with normalization by the total number of carbon atoms in a monomer (i.e., 8 for chitin carbons of C1-C6, CO, and CH_3). Similar approaches are also used for comparing different forms of fungal chitin. Good correlations give low RMSD values.

Principal Component Analysis and Linear Discriminant Analysis

We conducted PCA to facilitate the analysis of the species- and condition-dependent data of chitin chemical shifts. PCA is a form of multivariate analysis employed to reduce the many correlated variables to just a few new variables (the principal components)

that describe most of the variation in a dataset. Recently, PCA has been successfully employed to provide valuable insights on chemical shift data for small molecules (Tasic et al., 2002) and proteins (Kazumasa and Goto, 2007; Sakurai et al., 2019). The PCA was first conducted using MATLAB for the entire dataset from both the available literature and freshly measured spectra (**Supplementary Tables S2, S3**). A 62 × 8 matrix was composed, with each row representing a different chitin form identified in the NMR spectra, and each column corresponding to the chemical shifts observed for a ^{13}C atom at a particular location in the chitin structure. Similarly, PCA was also run separately for three subsets of chitin chemical shift data to compare 1) only the data from fungal chitin, 2) drug-free and drug-treated samples, and 3) optimal and high salinity conditions. For each PCA, a singular value decomposition (SVD) analysis was performed on the data matrix to generate orthogonal eigenvectors with values known as "loadings" or "PCA coefficients" arranged in a matrix by column. Loadings are normalized and used to describe the contribution made by each chemical shift, while the magnitude of the eigenvector shows how much of the variance in the data is explained by each eigenvector. The largest eigenvector defines the axis principal component 1 (PC1), and the next largest one defines PC2, etc. Each NMR dataset can be given a score based on the loadings and is projected onto the principal axes to show how the chemical conditions in that sample affect the observed chemical shifts. Samples of molecules within similar chemical environments are expected to cluster together in the "PC-space" if the dimension-reduction is successful. Because loadings describe a linear combination of the original variables, the relationship between the mean-centered data, score, and loadings are the matrix product: [PC score] = [data] × [PC loadings].

In addition, we performed linear discriminant analysis (LDA) to identify the factor that distinguishes the chitins produced in *Candida* species and other fungi. LDA was performed on the PCA scores, which provide linear discriminant (LD) loadings and LD scores. The scores of observations in separate classes fall approximately into a normal distribution with as little overlap with other classes as possible. The addition of more classes requires additional linear discriminants. Similar to PCA, the relationship between LD scores and LD loadings is: [LD score] = [data] × [LD loadings].

RESULTS AND DISCUSSION

Solid-State NMR Fingerprints of Chitin in Fungal Cell Walls

Solid-state NMR has been widely applied to differentiate the hydrogen-bonding patterns, identify the type of chitin, and determine the degree of acetylation of chitin and chitosan (by tracking the intensities of CO and CH_3 peaks) in model samples (Tanner et al., 1990; Heux et al., 2000; Kameda et al., 2004; Kono, 2004; Kasaai, 2010; King et al., 2017). The spectroscopic signatures of model chitin allomorphs are summarized in 2D^{13}C-^{13}C correlation spectra simulated and plotted using literature-reported chemical shifts (**Supplementary Table S2**) (Jang et al., 2004; Kono, 2004; Tanner et al., 1990; Brunner et al., 2009; Kaya et al., 2017; Kolbe et al., 2021) (**Figure 1D**). α-chitin has its C3 and C5 peaks distributed as two separated regions (72–73.7 and 75.4–76 ppm like a doublet) while most β-chitins have characteristic C3 and C5 signals sharply clustered in the 74–76 ppm region. The signals of γ-chitin are mixed with those of α- and β-allomorphs, with better alignment to the α-form. The same trend is retained in the double-quantum (DQ)-SQ correlation spectrum. The INADEQUATE spectrum, with an example shown in **Supplementary Figure S1**, was not explicitly used in this study but have been frequently measured for characterizing cellular samples.

Different from the model compounds, analysis of cellular systems using solid-state NMR spectroscopy has remained challenging due to the coexistence of a large variety of biomolecules, whose signals often exhibit significant overlap (Poulhazan et al., 2018; Narasimhan et al., 2019; Kelly et al., 2020; Zhao et al., 2020; Reif et al., 2021). Fortunately, the presence of nitrogen in the amide group has made chitin chemically unique among the structural polysaccharides in the cell wall. At the same time, the nitrogenated sugars in the intracellular content have already been filtered out using CP-based methods, which remove the signals of mobile sugars but selectively highlight the stiff molecules in the cell wall. The ^{15}N chemical shifts (~128 ppm) and the unique ^{13}C chemical shift of the nitrogen-linked carbon 2 (54–56 ppm) are the characteristic signals of chitin for initiating the resonance assignment. High-resolution 2D ^{13}C-^{13}C and ^{15}N-^{13}C correlation spectra collected on freshly prepared *A. fumigatus* mycelia resolved the signals of six major types of chitins (type a–f), together with two forms with some carbon sites being ambiguously assigned (types g and h) (**Figure 2A**; **Supplementary Figure S2**). The ^{13}C full width at half maximum (FWHM) linewidth is in the range of 0.5–0.7 ppm for the chitin in native cell walls.

The C5-C4 and C3-C4 cross-peaks showed comparable spectral patterns among the three *Aspergillus* samples, indicative of structural similarity (**Figure 2B**). *R. delemar*, however, had more extensive signals in this spectral region due to its uniquely high content of chitin and chitosan molecules (Melida et al., 2015; Ghormade et al., 2017; Lecointe et al., 2019). The spectra of *C. albicans* and *C. auris* looked alike, but their spectral patterns differ from the other filamentous fungi studied. Comparing to α and γ chitin, the characteristic signals of β-chitin were less overlapped with the spectra of all the fungal samples. Chains in β-chitin are arranged in a parallel way, with only intramolecular H-bonds. This results in a unique and less tightly packed structure for β-chitin, which is swollen in water and exhibiting high reactivity. Most of the literature-reported chemical shifts (**Supplementary Table S2**) from the α-allomorph are enclosed in the spectral envelope of the fungal samples studied here. Still, the expected signals of β-chitin mostly fell out of the spectral region.

Caspofungin inhibits the β-1,3-glucan synthesis, but when above the minimal inhibitory concentration, it causes a paradoxical effect enhancing the production of chitin to compensate for the loss of β-1,3-glucan (Loiko and

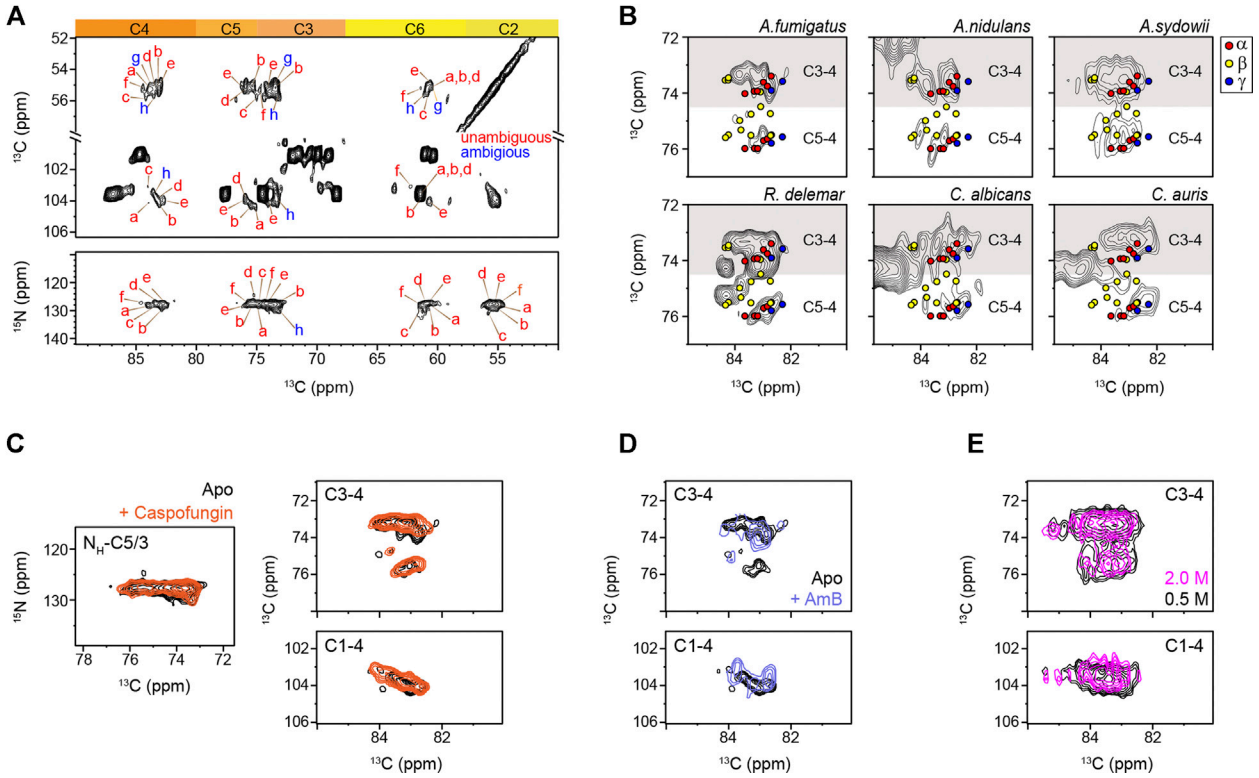


FIGURE 2 | Peak multiplicity of chitin in different fungi. **(A)** Representative signals of different chitin types in *A. fumigatus*. ^{13}C-^{13}C (top) and ^{15}N-^{13}C (bottom) correlation spectra resolved different forms of chitin molecules. Chitin forms with all carbon sites unambiguously resolved are labeled in red (types a–f), while the ambiguous forms are in blue (types g and h), with the ambiguous (partially resolved) carbon sites underlined. **(B)** Comparison of chitin signals in different fungi. The C5-C4 and C3-C4 regions are shown. Colored dots denote the data from three crystalline forms of chitin: α-chitin (red), β-chitin (yellow), and γ-chitin (blue). **(C)** 2D ^{15}N-^{13}C and ^{13}C-^{13}C correlation spectra of *A. fumigatus* without drug (apo; black) and with caspofungin treatment (orange). **(D)** 2D ^{13}C-^{13}C spectra of *A. fumigatus* without drug (apo; black) and with amphotericin B (AmB; blue). (**E**) Overlay of 2D ^{13}C-^{13}C correlation spectra collected on two *A. sydowii* samples cultured with 0.5 M NaCl (black) and 2 M NaCl (magenta).

Wagener, 2017). Consistently, the intensities of chitin peaks were enhanced relative to other cell wall components (**Supplementary Figure S3**), but no major changes were observed in the chemical shifts (**Figure 2C**). Therefore, the increased amount of chitin has insignificant effects on the structure of this molecule. Similarly, the addition of AmB that targets ergosterol in fungal membranes (Anderson et al., 2014) only redistributed the intensities among chitin subtypes without inducing new signals (**Figure 2D**) The robustness of the chitin structure is further confirmed by the comparable signals observed in the saprophytic *A. sydowii* samples cultured with either optimal or high salinities (**Figure 2E**) (Perez-LIano et al., 2020). Although chitin structure altered moderately among different fungi, it remained resistant to these external stresses (**Supplementary Tables S4, S5**). These observations are not surprising because AmB and caspofungin do not directly target chitin. Nikkomycin is the most notable chitin synthesis inhibitor and is thus of significant interest for further investigations (Steinbach and Stevens, 2003; Nix et al., 2009; Li et al., 2019). Recently combinatorial biosynthetic approaches have been used integrating echinocandin and chitin inhibitors which show potent antifungal activity (Li et al., 2019).

Comparison of Chitin Structures Using Chemical Shift Analysis

We compared the ^{13}C chemical shifts obtained on the 45 chitin forms in nine fungal samples (**Supplementary Table S3**) with the 17 datasets reported in the literature (**Supplementary Table S2**) (Jang et al., 2004; Kono, 2004; Tanner et al., 1990; Brunner et al., 2009; Kaya et al., 2017; Kolbe et al., 2021), generating a chemical shift RMSD heatmap (**Figure 3**). The 45 subforms identified and assigned in the intact fungal cell wall include eight chitin forms (a–h) in drug-free *A. fumigatus*, six forms (a–f) in each of the two *A. fumigatus* samples treated with either caspofungin or amphotericin B, four chitin forms (a′-d′) in *A. nidulans*, five forms (A–E) in each of the two *A. sydowii* samples cultured with 0.5 M or 2 M NaCl, three chitin forms (i,k) in *R. delemar*, and four chitin forms (l-o) in each of the two *Candida* samples. Each of the 765 comparisons was represented by an RMSD value based on 16 ^{13}C chemical shifts of C1-C6, CO, and CH_3 from two different chitin forms. Similar methods have been used to compare the

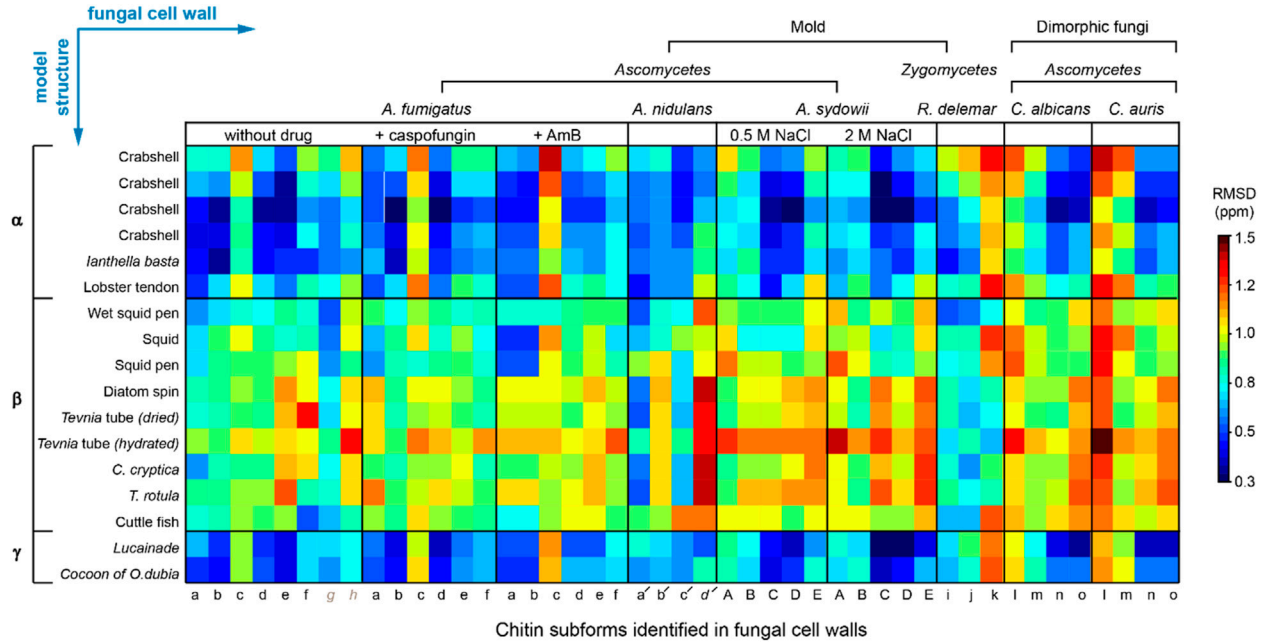


FIGURE 3 | ^{13}C chemical shift RMSD map comparing chitin structure. Data were compared between the observed 45 chitin forms in nine fungal cell walls (*x*-axis) and the three crystalline forms reported by literature (*y*-axis). Data from six fungal species were shown, including three species of *Ascomycetes* (*A. fumigatus*, *A nidulans* and *A. sydowii*), a sample from Zygomycetes (*R. delemar*), and two Ascomycetes yeast species (*C. albicans* and *C. auris*). Most chitin types showed similarity to α-chitin form. The color scale is shown, with units of ppm. Good correlation with RMSD less than 0.5 ppm (within NMR linewidth) are in dark blue. The forms with certain ambiguous carbon sites are labeled in italics and grey. The chemical shift values used for the analysis are provided in **Supplementary Tables S2, S3**.

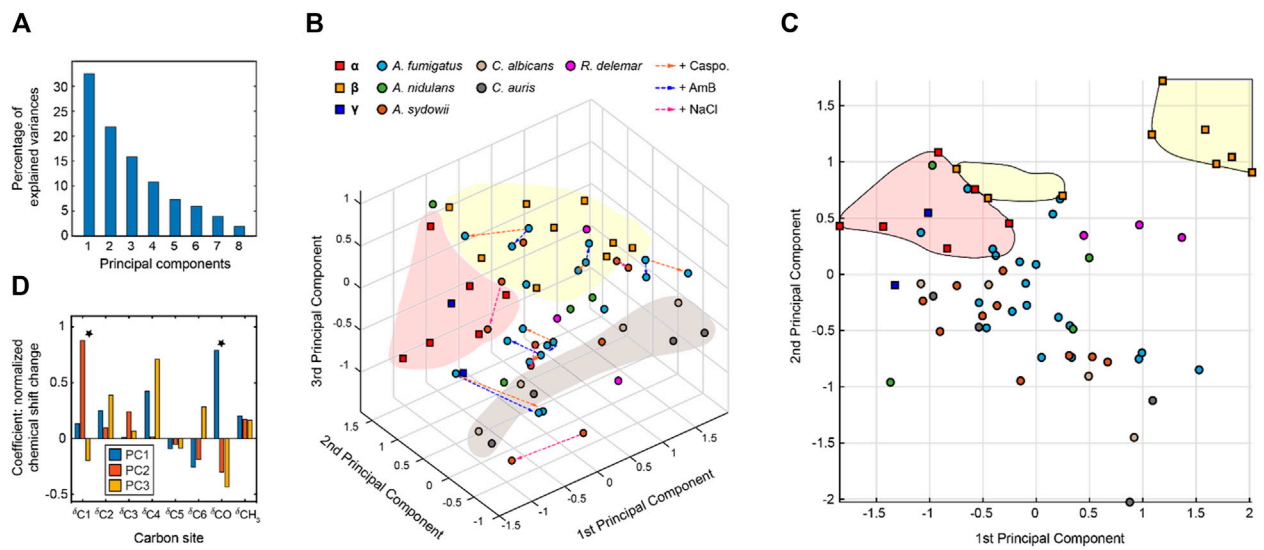


FIGURE 4 | Principal component analysis of chitin chemical shifts. **(A)** Variance explained by each principal component (PC). **(B)** PCA scores for chitin NMR chemical shifts projected onto principal component 1 (PC1) vs. PC2 vs. PC3. Model chitin allomorphs (α, β, and γ-types) are shown using squares while chitin forms identified in fungal cell walls are presented using circles. Shaded regions in red and yellow are used to enclose α- and β-type chitins, respectively. The shaded region in grey mainly contains data from *Candida* species. Data from different model samples and fungal species are color-coded. Arrows in orange, blue, and magenta represents the changes induced by caspofungin (Caspo.), the amphotericin B (AmB), and NaCl (from 0.5 to 2.0 M), respectively. **(C)** PCA scores of chitin chemical shifts projected onto PC1 and PC2 proving a better visualization of most chitin forms. **(D)** Loadings for each PC. Asterisks indicate the most pronounced differences for PC1 and PC2.

NMR data collected on other fibrillar biomolecules including cellulose and amyloid fibrils (Elkins et al., 2016; Wang et al., 2016; Qiang et al., 2017). We found that fungal chitin correlated relatively well with α-chitin. Small RMSD values below the spectroscopic resolution (0.5 ppm) were observed for some datasets of *A. fumigatus* and *C. albicans*. Reasonable correlations between the cell wall chitin and the γ-chitin model structure were also noted, which can be understood by treating γ-chitin as a derivative of α-chitin due to their structural similarities. In contrast, β-chitins failed to correlate with our observations, with large RMSD typically in the range of 0.7–1.6 ppm. Exceptions were observed for *R. delemar* (**Figure 3**), suggesting the formation of structurally unique chitin domains in this fungus.

The NMR chemical shift data were subjected to PCA. As a dimension-reduction analysis tool, a useful PCA result necessitates that the importance of each consecutive PC declines rapidly. PCs are constructed by the SVD algorithm in an unsupervised manner, beginning with a new axis that maximizes the variance of all data points when projected onto it, then constructing orthogonal axes according to the same criteria. The eigenvectors returned from the SVD calculation are shown in **Figure 4A**, with the sum normalized to 100, showing the percent of variance in the data explained by each PC. With the first three PCs explaining 70% of the variance in the data, a safe majority of the variance is now explained in those three variables, and the first three PCs should be able to account for the major factors contributed to the chemical shift.

The 3D PCA score plot composed using the first three PCs (**Figure 4B**) illustrates the relationship between each chitin sample in the PC space. Consistent with the heatmap representation, principal component 1 (PC1) primarily differentiated the α and β chitin standards, with the γ-chitin standards more closely associated with the former. This is more clearly recognizable in the 2D presentation of PC1 vs. PC2 (**Figure 4C**), that the spreading of α and β chitins are on the negative side and positive sides of PC1, respectively. We only observed a relatively small amount of stretching of β-chitins to the negative side. In addition, γ-chitin are distributed mostly to the α-chitin side. Therefore, it is likely that PC1 can sense the difference in hydrogen bonding and chain-packing. This is confirmed by the loadings where the first principal component experiences the most significant change at the carbonyl group (**Figure 4D**). Together, PC1 and PC2 can clearly resolve most forms of β-chitins as a self-isolated group. *Candida* chitins and β-chitins show up on the two extreme positions of PC2, with scores distributed somewhat evenly between −1 and 1 of PC2 and PC3.

The PCA loadings shown in **Figure 4D** are the weight given to each original variable (chemical shifts) in the linear combination that defines each PC, from which one can gather the relative magnitude and direction (as indicated by the sign) of change in those variables expected to occur over positive displacement in the respective PC score. The loadings show that while PC1 is mostly concerned with the carbonyl, PC2 focuses on the C1 atom, while PC3 and PC1 focus on C4 atom that also (together with C1) participates in the glycosidic linkages of chitin molecule.

To only focus on fungal chitin, we conducted a separate PCA by excluding the data from α, β, and γ model allomorphs (**Supplementary Figure S4**). PCA scores for all fungi chitins indicate that similarities between chitins within a single fungal species are sparse, as many allomorphs of the same species can be found at opposite extremes of both PC1 and PC2, accounting together for almost 60% of variation. Two other PCAs were conducted to respectively focus on the effect of drug and salt conditions (**Supplementary Figures S5, S6**). It should be noted that the changes caused by antifungal drugs and increased salinity are trivial when compared with the large structural dispersion of native chitin molecules.

In addition, partial structural similarities were noted for some chitin subtypes residing in different fungal strains (**Figure 5A**). For *A. fumigatus*, a few reasonably good correlations can be found with *A. nidulans* and *A. sydowii, Candida* species, and *R. delemar*. These observations revealed the partial alignment of chitin structure in different species. The best correlation was found between the type-d chitin of *A. fumigatus* and the type-D form of *A. sydowii*, with a small RMSD (0.19 ppm) well below the NMR linewidth. Just like the *Aspergillus* samples, *R. delemar* is also a filamentous fungus, but it exhibited only a single modest correlation with *Aspergillus* species, indicating the structural uniqueness of the chitin produced in *R. delemar*.

The *Candida* samples prepared in this study were grown only as a yeast form. The two *Candida* species are highly similar to each other, with small RMSD values (0.16–0.32 ppm) when comparing each type of chitin between two *Candida* species. For example, the RMSD is 0.16 for the comparison of type-m chitins in *C. albicans* and *C. auris*. The RMSD is similarly good for the comparisons of type-n (0.21 ppm) and type-l (0.26 ppm) chitins, and only slightly larger for the type-m form (0.32 ppm). In contrast, the filamentous fungi (*Aspergillus* and *Rhizopus* species) studied here only exhibited partial similarities to the *Candida* species. It is possible that filamentous fungi require for their hypha a specific form of chitin because the strength to hold the tube-shaped mycelium should be different and stronger than holding a balloon shape like a yeast.

The results also aligned with the number and families of chitin synthase (CHS) genes seen in these species. In yeasts (*Candida* and *Saccharomyces* for example), 3 to 4 CHS genes have been encountered belonging to the families I, II and IV. In *Aspergillus* and *Rhizopus*, however, 9 to 23 genes have been found and they not only belong to the three classes (I, II and IV) that were also identified in yeasts, but also have contributions from additional classes (III, V, VII, VI or VIII) (Lenardon et al., 2007; Ma et al., 2009; Muszkieta et al., 2014).

To directly identify the structural factor that differentiates the chitin types in yeasts and filamentous fungi, we conducted linear discriminant analysis (LDA). Different from the PCA method described above, LDA is a supervised learning method. LDA can pinpoint the variables that distinguish between the observations that have already been arranged into classes by their properties of interest. Here, we categorized the data into two separate classes to distinguish *Candida* strains (grown as yeasts) from other fungal species (grown as mycelium), which produced a linear discriminant (**Figure 5B**). Their probability distributions (**Figure 5C**) only overlapped slightly, and the loadings (**Figure 5D**) indicated that *Candida* chitin and the chitins of

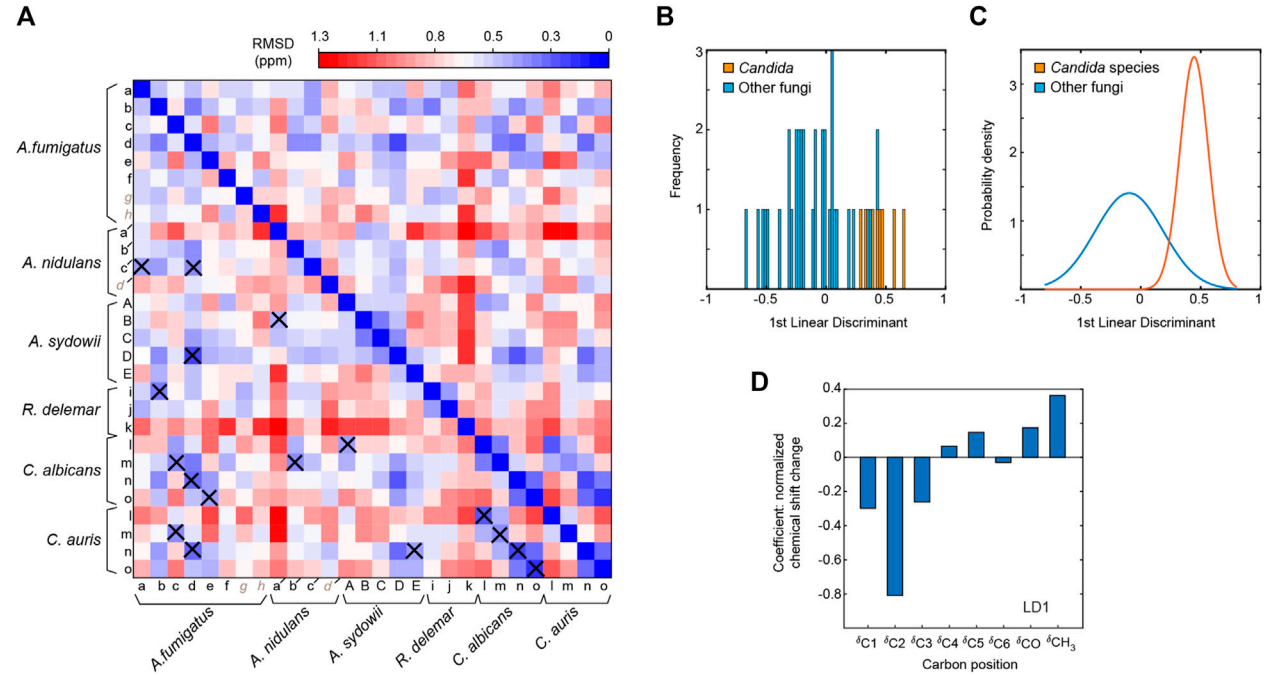


FIGURE 5 | Comparison of chitin forms identified in different fungal species. **(A)** Chemical shift RMSD heatmap comparing the chitin forms observed in different fungi. Good correlations with RMSD of less than 0.5 ppm are highlighted using crosses. **(B)** Linear discriminant analysis with *Candida* fungi (*C. albicans* and *C. auris*) classified differently from other fungal species, with linear discriminant 1 (LD1) scores shown in a histogram. This panel mainly shows the frequency in which LD1 scores fall into a particular range (the width of each bar). **(C)** Gaussian probability distributions of LD1 scores. The *Candida* data falls into a smaller range than the other fungi, therefore, there is a much higher probability that a *Candida* species will fall near their statistical mean. **(D)** LD1 loadings corresponding to the chemical shifts of each carbon site.

other fungal species could be best distinguished by the chemical shifts of C2 and CH_3, thus revealing the key sites for tracking fungal chitin structure.

The results provided three structural implications. First, the structure of chitin is highly polymorphic in fungal cell walls. At this moment, it is unclear whether the observed polymorphism is related to the diverse groups of chitin synthases involved in the biosynthesis of this polymer, which should be further investigated using functional genomics and spectroscopic approaches. It also raised a major question on the individual function of all the CHS genes (>20 genes in the *Zygomycetes*). This study raises unanswered questions about the function of the different classes of chitin synthases in the cell wall structuration. Based on the ssNMR data presented here it does suggest that all CHS synthesized a chitin with very similar structure. The actual biological role of each CHS should be totally dependent on the cellular localization of each synthase in the cell wall as recently suggested (Walker et al., 2013).

Second, the model structures of α-chitins, as characterized using the highly crystalline material isolated and purified from marine sources, are remarkably preserved among different fungi. This is intriguing as the interactions with other polysaccharides, often by covalent linkages in fungal cell walls (Gow et al., 2017), did not substantially perturb the structure of chitin. This result agrees with the low number of linkages identified biochemically in the β-1,3-glucan-chitin core of *A. fumigatus* cell wall and the poor growth phenotype resulting from the deletion of the CRH genes coding for the glycosyltransferases that are responsible for forming glucan-chitin linkages (Latgé et al., 2017). It is a supplementary argument to suggest that these chitin-glucan covalent connections might not be structurally important for the building of the cell wall.

Third, the structure of chitin is resistant to environmental stimuli, such as non-chitin-focused drug treatment as well as hypersaline environment and osmotic pressure. The structural robustness of chitin and its central role in mechanically supporting the cell wall confirmed the suitability of chitin as a potential target for the development of novel antifungal compounds. It also indicated that the increase in chitin concentration in the cell wall is a survival response, which is not depending on the stress proposed. At this moment, it remains unknown how to reconcile the microscopic structure of the different chitin microfibrils seen in electron microscopy (Lenardon et al., 2007; Lenardon et al., 2010; Muszkieta et al., 2014) with the atomic level ssNMR data.

Spectroscopic and Structural Features of Fungal Chitosan

Deacetylation of chitin leads to chitosan. Chitosan exists in a semicrystalline form in solids but can be solubilized by acidic solutions. In the fungal cell wall, chitosan has been proposed to serve as a backbone to bind other biomolecules, such as

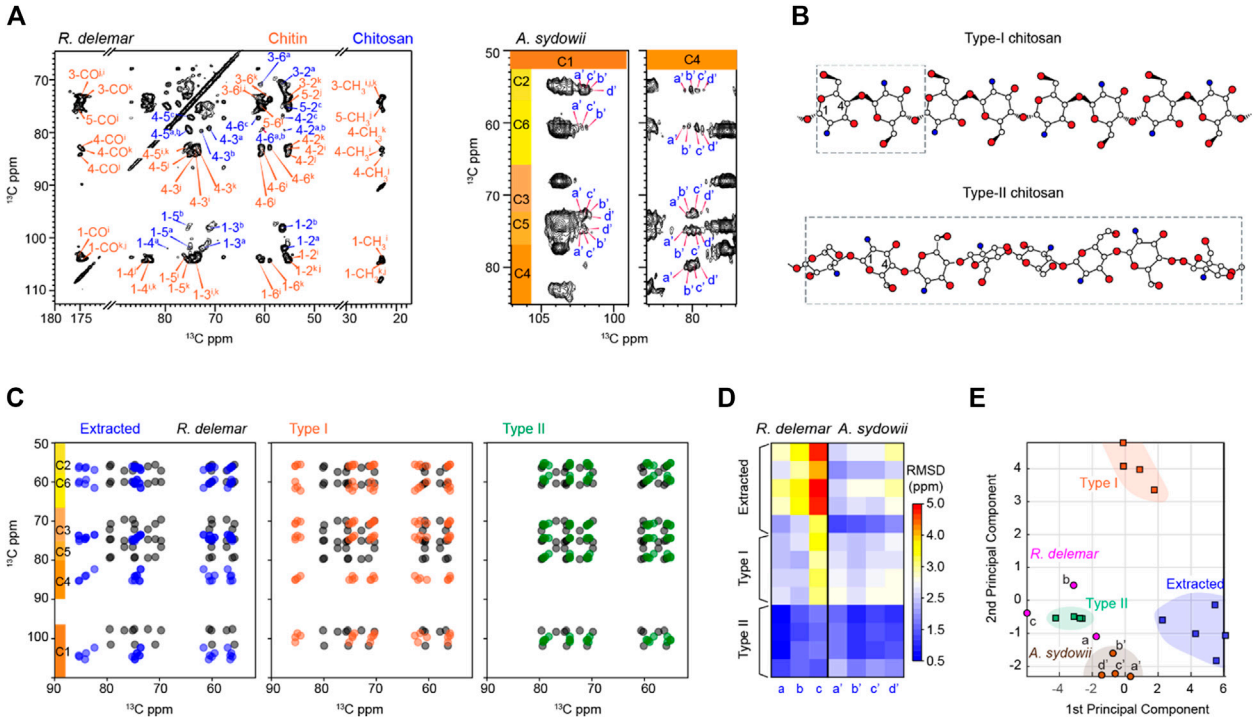


FIGURE 6 | *R. delemar* and *A. sydowii* cell walls are rich in chitosan. **(A)** Representative 2D ^{13}C-^{13}C CORD spectrum of *R. delemar* and *A. sydowii* cells showing many sets of chitosan signals (blue). **(B)** Representative structures of Type-I and Type-II chitosan molecules. Nitrogen (blue), oxygen (red), and carbon (white) atoms are shown but hydrogen atoms are not included for simplicity. The repeating units are shown in dash line boxes. Structure schemes are adapted from (Okuyama et al., 2000). **(C)** Simulated spectra of *R. delemar* chitosan (black) overlaid with the literature-reported chitosan forms including extracted chitosan **(blue; left panel)**, Type-I salts with inorganic acids **(orange; middle panel)**, and Type-II salts with inorganic acids **(green; right panel)**. **(D)** ^{13}C chemical shift RMSD map for the comparison between fungal cell wall chitosan (*X*-axis) and model samples (*Y*-axis). The color scale unit is ppm. **(E)** PCA scores of chitosan. The data analyzed include Type-I (orange squares) and Type-II (green squares) salts with inorganic acids, extracted chitosan (blue square), as well as the chitosan forms identified in *R. delemar* (magenta circles) and *A. sydowii* (brown circles).

dityrosines or melanin (Chrissian et al., 2020a; Chrissian et al., 2020b). The NMR signals of chitosan are resolved from those of chitin by the absence of CH_3 and CO peaks at 22 and 174 ppm (**Supplementary Figure S7**). The substantial modification in the chemical structure and the hydrogen-bonding patterns induce unique chemical shifts at most carbon sites as shown by **Figure 6A**. The structures of two major chitosan forms, Types I and II salts with inorganic acids, have been reported (**Figure 6B**), which exhibited different helical conformations (Saito et al., 1987; Ogawa et al., 2004; Franca et al., 2008). Type-I chitosan has a fully extended two-fold helical structure. The repeating unit of type-II chitosan is four times longer than that of type-I, with a relaxed two-fold helix and a tetrasaccharide repeat in a helical asymmetric unit. Overlay of the spectra predicted using the chemical shifts available in the literature and our dataset revealed that *R. delemar* chitosan could not structurally align with those extracted from various sources such as crab tendon, crab shell, and shrimp shell (**Figure 6C**). The same discrepancy was also present for the Type-I compound, but a better correlation was observed with the Type-II structure.

No chitosan signal was observed in these fresh *A. fumigatus* samples. This is in agreement with a recent genomic study which indicates that the deletion of all deacetylase genes in *A. fumigatus* does not lead to any significant growth phenotype (Mouyna et al., 2020). Interestingly, the occurrence of a significant amount of chitosan in xerophilic *Aspergillus* species may indicate that the fungus synthesizes chitosan to make the cell wall more flexible to fight against the increase in osmotic pressure.

The type-c chitosan in *R. delemar* exhibited bad correlations with the chitosan prepared using extracted chitin (RMSD ~5 ppm) and Type-I chitosan in inorganic salt. RMSD values as large as that should be originated from totally different structures. In contrast, the type-c chitosan correlated reasonably with Type-II chitosan (RMSD <1.5 ppm) (**Figure 6D**). Similar trends were observed for the other two types (a and b) of chitosan molecules. For example, comparison of chitosan-a in *R. delemar with* Type-II model structures gave very small RMSDs of 0.6–0.8 ppm. The results indicate that chitin chitosan differs from the extracted forms or the Type I structure, but closely resembles the Type-II structure. This trend was projected in the RMSD heatmap of ^{13}C chemical shifts for both *R. delemar* and *A. sydowii* (**Figure 6D**). In the PCA plot, chitosan signals were separated remarkably well by the first two principal components, which account for 89% of the variation in the data (**Figure 6E**; **Supplementary Figure S8**). *R. delemar* and *A. sydowii* samples shared more in common with the Type-II

chitosan standards but lacked structural similarity to the Type-I standard and extracted chitosan. Therefore, chitosan in the fungal cell wall only has moderate correlations to the Type-II standard structure.

It should be noted that the RMSD values between different chitosan forms are substantially larger than those calculated for chitin. The NMR data actually suggest a new type of chitosan structure that is different from those previously characterized. It is also intriguing that chitosan molecules in extracted materials and intact fungal cell walls are structurally distinct. A possible reason is the solubilization and extraction procedures used in previous studies might have restructured this molecule before subjection to structural characterization. For example, alkali treatment was known to induce chitin deacetylation. The distinct organization of molecules in arthropods and fungi, as well as the potential difference in the degree of deacetylation (Liu et al., 2019), might also contribute to the observed discrepancy. This differs from the case of chitin, which is an insoluble polymer and often found in the crosslinked core of fungal cell walls, thus being more resistant to isolation and processing procedures. More in-depth investigations are needed to identify the biochemical reason driving the structural complexity of chitosan and to fully understand its function-related structures in fungal cell walls.

CONCLUSION

The high-resolution dataset enabled by solid-state NMR spectroscopy has made it possible to analyze and compare the structural features of cell wall polysaccharides using statistical approaches. Such protocols will accommodate the rapidly expanding ssNMR dataset and open new research avenues related to the structural investigations of cellular and extracellular biomolecules as well as natural and artificial biomaterials (Arnold et al., 2018; Bougault et al., 2019; Kang et al., 2019; Ehren et al., 2020; Kelly et al., 2020). The polymorphic structure of chitin and its resistance to external stress was determined in fungal species of biomedical and environmental significance. This information has the potential to facilitate the development of antifungal strategies targeting the unique structures of chitin or its biosynthesis.

DATA AVAILABILITY STATEMENT

The datasets presented in the study are included in the article and **Supplementary Material**. Additional data can be requested from the corresponding author.

AUTHOR CONTRIBUTIONS

LF and MW were responsible for sample preparation; LF, MW, AL, and NW are responsible for conducting the NMR experiments; JP and LZ were responsible for PCA and LDA analysis. LF, MW, JP-L, PW and TW were responsible for analyzing and interpreting the NMR results. PW and TW were responsible for conceptualization and supervision. All authors were responsible for writing the manuscript.

FUNDING

This work was supported by the National Institute of Health (NIH) grant AI149289. A portion of this work was performed at the National High Magnetic Field Laboratory, which is supported by the National Science Foundation Cooperative Agreement No. DMR-1644779 and the State of Florida. A portion of the NMR dataset was collected at the Environmental Molecular Sciences Laboratory (grid.436923.9), a DOE Office of Science scientific user facility sponsored by the Department of Energy's Office of Biological and Environmental Research and located at PNNL under contract DE-AC05-76RL01830.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.727053/full#supplementary-material

REFERENCES

Anderson, T. M., Clay, M. C., Cioffi, A. G., Diaz, K. A., HIsao, G. S., Tuttle, M. D., et al. 2014). Amphotericin Forms an Extramembranous and Fungicidal Sterol Sponge. *Nat. Chem. Biol.*, 10, 400–406.doi:10.1038/nchembio.1496

Arnold, A. A., Bourgouin, J.-P., Genard, B., Warschawski, D. E., Tremblay, R., and Marcotte, I. (2018). Whole Cell Solid-State NMR Study of Chlamydomonas Reinhardtii Microalgae. *J. Biomol. NMR*, 70(2), 123–131.doi:10.1007/s10858-018-0164-7

Bougault, C., Ayala, I., Vollmer, W., Simorre, J.-P., and Schanda, P. (2019). Studying Intact Bacterial Peptidoglycan by Proton-Detected NMR Spectroscopy at 100 kHz MAS Frequency. *J. Struct. Biol.*, 206(1), 66–72.doi:10.1016/j.jsb.2018.07.009

Brown, G. D., Denning, D. W., Gow, N. A., Levitz, S. M., Netea, M. G., and White, T. C. (2012). Hidden Killers: Human Fungal Infections. *Sci. Transl. Med.*, 4(165, 165rv13).doi:10.1126/scitranslmed.3004404

Brunner, E., Ehrlich, H., Schupp, P., Hedrich, R., Hunoldt, S., Kammer, M., et al. 2009). Chitin-based Scaffolds Are an Integral Part of the Skeleton of the marine Demosponge Ianthella Basta. *J. Struct. Biol.*, 168(3), 539–547.doi:10.1016/j.jsb.2009.06.018

Chrissian, C., Camacho, E., Fu, M. S., Prados-Rosales, R., Chatterjee, S., Cordero, R. J. B., et al. 2020a). Melanin Deposition in Two Cryptococcus Species Depends on Cell-wall Composition and Flexibility. *J. Biol. Chem.*, 295(7), 1815, 1828.doi:10.1074/jbc.ra119.011949

Chrissian, C., Lin, C. P.-C., Camacho, E., Casadevall, A., Neiman, A. M., and Stark, R. E. (2020b). Unconventional Constituents and Shared Molecular Architecture of the Melanized Cell Wall of C. Neoformans and Spore Wall of *S. cerevisiae*. *JoF*, 6(4), 329.doi:10.3390/jof6040329

Deringer, V. L., Englert, U., and Dronskowski, R. (2016). Nature, Strength, and Cooperativity of the Hydrogen-Bonding Network in α-Chitin. *Biomacromolecules*, 17(3), 996–1003.doi:10.1021/acs.biomac.5b01653

Ehren, H. L., Appels, F. V. W., Houben, K., Renault, M. A. M., Wösten, H. A. B., and Baldus, M. (2020). Characterization of the Cell wall of a Mushroom Forming Fungus at Atomic Resolution Using Solid-State NMR Spectroscopy. *Cell Surf*, 6, 100046.doi:10.1016/j.tcsw.2020.100046

Ehrlich, H., Bazhenov, V. V., Debitus, C., de Voogd, N., Galli, R., Tsurkan, M. V., et al. 2017). Isolation and Identification of Chitin from Heavy Mineralized Skeleton of Suberea Clavata (Verongida: Demospongiae: Porifera) marine Demosponge. *Int. J. Biol. Macromolecules*, 104(PtB), 1706–1712.doi:10.1016/j.ijbiomac.2017.01.141

Ehrlich, H., Krautter, M., Hanke, T., Simon, P., Knieb, C., Heinemann, S., et al. (2007). First Evidence of the Presence of Chitin in Skeletons of marine Sponges. Part II. Glass Sponges (Hexactinellida: Porifera). *J. Exp. Zool.*, 308B(4), 473–483.doi:10.1002/jez.b.21174

Elkins, M. R., Wang, T., Nick, M., Jo, H., Lemmin, T., Prusiner, S. B., et al. 2016). Structural Polymorphism of Alzheimer's β-Amyloid Fibrils as Controlled by an E22 Switch: A Solid-State NMR Study. *J. Am. Chem. Soc.*, 138, 9840–9852.doi:10.1021/jacs.6b03715

Erwig, L. P., and Gow, N. A. (2016). Interactions of Fungal Pathogens with Phagocytes. *Nat. Rev. Microbiol.*, 14(4), 163–176.doi:10.1038/nrmicro.2015.21

Franca, E. F., Lins, R. D., Freitas, L. C. G., and Straatsma, T. P. (2008). Characterization of Chitin and Chitosan Molecular Structure in Aqueous Solution. *J. Chem. Theor. Comput.*, 4(12), 2141–2149.doi:10.1021/ct8002964

Ghormade, V., Pathan, E. K., and Deshpande, M. V. (2017). Can Fungi Compete with marine Sources for Chitosan Production? *Int. J. Biol. Macromolecules*, 104, 1415–1421.doi:10.1016/j.ijbiomac.2017.01.112

Gow, N. A. R., Latge, J. P., and Munro, C. A. (2017). The Fungal Cell Wall: Structure, Biosynthesis, and Function. *Microbiol. Spectr.*, 5(3).doi:10.1128/microbiolspec.FUNK-0035-2016

Heux, L., Brugnerotto, J., Desbrières, J., Versali, M.-F., and Rinaudo, M. (2000). Solid State NMR for Determination of Degree of Acetylation of Chitin and Chitosan. *Biomacromolecules*, 1(4), 746–751.doi:10.1021/bm000070y

Hou, G., Yan, S., Trébosc, J., Amoureux, J.-P., and Polenova, T. (2013). Broadband Homonuclear Correlation Spectroscopy Driven by Combined R2nv Sequences under Fast Magic Angle Spinning for NMR Structural Analysis of Organic and Biological Solids. *J. Magn. Reson.*, 232, 18–30.doi:10.1016/j.jmr.2013.04.009

Jang, M. K., Kong, B. G., Jeong, Y. I., Lee, C. H., and Nah, J. W. (2004). Physicochemical Characterization of α-Chitin, β-Chitin, and γ-Chitin Separated from Natural Resources. *J. Polym. Sci. A Polym. Chem.*, 42(14), 3423–3432.

Jarvis, M. (2003). Cellulose Stacks up. *Nature*, 426(6967), 611–612.doi:10.1038/426611a

Jayakumar, R., Menon, D., Manzoor, K., Nair, S. V., and Tamura, H. (2010). Biomedical Applications of Chitin and Chitosan Based Nanomaterials-A Short Review. *Carbohydr. Polym.*, 82(2), 227–232.doi:10.1016/j.carbpol.2010.04.074

Kameda, T., Miyazawa, M., Ono, H., and Yoshida, M. (2004). Hydrogen Bonding Structure and Stability of Alpha-Chitin Studied by 13C Solid-State NMR. *Macromol. Biosci.*, 5(2), 103–106.doi:10.1002/mabi.200400142

Kameda, T., Miyazawa, M., Ono, H., and Yoshida, M. (2005). Hydrogen Bonding Structure and Stability Of?-Chitin Studied by13C Solid-State NMR. *Macromol. Biosci.*, 5(2), 103–106.doi:10.1002/mabi.200400142

Kang, X., Kirui, A., Dickwella Widanage, M. C., Mentink-Vigier, F., Cosgrove, D. J., and Wang, T. (2019). Lignin-polysaccharide Interactions in Plant Secondary Cell walls Revealed by Solid-State NMR. *Nat. Commun.*, 10(1), 347.doi:10.1038/s41467-018-08252-0

Kang, X., Kirui, A., Muszyński, A., Widanage, M. C. D., Chen, A., Azadi, P., et al. 2018). Molecular Architecture of Fungal Cell walls Revealed by Solid-State NMR. *Nat. Commun.*, 9, 2747.doi:10.1038/s41467-018-05199-0

Kang, X., Zhao, W., Dickwella Widanage, M. C., Kirui, A., Ozdenvar, U., and Wang, T. (2020). CCMRD: A Solid-State NMR Database for Complex Carbohydrates. *J. Biomol. NMR*, 74, 239–245.doi:10.1007/s10858-020-00304-2

Kasaai, M. R. (2010). Determination of the Degree of N-Acetylation for Chitin and Chitosan by Various NMR Spectroscopy Techniques: A Review. *Carbohydr. Polym.*, 79(4), 801–810.doi:10.1016/j.carbpol.2009.10.051

Kaya, M., Mujtaba, M., Ehrlich, H., Salaberria, A. M., Baran, T., Amemiya, C. T., et al. 2017). On Chemistry of γ-chitin. *Carbohydr. Polym.*, 176, 177–186.doi:10.1016/j.carbpol.2017.08.076

Kazumasa, S., and Goto, Y. (2007). Principal Component Analysis of the pH-dependent Conformational Transitions of Bovine β-lactoglobulin Monitored by Heteronuclear NMR. *Proc. Natl. Acad. Sci. USA*, 104(39), 15346–15351. doi:10.1073/pnas.0702112104

Kelly, J. E., Chrissian, C., and Stark, R. E. (2020). Tailoring NMR Experiments for Structural Characterization of Amorphous Biological Solids: A Practical Guide. *Solid State. Nucl. Magn. Reson.*, 109, 101686.doi:10.1016/j.ssnmr.2020.101686

King, C., Stein, R. S., Shamshina, J. L., and Rogers, R. D. (2017). Measuring the Purity of Chitin with a Clean, Quantitative Solid-State NMR Method. *ACS Sustain. Chem. Eng.*, 5(9), 8011–8016.doi:10.1021/acssuschemeng.7b01589

Kirui, A., Dickwella Widanage, M. C., Mentink-Vigier, F., Wang, P., Kang, X., and Wang, T. (2019). Preparation of Fungal and Plant Materials for Structural Elucidation Using Dynamic Nuclear Polarization Solid-State NMR. *J. Vis. Exp.* , 144, e59152.doi:10.3791/59152

Kolbe, F., Ehren, H. L., Kohrs, S., Butscher, D., Reiß, L., Baldus, M., et al. (2021). Solid-State NMR Spectroscopic Studies of 13C, 15N, 29Si-Enriched Biosilica from the Marine Diatom *Cyclotella cryptica*. *Disc. Mater.*, 1(1), 1–12.

Kono, H. (2004). Two-dimensional Magic Angle Spinning NMR Investigation of Naturally Occurring Chitins: Precise1H and13C Resonance Assignment of ?- and ?-chitin. *Biopolymers*, 75(3), 255–263.doi:10.1002/bip.20124

Latgé, J.-P., Beauvais, A., and Chamilos, G. (2017). The Cell Wall of the Human Fungal PathogenAspergillus Fumigatus: Biosynthesis, Organization, Immune Response, and Virulence. *Annu. Rev. Microbiol.*, 71, 99–116.doi:10.1146/annurev-micro-030117-020406

Latge, J. P., and Calderone, R. A. (2006). The Fungal Cell Wall. In *Growth, Differentiation and Sexuality* (pp. 73–104). Berlin: Springer,

Latgé, J. P., and Chamilos, G. (2020). Aspergillus fumigatus and Aspergillosis in 2019. *Clin. Microbiol. Rev.*, 33, e00140-00118. doi:10.1128/CMR.00140-18

Latgé, J. P. (2007). The Cell wall: a Carbohydrate armour for the Fungal Cell. *Mol. Microbiol.*, 66(2), 279–290.doi:10.1111/j.1365-2958.2007.05872.x

Lecointe, K., Cornu, M., Leroy, J., Coulon, P., and Sendid, B. (2019). Polysaccharides Cell Wall Architecture of Mucorales. *Front. Microbiol.*, 10, 469.doi:10.3389/fmicb.2019.00469

Lenardon, M. D., Munro, C. A., and Gow, N. A. (2010). Chitin Synthesis and Fungal Pathogenesis. *Curr. Opin. Microbiol.*, 13(4), 416–423.doi:10.1016/j.mib.2010.05.002

Lenardon, M. D., Whitton, R. K., Munro, C. A., Marshall, D., and Gow, N. A. R. (2007). Individual Chitin Synthase Enzymes Synthesize Microfibrils of Differing Structure at Specific Locations in the Candida Albicans Cell wall. *Mol. Microbiol.*, 66(5), 1164–1173.doi:10.1111/j.1365-2958.2007.05990.x

Li, Y., Sun, H., Zhu, X., Bian, C., Wang, Y., and Si, S. (2019). Identification of New Antifungal Agents Targeting Chitin Synthesis by a Chemical-Genetic Method. *Molecules*, 24(17), 3155.doi:10.3390/molecules24173155

Liu, W., Ma, Y., Ai, L., Li, W., Li, W., Li, H., et al. 2019). Enzymatic Degradation of Nanosized Chitin Whiskers with Different Degrees of Deacetylation. *ACS Biomater. Sci. Eng.*, 5(10), 5316–5326.doi:10.1021/acsbiomaterials.9b00796

Loiko, V., and Wagener, J. (2017). The Paradoxical Effect of Echinocandins in Aspergillus fumigatus Relies on Recovery of the β-1,3-Glucan Synthase Fks1 *Antimicrob. Agents Chemother.*, 61, e01690-01616.doi:10.1128/AAC.01690-16

Ma, L.-J., Ibrahim, A. S., Skory, C., Grabherr, M. G., Burger, G., Butler, M., et al. 2009). Genomic Analysis of the Basal Lineage Fungus Rhizopus Oryzae Reveals a Whole-Genome Duplication. *Plos Genet.* 5(7), e1000549.doi:10.1371/journal.pgen.1000549

Mélida, H., Sain, D., Stajich, J. E., and Bulone, V. (2015). Deciphering the Uniqueness of Mucoromycotina Cell walls by Combining Biochemical and Phylogenomic Approaches *Environ. Microbiol.*, 17(5), 1649–1662.doi:10.1111/1462-2920.12601

Mouyna, I., Dellière, S., Beauvais, A., Gravelat, F., Snarr, B., Lehoux, M., et al. 2020). What Are the Functions of Chitin Deacetylases in Aspergillus fumigatus? *Front. Cel. Infect. Microbiol.*, 10, 28.doi:10.3389/fcimb.2020.00028

Muszkieta, L., Aimanianda, V., Mellado, E., Gribaldo, S., Alcàzar-Fuoli, L., Szewczyk, E., et al. 2014). Deciphering the Role of the Chitin Synthase Families 1 and 2 in Thein Vivoandin Vitrogrowth ofAspergillus Fumigatusby Multiple Gene Targeting Deletion. *Cell Microbiol.*, 16(12), 1784–1805.doi:10.1111/cmi.12326

Narasimhan, S., Scherpe, S., Lucini Paioni, A., van der Zwan, J., Folkers, G. E., Ovaa, H., et al. (2019). DNP-Supported Solid-State NMR Spectroscopy of Proteins inside Mammalian Cells. *Angew. Chem. Int. Ed.*, 58(37), 12969–12973.doi:10.1002/anie.201903246

Nix, D. E., Swezey, R. R., Hector, R., and Galgiani, J. N. (2009). Pharmacokinetics of Nikkomycin Z after Single Rising Oral Doses. *Antimicrob. Agents Chemother.*, 53(6), 2517–2521.doi:10.1128/aac.01609-08

Ogawa, K., Yui, T., and Okuyama, K. (2004). Three D Structures of Chitosan. *Int. J. Biol. Macromolecules*, 34, 1–8.doi:10.1016/j.ijbiomac.2003.11.002

Okuyama, K., Noguchi, K., Kanenari, M., Egawa, T., Osawa, K., and Ogawa, K. (2000). Structural Diversity of Chitosan and its Complexes. *Carbohydr. Polym.*, 41(3), 237–247.doi:10.1016/s0144-8617(99)00142-3

Pauli, J., Baldus, M., van Rossum, B., de Groot, H., and Oschkinat, H. (2001). Backbone and Side-Chain13C and15N Signal Assignments of the α-Spectrin SH3 Domain by Magic Angle Spinning Solid-State NMR at 17.6 Tesla. *ChemBioChem.*, 2, 272–281.doi:10.1002/1439-7633(20010401)2:4<272:aid-cbic272>3.0.co;2-2

Perez-LIano, Y., Rodriguez-Pupo, E. C., Druzhinina, I. S., henthamara, K., Cai, F., Gude-Cimerman, N., et al. (2020). Stress Reshapes the Physiological Response ofHalophile Fungi to Salinity. *Cells*, 9(3), 525. doi:10.3390/cells9030525

Pillai, C. K. S., Paul, W., and Sharma, C. P. (2009). Chitin and Chitosan Polymers: Chemistry, Solubility and Fiber Formation. *Prog. Polym. Sci.*, 34, 641–678.doi:10.1016/j.progpolymsci.2009.04.001

Poulhazan, A., Arnold, A. A., Warschawski, D. E., and Marcotte, I. (2018). Unambiguous *Ex Situ* and in Cell 2D 13C Solid-State NMR Characterization of Starch and its Constituents. *Int. J. Mol. Sci.*, 19(12).doi:10.3390/ijms19123817

Qiang, W., Yau, W.-M., Lu, J.-X., Collinge, J., and Tycko, R. (2017). Structural Variation in Amyloid-β Fibrils from Alzheimer's Disease Clinical Subtypes. *Nature*, 541, 217–221.doi:10.1038/nature20814

Reif, B., Ashbrook, S. E., Emsley, L., and Hong, M. (2021). Solid-state NMR Spectroscopy. *Nat. Rev. Methods Primers*, 1, 2.doi:10.1038/s43586-020-00002-1

Rinaudo, M. (2006). Chitin and Chitosan: Properties and Applications. *Prog. Polym. Sci.*, 31(7), 603–632.doi:10.1016/j.progpolymsci.2006.06.001

Rudall, K. M. (1963). The Chitin/Protein Complexes of Insect Cuticles. In J. W. L. Beament, J. E. Treherne, and V. B. Wigglesworth (Eds.), *Advances in Insect Physiology* (pp. 257–313.): Academic Pressdoi:10.1016/s0065-2806(08)60177-0

Saito, H., Tabeta, R., and Ogawa, K. (1987). High-resolution Solid-State Carbon-13 NMR Study of Chitosan and its Salts with Acids: Conformational Characterization of Polymorphs and Helical Structures as Viewed from the Conformation-dependent Carbon-13 Chemical Shifts. *Macromolecules*, 20(10), 2424–2430.doi:10.1021/ma00176a017

Sakurai, K., Maeno, A., Lee, Y.-H., and Akasaka, K. (2019). Conformational Properties Relevant to the Amyloidogenicity of β2-Microglobulin Analyzed Using Pressure- and Salt-dependent Chemical Shift Data. *J. Phys. Chem. B*, 123, 836–844.doi:10.1021/acs.jpcb.8b11408

Sawada, D., Nishiyama, Y., Langan, P., Forsyth, V. T., Kimura, S., and Wada, M. (2012a). Direct Determination of the Hydrogen Bonding Arrangement in Anhydrous β-Chitin by Neutron Fiber Diffraction. *Biomacromolecules*, 13(1), 288–291.doi:10.1021/bm201512t

Sawada, D., Nishiyama, Y., Langan, P., Forsyth, V. T., Kimura, S., and Wada, M. (2012b). Water in Crystalline Fibers of Dihydrate β-Chitin Results in Unexpected Absence of Intramolecular Hydrogen Bonding. *PLOS one*, 7(6), e39376.doi:10.1371/journal.pone.0039376

Sikorski, P., Hori, R., and Wada, M. (2009). Revisit of α-Chitin Crystal Structure Using High Resolution X-ray Diffraction Data. *Biomacromolecules*, 10(5), 1100–1105.doi:10.1021/bm801251e

Steinbach, W. J., and Stevens, D. A. (2003). Review of Newer Antifungal and Immunomodulatory Strategies for Invasive Aspergillosis. *Clin. Infect. Dis.*, 37 Suppl. 3, S157–S187.doi:10.1086/376523

Tanner, S. F., Chanzy, H., Vincendon, M., Roux, J. C., and Gaill, F. (1990). High-resolution Solid-State Carbon-13 Nuclear Magnetic Resonance Study of Chitin. *Macromolecules*, 23(15), 3576–3583.doi:10.1021/ma00217a008

Tasic, L., Abraham, R. J., and Rittner, R. (2002). Substituent Effects on1H and13C NMR Chemical Shifts in ?-monosubstituted Ethyl Acetates: Principal Component Analysis and1H Chemical Shift Calculations. *Magn. Reson. Chem.*, 40, 449–454.doi:10.1002/mrc.1046

Walker, L. A., Lenardon, M. D., Preechasuth, K., Munro, C. A., and Gow, N. A. (2013). Cell wall Stress Induces Alternative Fungal Cytokinesis and Septation Strategies. *J. Cel Sci.*, 126(Pt 12), 2668–2677.doi:10.1242/jcs.118885

Wang, T., Yang, H., Kubicki, J. D., and Hong, M. (2016). Cellulose Structural Polymorphism in Plant Primary Cell Walls Investigated by High-Field 2D Solid-State NMR Spectroscopy and Density Functional Theory Calculations. *Biomacromolecules*, 17(6), 2210–2222.doi:10.1021/acs.biomac.6b00441

Yui, T., Taki, N., Sugiyama, J., and Hayashi, S. (2007). Exhaustive crystal Structure Search and crystal Modeling of β-chitin. *Int. J. Biol. Macromolecules*, 40(4), 336–344.doi:10.1016/j.ijbiomac.2006.08.017

Zhao, W., Fernando, L. D., Kirui, A., Deligey, F., and Wang, T. (2020). Solid-state NMR of Plant and Fungal Cell walls: A Critical Review. *Solid State. Nucl. Magn. Reson.*, 107, 101660.doi:10.1016/j.ssnmr.2020.101660

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

REVIEW
published: 11 October 2021
doi: 10.3389/fmolb.2021.741581

Emerging Contributions of Solid-State NMR Spectroscopy to Chromatin Structural Biology

*Bryce E. Ackermann and Galia T. Debelouchina**

Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA, United States

The eukaryotic genome is packaged into chromatin, a polymer of DNA and histone proteins that regulates gene expression and the spatial organization of nuclear content. The repetitive character of chromatin is diversified into rich layers of complexity that encompass DNA sequence, histone variants and post-translational modifications. Subtle molecular changes in these variables can often lead to global chromatin rearrangements that dictate entire gene programs with far reaching implications for development and disease. Decades of structural biology advances have revealed the complex relationship between chromatin structure, dynamics, interactions, and gene expression. Here, we focus on the emerging contributions of magic-angle spinning solid-state nuclear magnetic resonance spectroscopy (MAS NMR), a relative newcomer on the chromatin structural biology stage. Unique among structural biology techniques, MAS NMR is ideally suited to provide atomic level information regarding both the rigid and dynamic components of this complex and heterogenous biological polymer. In this review, we highlight the advantages MAS NMR can offer to chromatin structural biologists, discuss sample preparation strategies for structural analysis, summarize recent MAS NMR studies of chromatin structure and dynamics, and close by discussing how MAS NMR can be combined with state-of-the-art chemical biology tools to reconstitute and dissect complex chromatin environments.

Keywords: gene regulation, magic angle spinning, chemical biology, nucleosome dynamics, histone dynamics

OPEN ACCESS

Edited by:
Loren B. Andreas,
Max Planck Institute for Biophysical Chemistry, Germany

Reviewed by:
Christopher Jaroniec,
The Ohio State University, United States
Lars Nordenskiöld,
Nanyang Technological University, Singapore
Xiangyan Shi,
Shenzhen MSU-BIT University, China, in collaboration with reviewer LN

***Correspondence:**
Galia T. Debelouchina
gdebelouchina@ucsd.edu

Specialty section:
This article was submitted to Structural Biology, a section of the journal Frontiers in Molecular Biosciences

Received: *14 July 2021*
Accepted: *20 September 2021*
Published: *11 October 2021*

Citation:
Ackermann BE and Debelouchina GT (2021) Emerging Contributions of Solid-State NMR Spectroscopy to Chromatin Structural Biology. Front. Mol. Biosci. 8:741581. doi: 10.3389/fmolb.2021.741581

INTRODUCTION

In the nuclei of eukaryotic cells, approximately 2 m of DNA must be packaged and organized for efficient gene regulation and DNA replication (**Figure 1A**). On a global level, this is achieved by separation into active gene regions called euchromatin and silent gene compartments known as heterochromatin. A fundamental molecular building block for such organization is the nucleosome, a complex of histone proteins (2 each of H2A, H2B, H3, and H4) assembled into an octamer and wrapped by ~147 base pairs of DNA. These nucleosome units repeat along the length of each chromosome to generate a "'beads-on-a-string" polymer called chromatin (Woodcock et al., 1976) (**Figure 1B**). The interactions between histones and DNA are primarily electrostatic in nature, where the peripheral basic residues of the histone octamer intercalate into the phosphate backbone of wrapped DNA, while the dynamic histone termini (tails) transiently explore exposed DNA (**Figure 1C**). Since the majority of the genome is bound by histones, any modification of these interactions, however small, can alter genetic outcomes. Post-translational modifications (PTMs), for example, differentially mark histone tails to recruit specific histone readers (Kouzarides, 2007;

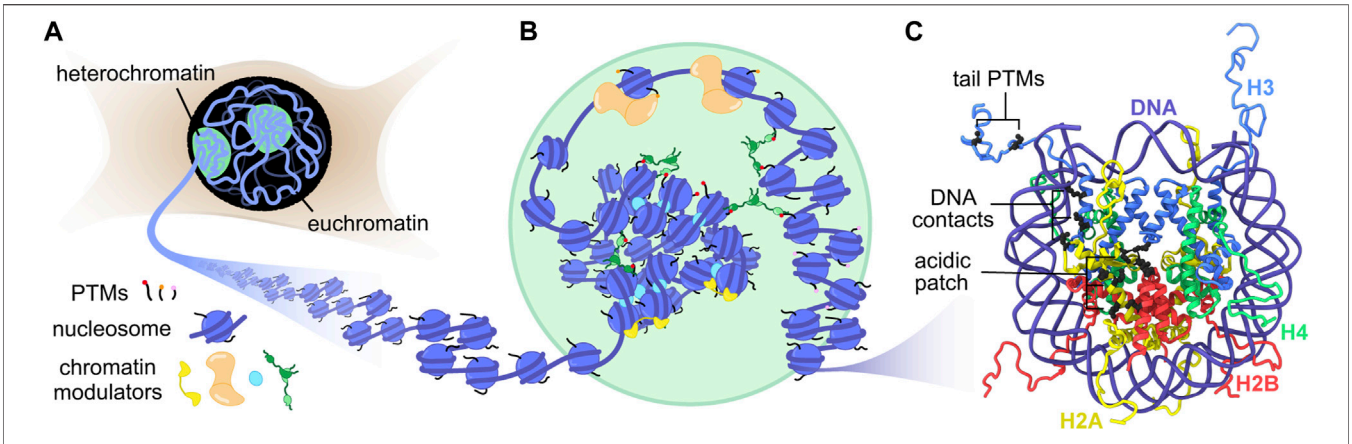


FIGURE 1 | Genome organization from the nucleus to the nucleosome. **(A)** The nucleus contains two distinct chromatin states, heterochromatin and euchromatin. Compact heterochromatin compartments may form by phase separation. **(B)** Chromatin fibers in different states contain distinct PTM signatures and interact with specific chromatin modulators. **(C)** The structure of the nucleosome with highlighted regions of interest for MAS NMR studies (PDB:1KX5) (Davey et al., 2002).

Prakash and Fournier, 2017) or to initiate DNA unwrapping (Bowman and Poirier, 2015). Macromolecular complexes that recognize such PTMs can further impact chromatin organization by cross-linking nucleosomes that are megabases apart in sequence (Rao et al., 2014; Strom et al., 2021), or by shifting the position of nucleosomes to expose new DNA sites for transcription initiation. Despite continuing progress towards determining the structure of chromatin in cells (Hsieh et al., 2015; Ricci et al., 2015; Nozaki et al., 2017; Ou et al., 2017; Risca et al., 2017; Cai et al., 2018; Xu et al., 2018; Ohno et al., 2019; Otterstrom et al., 2019; Krietenstein et al., 2020; Su et al., 2020), the impressive span of length scales involved, from small chemical modifications in the Ångstrom range to whole chromosome rearrangements on the micrometer scale, creates a tremendous challenge for structural biologists and biophysicists.

Bottom-up approaches using purified components are well suited to systematically probe the interplay between nucleosomes and chromatin structure on the Ångstrom and nanometer scale (Allis and Muir, 2011). The isolation of mononucleosomes from their polymer context enables high-resolution structural biology where the rigid histone core can be studied by X-ray crystallography and cryo-EM, and the histone tails by solution NMR spectroscopy (McGinty et al., 2016; Zhou et al., 2019). Virtually any chromatin and DNA process has now been mapped on the nucleosome, including chromatin remodeling, transcription, histone/DNA modification, gene repression, and DNA repair (Jang and Song, 2019; Zhou et al., 2019; McGinty and Tan, 2021; Min and Liu, 2021). These studies have highlighted several cornerstones in nucleosome recognition and modification. The H2A-H2B acidic patch, a lone cluster of glutamate and aspartate residues on the nucleosome surface, serves as a landing pad to anchor chromatin modulators as well as nearby nucleosomes (Kalashnikova et al., 2013; Chen et al., 2017; McGinty and Tan, 2021) (**Figure 1C**). The histone tails are dynamic, enabling access to both modification and recognition. PTMs and sequence variation of histones dictate the dynamics of nucleosome sliding and unwrapping. The histone-induced bending of nucleosomal DNA elicits unique recognition motifs for protein interaction.

Building upon these studies, the interactions of many adjacent nucleosomes can be addressed. The chromatin context is important for biomolecular recognition; some chromatin modulator complexes are much larger than a nucleosome and can sense nearby nucleosomes (Yang et al., 2006; He et al., 2020), many architectural proteins are multivalent and can simultaneously interact with several nucleosomes (Machida et al., 2018; Poepsel et al., 2018), and neighboring nucleosomes can stack atop each other, thereby competing with chromatin modulators for binding sites (Bilokapic et al., 2018; Sanulli et al., 2019; Alvarado et al., 2021). The fiber context is also necessary for packaging, as nucleosomes are strung together and densities get closer to those observed in cells (10–100 mg/ml) (Imai et al., 2017; Hancock, 2018; Kim and Guck, 2020), chromatin can undergo phase separation into a highly viscous solid-like material (Strickfaden et al., 2020). Phase separation is a promising model for cellular compartmentalization (and more recently chromatin compaction) that depends on a large number of transient multivalent interactions (Gibson et al., 2019; Sanulli et al., 2019). While this setting better represents the native nucleosome competition and the physical forces of compaction, the heterogeneity and density of such nucleosome polymers become intractable to most structural techniques.

NMR spectroscopy rises to the forefront of techniques uniquely capable of probing atomic structural and dynamic information for complex samples. Solution NMR of proteins benefits from fast molecular rotation to average line-broadening anisotropic interactions. However, larger proteins or protein complexes may tumble too slowly for efficient averaging and may require the tools of solid-state NMR. Magic-angle spinning solid-state NMR spectroscopy (MAS NMR) achieves partial averaging by rapid (tens of kilohertz) rotation of the sample at 54.7° (the magic angle) relative to the

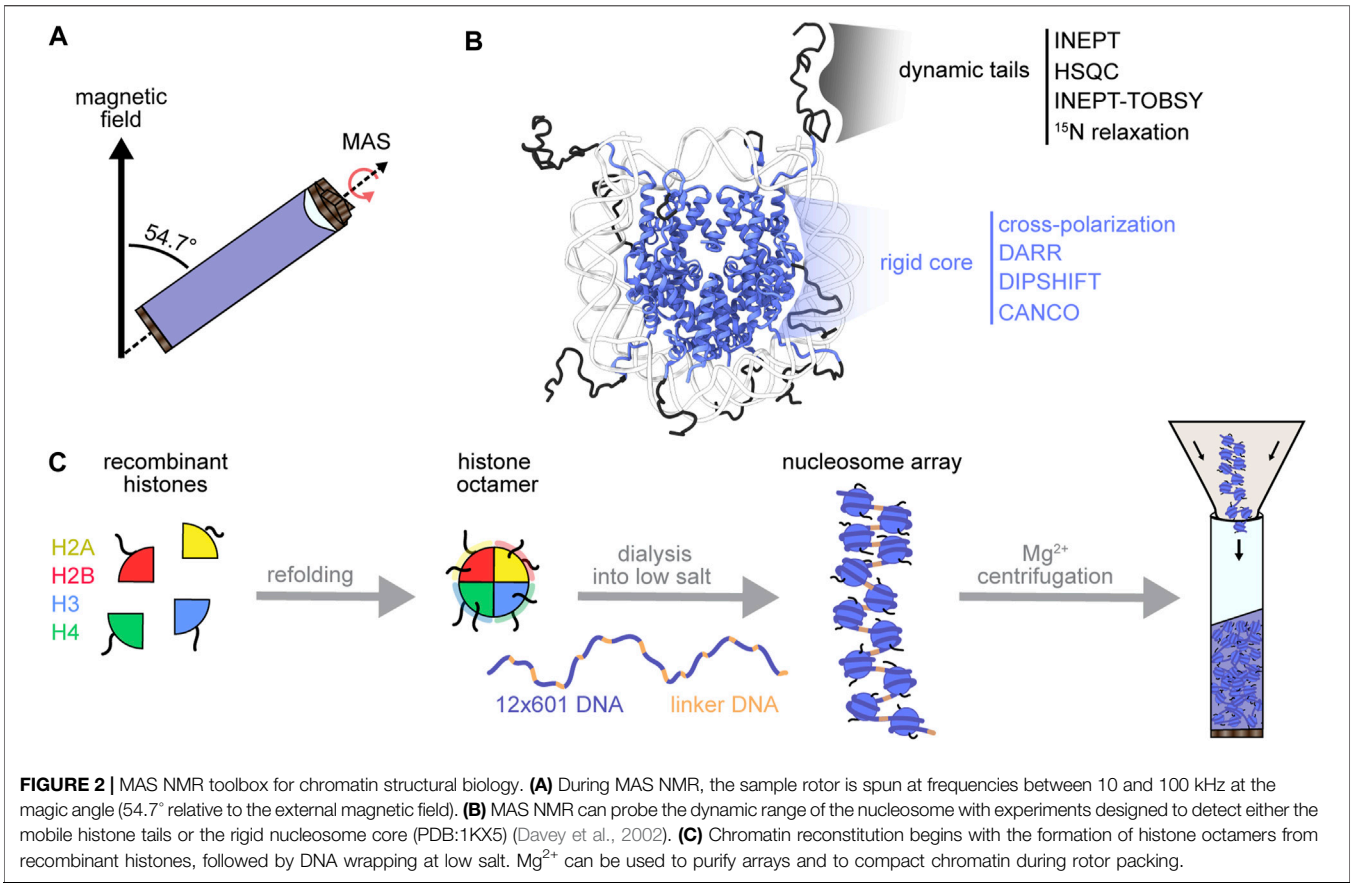


FIGURE 2 | MAS NMR toolbox for chromatin structural biology. **(A)** During MAS NMR, the sample rotor is spun at frequencies between 10 and 100 kHz at the magic angle (54.7° relative to the external magnetic field). **(B)** MAS NMR can probe the dynamic range of the nucleosome with experiments designed to detect either the mobile histone tails or the rigid nucleosome core (PDB:1KX5) (Davey et al., 2002). **(C)** Chromatin reconstitution begins with the formation of histone octamers from recombinant histones, followed by DNA wrapping at low salt. Mg^{2+} can be used to purify arrays and to compact chromatin during rotor packing.

external magnetic field (**Figure 2A**). MAS thereby enables detection of large macromolecules in various material states (amorphous, crystalline, gel-like, liquid) (Quinn and Polenova, 2017; van der Wel, 2018). MAS NMR can detect both slow and fast biomolecular motions by using solution-state (INEPT) experiments that depend on through-bond scalar-couplings (Morris and Freeman, 1979; Andronesi et al., 2005) or by using solid-state cross-polarization (CP) experiments (Hartmann and Hahn, 1962; Pines et al., 1973) that rely on through-space dipolar interactions (**Figure 2B**). These two experiment types allow for qualitative dynamics-based spectral filtering (Matlahov and van der Wel, 2018), where INEPT selects for nuclei that undergo dynamic isotropic motion, and CP builds up signal for nuclei in rigid networks that maintain strong dipolar couplings. MAS NMR can also resolve distance parameters by recoupling spatially informative anisotropic interactions via pulse sequences matched to the sample spinning rate. These features of MAS NMR suit chromatin, a heterogenous polymer that forms an amorphous solid *in vitro* and *in nucleo* (Strickfaden et al., 2020), and that bears nucleosomes with both rigid histone cores and dynamic histone tails (Fierz and Poirier, 2019). Therefore, MAS NMR can detect the dynamic range of nucleosomes while embedded in the chromatin context. In this review, we will cover the practicalities of sample preparation for MAS NMR, highlight the current applications of MAS NMR to chromatin, and finally outline the next frontier of biologically compelling chromatin design and structural analysis. Readers interested in chromatin applications of other structural techniques such as solution NMR spectroscopy, X-ray crystallography and cryo-EM are referred to several recent comprehensive reviews on these topics (van Emmerik and van Ingen, 2019; Zhou et al., 2019; McGinty and Tan, 2021).

CHROMATIN SAMPLE PREPARATION FOR MAS NMR

The chromatin fiber, while richly decorated in cells with PTMs, histone variants and chromatin modulators, can be minimally reconstructed using purified DNA and recombinant histones (**Figure 2C**). Mononucleosomes and nucleosome arrays were first reconstituted by depositing histone octamers onto alpha-satellite DNA and tandem-repeat 5S rDNA, respectively (Simpson et al., 1985; Luger et al., 1997; Fierz and Muir, 2012). The discovery of the high-affinity 601 positioning sequence soon revolutionized chromatin construction by enabling the efficient assembly of highly homogenous chromatin preparations (Lowary and Widom, 1998). The 601 DNA is now favored in most *in vitro* biochemical and structural chromatin studies. Nucleosome arrays have been standardized to include twelve 601 sequence repeats, partitioned by selected lengths of linker DNA (Huynh et al., 2005). The 12-mer

DNA, along with a short helper DNA strand that assists later steps, are cloned for bacterial production on a large scale (tens of milligrams). Histone octamers, on the other hand, are formed by refolding a stoichiometric mixture of the four histone proteins, typically produced recombinantly in *E. coli* and thus devoid of PTMs. For NMR studies, the histone of interest is expressed in isotopically enriched media which leads to a selectively labeled histone type within the otherwise NMR-silent nucleosome complex. Finally, the repetitive DNA and helper DNA strands are excised from their parent plasmids and loaded with octamers by the salt dialysis assembly method to produce well-defined and spaced nucleosome arrays (Dorigo et al., 2003; Dyer et al., 2003; Fyodorov and Kadonaga, 2003) (**Figure 2C**). The helper DNA, which has a weaker affinity for histones, is essential for soaking up the excess histone octamers required to saturate the twelve nucleosome sites in the array. Nucleosome arrays can then be purified from helper DNA by a facile Mg^{2+} precipitation step, in contrast to mononucleosomes which require sucrose gradient or preparative gel purifications to remove leftover DNA. Homotypic and heterotypic post-translationally modified nucleosome arrays can be produced using native chemical ligation, expressed protein ligation, unnatural amino acid incorporation and other chemical approaches (Müller and Muir, 2015; Muller et al., 2016). While these methods typically yield small amounts of modified nucleosome arrays, some can be adapted to the high sample demands of NMR spectroscopy. We will briefly discuss those in the last section of this review.

The development of efficient and reliable protocols for chromatin assembly has led to fundamental insights into chromatin structure. For example, it is now well known that the H4 tail and the H2A/H2B acidic patch are critical for inter-nucleosome stacking (Dorigo et al., 2003; Zhou et al., 2007; Lu et al., 2008) while lysine acetylation on the H4 tail can open up the chromatin fiber and expose DNA (Shogren-Knaak, 2006; Mishra et al., 2016). Chromatin reconstitution was also vital for uncovering the structural continuum of extended 10-nm fibers, folded 30-nm fibers, interdigitated fibers, and irregular globules that depend on Mg^{2+} concentration, DNA linker length, linker histone H1, and array concentration (Baldi et al., 2018; Maeshima et al., 2019; Adhireksan et al., 2020; Brouwer et al., 2021). Despite the detection of structured high-order fibers *in vitro*, recent studies emphasize that interphase chromatin appears to compact irregularly at densities up to 100 mg/ml (Poirier et al., 2009; Hihara et al., 2012; Imai et al., 2017; Cai et al., 2018; Hancock, 2018; Kilic et al., 2018; Audugé et al., 2019). Careful preparation and explicit characterization of chromatin states stands as a crucial step to identify the specialized structures that relate to functions such as gene repression, mitosis, and transcription (Luger et al., 2012).

MAS NMR is well suited to tackle macroscopically heterogenous samples that emulate the irregularity of nuclear chromatin. Non-crystalline samples are routinely packaged for analysis by ultracentrifugation (Bertini et al., 2013; Mandal et al., 2017), where high gravity values generate maximally concentrated sediments of macromolecules while retaining sample hydration and stability for years (Fragai et al., 2013; Wiegand et al., 2020). Sedimentation can be performed in several ways; the sample can be centrifuged inside a tube and then transferred to a rotor, centrifuged directly into a rotor, or sedimented inside the rotor during MAS (Bertini et al., 2013; Mandal et al., 2017). The rate of sedimentation depends on the degree of chromatin compaction (Dorigo et al., 2003) and the process can be sped up by the addition of Mg^{2+} ions. Divalent cations shield the repulsive electrostatic interactions between nucleosomes to compact and eventually aggregate nucleosome arrays. Despite this benefit for MAS NMR studies, Mg^{2+} presents a conundrum for sample preparation due to its profound effect on chromatin structure (Berezhnoy et al., 2016) and inconclusive results regarding its physiological concentrations in the nucleus (Maeshima et al., 2018). As a result, it is important to study chromatin in various buffer conditions and rigorously compare the outcomes. Arrays and mononucleosomes have been prepared for MAS NMR with and without Mg^{2+}-induced precipitation (Gao et al., 2013; le Paige et al., 2021). Mononucleosome sediments with low levels of Mg^{2+} lacked the long-range order expected for stacked fibers formed by high Mg^{2+} concentrations (le Paige et al., 2021). Such irregular packing may illuminate the transient nucleosome-nucleosome interactions that dominate when nucleosomes are not restricted into ordered arrays (Bilokapic et al., 2018; Sanulli et al., 2019; Alvarado et al., 2021). Thus, samples prepared by sedimentation and low Mg^{2+} concentrations may prove crucial for resolving the transient interactions that lead to chromatin compaction and regulation (Gibson et al., 2019; Khanna et al., 2019; Sanulli et al., 2019; Kantidze and Razin, 2020).

MAS NMR OF HISTONE TAILS

The flexible N- and C-terminal histone tails extend from the nucleosome, each with a distinct interaction profile and PTM landscape. The histone tails have been the subject of numerous biochemical and biophysical studies that have revealed their important role in chromatin structure, function, and regulation (Peng et al., 2021). While these dynamic segments are rarely visible on their own in X-ray and cryo-EM structural models, they can be studied by both solution and MAS NMR spectroscopy. The first MAS NMR study by Gao et al. used INEPT-based experiments to map the amino acid specific dynamics of the H3 and H4 tails within arrays at different Mg^{2+} concentrations (Gao et al., 2013). The histone tails remained dynamic regardless of the Mg^{2+} concentration used (0–5 mM), arguing for a much more flexible landscape than suggested previously by X-ray crystallography and hydrogen-deuterium exchange experiments (Luger et al., 1997; Kato et al., 2009). 2D ^{1}H-^{13}C and ^{1}H-^{15}N correlations revealed identifiable amino acids signatures that made it possible to compare the information content of INEPT-based experiments under solution and MAS NMR conditions. For example, Val 35, a unique amino acid in the H3 tail, was visible in both solution and MAS NMR spectra, while residues up to Val 21 were detected in the MAS NMR experiments of H4, in contrast to solution NMR experiments where only residues up to Ala 15 were present (Zhou et al., 2012; Morrison et al., 2021; Rabdano et al.,

2021). Subsequent MAS NMR studies of mononucleosomes by Shi et al. also confirmed these observations (Shi et al., 2018). This discrepancy is most likely due to the type of experiments used (i.e., sidechain vs. backbone correlations), with ^{13}C MAS NMR experiments holding a slight advantage regarding the detection of slower moving tail sidechains such as those that are close to the DNA interface (Shi et al., 2018; Shi et al., 2020b). While the detected tail boundary is consistent between arrays and mononucleosomes in MAS NMR experiments, a closer look into the ^{15}N spin relaxation rates of H3 tails reveals mobility differences (Zandian et al., 2021). Here, T_1 and $T_{1\rho}$ relaxation measurements were used to quantify residue-specific rotational correlation times. The H3 tail was most dynamic in 147 bp mononucleosomes, the mobility was diminished by linker DNA extensions and was further reduced in nucleosome arrays. These results are consistent with solution NMR studies that show H3 tail dynamics to be regulated by transient DNA contacts (Stutzer et al., 2016).

Histone tail interactions with DNA can tuned by the chemical properties of specific PTMs. H4 lysine 20 trimethylation (H4 K20me3), for example, is a hydrophobic modification thought to increase the compaction of nucleosome arrays by altering the adjacent residue side-chain interactions with DNA (Lu et al., 2008). Shoaib et al. used MAS NMR to show that H4 K20 mono- and tri-methylation differentially dictate tail conformation and lead to either open or closed chromatin states, respectively (Shoaib et al., 2021). These conclusions were based on genomic accessibility studies, Mg^{2+} precipitation experiments and 2D ^{1}H-^{13}C INEPT correlations that focused on Val 21, a residue that is, adjacent to the modification site. Surprisingly, the Val 21 Hα-Cα correlation was split into two peaks for the non- and tri-methylated samples but only one peak was visible for the monomethylated H4 tail. While it is not yet clear how these changes in H4 tail dynamics and structure bring about open or closed chromatin states, this study exemplifies the power of MAS NMR to monitor tail dynamics in the context of nucleosome arrays where the inter-nucleosome contacts are essential to the structural hypothesis.

MAS NMR has also revealed that histone tails can sense their local environment and drive chromatin compaction while remaining dynamic and potentially accessible to regulatory proteins. A recent study by le Paige et al. confirmed that the dynamic tails sustain accessibility within dense sedimented samples (le Paige et al., 2021). In this case, the interactions of the histone reader domain PHD2 with histone H3 were compared by INEPT-based experiments in nucleosome sediments and under dilute conditions. For MAS NMR, PHD2 was co-sedimented with nucleosomes during ultracentrifugation, while the dilute samples could be analyzed by solution NMR. The fast MAS rate of 50 kHz afforded ^{1}H-detection and generated highly resolved spectra of the histone tail backbone. Crucially, the spectral quality provided residue-specific assignments to directly compare the isotropic chemical shifts with solution NMR experiments. PHD2 binding produced comparable ^{1}H-^{15}N resonance broadening of residues 3–10 in both conditions. While the weak affinity of PHD2 complicated quantitative comparison of binding interactions, this study confirmed the permeability of dense chromatin environments to histone readers. This strategy will undoubtedly prove useful in future experiments aimed to dissect the interactions between histone readers and their nucleosome binding sites.

MAS NMR OF THE NUCLEOSOME CORE

While the dynamic histone tails are a major target for modification and binding, sequence variations, PTMs and disease-related mutations also embroider the nucleosome core and can disturb fundamental processes like nucleosome sliding, histone exchange, and DNA wrapping (Bowman and Poirier, 2015; Fenley et al., 2018; Bryant et al., 2020; Bagert et al., 2021). In addition, many chromatin readers interact with the nucleosome core which provides additional interaction surfaces through the H2A/H2B acidic patch and the histone/DNA interface. Solution NMR, and in particular methyl-TROSY spectroscopy, has illuminated nucleosome core motions resulting from histone mutations (Kitevski-LeBlanc et al., 2018), DNA methylation (Abramov et al., 2020), and interacting proteins (Kato et al., 2011; Sanulli et al., 2019). However, only a few amino acid types are typically detectable due to the necessary methyl-labeling scheme. A major advantage of MAS NMR is the ability to characterize the entire nucleosome core in the same samples and conditions as those used to analyze the dynamic histone tails. The Nordenskiöld lab has pioneered MAS NMR of the nucleosome core by extensively assigning the highly resolved histone H3 and H4 spectra obtained with CP-based ^{15}N-^{13}C-^{13}C three-dimensional experiments (Shi et al., 2018; Shi et al., 2020a). In the first of several studies, histone H4-labeled mononucleosomes and 12-mer arrays were Mg^{2+}-precipitated and subjected to CANCO and quantitative DIPSHIFT experiments which capture motion on a wide ns-ms timescale (Munowitz et al., 1982; Shi et al., 2018). Motions at the ns-scale involve side-chain fluctuations and μs-ms motions correspond to larger domain movements (Lewandowski, 2013). The array and mononucleosome samples were consistent in overall structure and dynamics. Small signal intensity differences were observed for residues adjacent to the N-terminal H4 tail, suggesting that the base of the tail is more rigid in the chromatin fiber. The mononucleosome samples in this study were precipitated with 20 mM Mg^{2+} which generates columnar stacked assemblies that may dictate histone dynamics and influence interpretation (Berezhnoy et al., 2016). Nonetheless, the precedent of quantitative dynamic parameters for each histone residue created a platform for probing other histones and DNA sequences.

Histone H3 spectra soon followed, illustrating that both H3 and H4 experience some ns-μs motions and have regional clusters of moderately altered μs-ms motions and highly dynamic termini (Shi et al., 2018; Shi et al., 2020b). Together, these results suggest histone H3 and H4 form stable folds but can undergo local intermediate motions. When these motions are mapped on the nucleosome structure, small neighboring clusters of dynamic sites connect from the nucleosome core to the DNA-bound periphery. Correlative motions of residues that contact DNA may be

important for regulating biological processes like DNA wrapping, sliding, and nucleosome assembly (Bowman and Poirier, 2015; Sinha et al., 2017; Sanulli et al., 2019; Bagert et al., 2021). The connection between histone and DNA dynamics was supported by comparing nucleosomes formed with the 601 sequence to those prepared with a tandem-repeat (TTAGGG) telomere sequence. The TTAGGG nucleosomes displayed a greater range of motions in the cluster network compared to the 601 nucleosomes consistent with previous experiments which showed that telomeric nucleosomes are less stable and wrap DNA less tightly (Shi et al., 2020b; Soman et al., 2020). Reduced nucleosome stability may translate into more flexible chromatin fibers that in turn enhance the potential for phase separation at telomeres (Sanulli et al., 2019; Farr et al., 2021). Here, MAS NMR greatly contributed to establishing a connection between histone core dynamics and the compaction of chromatin fibers. Further broadening the scope of DNA to include other genetic sequences and DNA modifications will allow MAS NMR to bridge the vast expanse of genomic data with nucleosome-specific dynamics.

MAS NMR has also been used to detect interactions between the nucleosome core and regulatory proteins. The Baldus and van Ingen labs have demonstrated the sensitivity of MAS NMR to chemical environment changes when the H2A/H2B acidic patch is bound by a peptide segment of the viral LANA protein (Xiang et al., 2018). In this study, LANA and mononucleosomes were co-sedimented with 2 mM Mg^{2+}. ^{1}H-detected CP experiments were used to assign 93% of the H2A core backbone, locate the LANA binding site, and independently model the binding conformation in agreement to the crystal structure. ^{1}H-detection was crucial to observe significant chemical shift perturbations (CSPs) that were diminished in the ^{13}C and ^{15}N dimensions. Importantly, the absence of peak-splitting suggests that the LANA peptide bound both sides of the nucleosome simultaneously. Such an observation is important because the LANA peptide affects chromatin compaction by shielding the acidic patch (Chodaparambil et al., 2007). Detection of acidic patch interactions by ^{1}H-detected MAS NMR holds promise for studying the myriad of chromatin modulators that use basic disordered segments to bind the nucleosome (Mashtalir et al., 2020; Teles et al., 2020).

MAS NMR OF CHROMATIN MODULATORS

Almost all MAS NMR studies of chromatin so far have focused on the nucleosome perspective, while the structure, dynamics, and interactions of chromatin modulators have largely remained unexplored. Here, MAS NMR can offer unique advantages as often chromatin modulators are too large to study by solution NMR or they form dynamic, viscous and heterogenous phase separated states that cannot be dissected by single particle techniques such as cryo-EM. Most importantly, however, such studies can be performed in the presence of mononucleosomes or nucleosome arrays where isotopic labeling of each protein one at a time can provide an opportunity to analyze a single component at high resolution in the complex chromatin-modulator assembly. A study from our lab used MAS NMR to illuminate the liquid-liquid phase separation of phosphorylated heterochromatin protein 1α (pHP1α) in the presence and absence of nucleosome arrays (Ackermann and Debelouchina, 2019). HP1α is a key chromatin modulator responsible for the formation of heterochromatin domains in the nucleus where silenced genes are sequestered (Cheutin, 2003). N-terminal serine phosphorylation was previously found to enhance the phase separation behavior of HP1α and to promote the transition from dynamic liquid to an arrested gel state (Larson et al., 2017). To measure the phenomenon in a more physiological chromatin environment, phase separated and isotopically labeled pHP1α was packaged into MAS rotors with and without H3 lysine 9 trimethylated nucleosome arrays, the native binding partner for pHP1α. Both INEPT and CP experiments were employed to measure the sample dynamics during gelation. The addition of chromatin appeared to disrupt the pHP1α-pHP1α gel network as detected by the lower overall CP signal for pHP1α. In addition, 2D INEPT-TOBSY and CP-based DARR ^{13}C-^{13}C correlations revealed changes in the dynamics of specific serine residues as pHP1α transitioned from the liquid to the gel state. In this case, MAS NMR provided an opportunity to observe in real time molecular interactions that underlie the formation of phase separated chromatin environments. As many chromatin modulators are now known to undergo phase separation (Weber, 2019), these unique capabilities of MAS NMR can no doubt be exploited further to provide a much needed molecular view of this complex biological process.

CHEMICAL BIOLOGY TOOLS FOR MAS NMR OF MODIFIED CHROMATIN SAMPLES

So far, MAS NMR studies of chromatin have largely used samples prepared with recombinant histones that are devoid of PTMs. Considering the importance of histone modifications in defining chromatin structure, dynamics and function, it is necessary to consider efficient and specific modification strategies that are compatible with isotopic labeling. In some serendipitous cases, enzymatic modification during expression in *E. coli* or after purification may produce homogeneously modified proteins. For example, we have used co-expression with the kinase CK2 to prepare isotopically labeled and well-defined phosphorylated HP1α samples with high yields (Ackermann and Debelouchina, 2019). More often, however, enzymatic approaches result in incomplete or off target modifications. Therefore, chemical approaches that introduce PTMs in a well-defined and controlled manner are highly desirable (Debelouchina and Muir, 2017).

Chromatin templates have long served as an expansive and challenging canvas for the development of chemical biology methods for protein modification (Allis and Muir, 2011). The histone proteins are relatively small (ranging in size from 100 to 150 amino acids) and practically devoid of cysteine residues (the lone Cys110 on H3.2 can easily be replaced with an alanine or serine residue without loss of structure or function). This makes the application of cysteine-based modification approaches

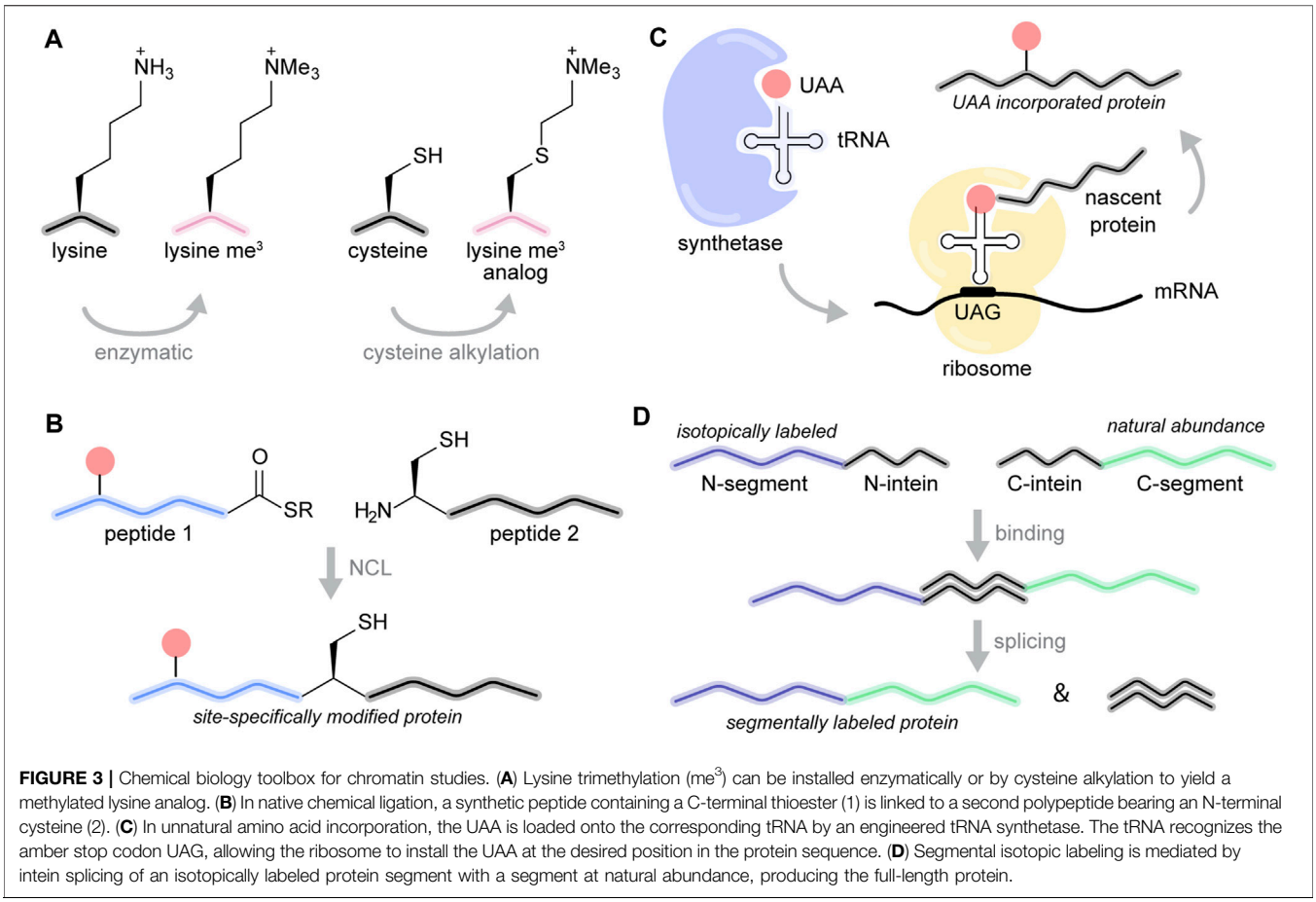


FIGURE 3 | Chemical biology toolbox for chromatin studies. (**A**) Lysine trimethylation (me^3) can be installed enzymatically or by cysteine alkylation to yield a methylated lysine analog. (**B**) In native chemical ligation, a synthetic peptide containing a C-terminal thioester (1) is linked to a second polypeptide bearing an N-terminal cysteine (2). (**C**) In unnatural amino acid incorporation, the UAA is loaded onto the corresponding tRNA by an engineered tRNA synthetase. The tRNA recognizes the amber stop codon UAG, allowing the ribosome to install the UAA at the desired position in the protein sequence. (**D**) Segmental isotopic labeling is mediated by intein splicing of an isotopically labeled protein segment with a segment at natural abundance, producing the full-length protein.

relatively straightforward (Chalker et al., 2009; Boutureira and Bernardes, 2015). Cysteine modification, for example, has been widely used to introduce spectroscopic probes, including fluorescent labels and paramagnetic relaxation enhancement (PRE) tags. More importantly, however, cysteine alkylation provides a convenient and efficient strategy to prepare lysine methylation mimics (Simon et al., 2007). In this case, the targeted lysine residue is replaced with a cysteine, and the cysteine is modified with mono-, di-, or trimethyl containing alkylating reagents (**Figure 3A**). The reaction is usually performed with the purified protein under denaturing conditions and is compatible with isotopic labeling provided that the protein can be refolded. While the resulting mimic contains a sulfur instead of a carbon atom at the γ position of the side chain, *in vitro* biochemical studies have shown that this mimic can faithfully reproduce the functional consequences of lysine methylation (Simon et al., 2007). We have successfully used this strategy to prepare large amounts of nucleosome array samples that contained H3 K9me3 (Ackermann and Debelouchina, 2019), the relevant modification for HP1α binding and heterochromatin formation. This strategy was also used to explore the effects of K20 methylation on the dynamics of the H4 tail (Shoaib et al., 2021).

Multiple PTMs can be introduced with a technique called native chemical ligation (NCL) (Dawson and Kent, 2000). In this case, the N-terminal segment of the protein (typically the first 10 to 50 residues) is made by solid-phase peptide synthesis and post-translationally modified amino acids can be introduced at specific positions in the sequence as desired (**Figure 3B**). The peptide ends with a C-terminal thioester necessary for the subsequent ligation step. At the same time, the remaining C-terminal segment of the protein can be prepared recombinantly in *E.coli* with or without isotopic labeling (Muir et al., 1998). This segment requires an N-terminal cysteine for ligation. Upon mixing, the synthetic peptide and the recombinant piece undergo a set of thio-esterification steps that result in a native peptide bond at the ligation junction. If necessary, the cysteine residue at the junction can be converted to an alanine by desulfurization (Yan and Dawson, 2001). It is also possible to adapt this technique to perform three-piece ligations and to introduce modifications in the middle or the C-terminal segment of the histone protein (Hackeng et al., 1999). The unprecedented chemical control and versatility of NCL has been used to construct large libraries (with more than 100 members) of uniquely modified nucleosomes and to elucidate the mechanisms of chromatin remodeling and multivalent PTM readout (Nguyen et al., 2014; Dann et al., 2017). While NCL has not yet been applied for MAS NMR of chromatin, it has been impactful in numerous solution and MAS NMR studies of other proteins (Kwon et al., 2015; Zoukimian et al., 2019).

Bypassing the need for cysteines and synthesis, single PTMs can be introduced into proteins using unnatural amino acid (UAA) incorporation by genetic means (amber suppression) (Lang and Chin, 2014) (**Figure 3C**). Typically, the amber stop codon (TAG/UAG) is used to signal the position where the unnatural amino acid will be placed. To interpret this message correctly, the ribosome requires an engineered tRNA that can recognize this codon and is loaded with the UAA. The tRNA and the gene for an appropriately engineered matching tRNA synthetase that loads the UAA onto the tRNA are typically encoded onto a separate plasmid. *E. coli* cultures are then transformed with the tRNA/tRNA synthetase plasmid and a plasmid that contains the gene for the protein of interest with the TAG mutation. The unnatural amino acid is added to the media, and expression and purification of the modified protein can be performed as usual. While amber suppression systems have been engineered for the introduction of more than 100 UAAs (Liu and Schultz, 2010), the UAAs of particular interest to the chromatin structural biologist are acetylated lysine, phosphorylated serine and UAA precursors that can be converted into methylated amino acids (Neumann et al., 2008; Pirman et al., 2015; Wang and Liu, 2017). The major drawback of amber suppression for MAS NMR studies of chromatin is that it results in much lower yields of the desired protein, a situation that can be severely exacerbated under isotopic labeling conditions. Nevertheless, we have successfully used this technique to introduce UAAs in different isotopically labeled proteins, including some that are relevant for chromatin studies (Lim et al., 2020).

Finally, it is important to mention intein-mediated segmental labeling (Muir et al., 1998; Shah and Muir, 2014), a useful tool for proteins with congested NMR spectra (Züger and Iwai, 2005; Schubeis et al., 2015; Frederick et al., 2017; Gupta and Tycko, 2018; Wiegand et al., 2018; Ciragan et al., 2020) (**Figure 3D**). Split inteins are a class of proteins found in unicellular organisms that can "stitch" together other protein segments with a native peptide bond. To prepare segmentally labeled samples for NMR spectroscopy, the desired protein is split into two segments. The N-terminal segment is attached to an N-terminal split intein while the C-terminal segment is attached to the matching C-terminal intein piece. The N- and C-segments can be prepared independently in different cultures, for example, one using ^{13}C, ^{15}N labeling and the other at natural abundance. Once the segments are purified, mixing of the segments results in an interaction between the intein pieces which adopt a functional intein fold. The intein mediates the formation of a native peptide bond between the two protein segments while excising itself in the process. This process requires a cysteine residue at the ligation junction. The result is a full-length protein with only a segment of the sequence visible by NMR, thereby simplifying the acquired spectra. This can aid assignment protocols (Iwai and Züger, 2007) and provide the opportunity to probe specific inter- or intra-molecular interactions (Frederick et al., 2017). There is now a large variety of efficient split intein pairs that can be used for this purpose (Pinto et al., 2020), including some that work well under denaturing conditions (Stevens et al., 2016). We envision that this technique will be extremely valuable in the resonance assignment and MAS NMR analysis of chromatin interacting proteins.

DISCUSSION

While still few in number, the recent applications of MAS NMR have demonstrated its versatile capabilities in the structural analysis of chromatin samples. Uniquely capable of characterizing both the rigid and dynamic components of mononucleosome and array preparations, precipitated, sedimented, and concentrated samples, chromatin MAS NMR has illuminated fundamental aspects of nucleosome dynamics, histone tail recognition, acidic patch docking, and heterochromatin phase separation. Most chromatin experiments so far have relied on dipolar and scalar based experiments performed at moderate spinning frequencies (10–20 kHz) and the detection and analysis of ^{13}C and ^{15}N signals. We expect that ^{1}H-detection and fast MAS (at 100 kHz and beyond) will continue to improve signal sensitivity and facilitate the resolution of chemical shift perturbations and chromatin interactions (Andreas et al., 2016; Struppe et al., 2017). Since chromatin assemblies contain multiple proteins and DNA, the preparation of large amounts of isotopically labeled samples can be time consuming and challenging, especially if chromatin interacting proteins or PTMs are included. We therefore expect that such samples will benefit tremendously from sensitivity-enhancement developments such as dynamic nuclear polarization and cryo-MAS probes (Lilly Thankamony et al., 2017; Hassan et al., 2020). Chemical biology tools such as cysteine alkylation, native chemical ligation and amber suppression enable the preparation of specifically decorated chromatin, while inteins allow the simplification of crowded histone or chromatin reader spectra. We envision that the combination of chemical biology tools and MAS NMR will provide the unprecedented opportunity to construct and dissect complex chromatin environments where dynamic multifaceted interactions can be interrogated one at a time.

AUTHOR CONTRIBUTIONS

BA prepared the figure images. BA and GD wrote the article.

FUNDING

This work was supported by NIH R35 GM138382 grant to GD and T32 GM008326 fellowship to BA.

REFERENCES

Abramov, G., Velyvis, A., Rennella, E., Wong, L. E., and Kay, L. E. (2020). A Methyl-TROSY Approach for NMR Studies of High-Molecular-Weight DNA with Application to the Nucleosome Core Particle. *Proc. Natl. Acad. Sci. USA* 117 (23), 12836–12846. doi:10.1073/pnas.2004317117

Ackermann, B. E., and Debelouchina, G. T. (2019). Heterochromatin Protein HP1α Gelation Dynamics Revealed by Solid-State NMR Spectroscopy. *Angew. Chem. Int. Ed.* 58 (19), 6300–6305. doi:10.1002/anie.201901141

Adhireksan, Z., Sharma, D., Lee, P. L., and Davey, C. A. (2020). Near-atomic Resolution Structures of Interdigitated Nucleosome Fibres. *Nat. Commun.* 11 (1), 4747. doi:10.1038/s41467-020-18533-2

Allis, C. D., and Muir, T. W. (2011). Spreading Chromatin into Chemical Biology. *ChemBioChem* 12 (2), 264–279. doi:10.1002/cbic.201000761

Alvarado, W., Moller, J., Ferguson, A. L., and de Pablo, J. J. (2021). Tetranucleosome Interactions Drive Chromatin Folding. *ACS Cent. Sci.* 7, 1019–1027. doi:10.1021/acscentsci.1c00085

Andreas, L. B., Jaudzems, K., Stanek, J., Lalli, D., Bertarello, A., Le Marchand, T., et al. (2016). Structure of Fully Protonated Proteins by Proton-Detected Magic-Angle Spinning NMR. *Proc. Natl. Acad. Sci. USA* 113 (33), 9187–9192. doi:10.1073/pnas.1602248113

Andronesi, O. C., Becker, S., Seidel, K., Heise, H., Young, H. S., and Baldus, M. (2005). Determination of Membrane Protein Structure and Dynamics by Magic-Angle-Spinning Solid-State NMR Spectroscopy†. *J. Am. Chem. Soc.* 127 (37), 12965–12974. doi:10.1021/ja0530164

Audugé, N., Padilla-Parra, S., Tramier, M., Borghi, N., and Coppey-Moisan, M. (2019). Chromatin Condensation Fluctuations rather Than Steady-State Predict Chromatin Accessibility. *Nucleic Acids Res.* 47 (12), 6184–6194. doi:10.1093/nar/gkz373

Bagert, J. D., Mitchener, M. M., Patriotis, A. L., Dul, B. E., Wojcik, F., Nacev, B. A., et al. (2021). Oncohistone Mutations Enhance Chromatin Remodeling and Alter Cell Fates. *Nat. Chem. Biol.* 17, 403–411. doi:10.1038/s41589-021-00738-1

Baldi, S., Krebs, S., Blum, H., and Becker, P. B. (2018). Genome-wide Measurement of Local Nucleosome Array Regularity and Spacing by Nanopore Sequencing. *Nat. Struct. Mol. Biol.* 25, 894–901. doi:10.1038/s41594-018-0110-0

Berezhnoy, N. V., Liu, Y., Allahverdi, A., Yang, R., Su, C.-J., Liu, C.-F., et al. (2016). The Influence of Ionic Environment and Histone Tails on Columnar Order of Nucleosome Core Particles. *Biophys. J.* 110 (8), 1720–1731. doi:10.1016/j.bpj.2016.03.016

Bertini, I., Luchinat, C., Parigi, G., and Ravera, E. (2013). SedNMR: On the Edge between Solution and Solid-State NMR. *Acc. Chem. Res.* 46 (9), 2059–2069. doi:10.1021/ar300342f

Bilokapic, S., Strauss, M., and Halic, M. (2018). Cryo-EM of Nucleosome Core Particle Interactions in Trans. *Sci. Rep.* 8 (1), 7046. doi:10.1038/s41598-018-25429-1

Boutureira, O., and Bernardes, G. J. L. (2015). Advances in Chemical Protein Modification. *Chem. Rev.* 115 (5), 2174–2195. doi:10.1021/cr500399p

Bowman, G. D., and Poirier, M. G. (2015). Post-Translational Modifications of Histones that Influence Nucleosome Dynamics. *Chem. Rev.* 115 (6), 2274–2295. doi:10.1021/cr500350x

Brouwer, T., Pham, C., Kaczmarczyk, A., de Voogd, W.-J., Botto, M., Vizjak, P., et al. (2021). A Critical Role for Linker DNA in Higher-Order Folding of Chromatin Fibers. *Nucleic Acids Res.* 49 (5), 2537–2551. doi:10.1093/nar/gkab058

Bryant, L., Li, D., Cox, S. G., Marchione, D., Joiner, E. F., Wilson, K., et al. (2020). Histone H3.3 beyond Cancer: Germline Mutations in Histone 3 Family 3A and 3B Cause a Previously Unidentified Neurodegenerative Disorder in 46 Patients. *Sci. Adv.* 6 (49), eabc9207. doi:10.1126/sciadv.abc9207

Cai, S., Böck, D., Pilhofer, M., and Gan, L. (2018). The *In Situ* Structures of Mono-, Di-, and Tri-nucleosomes in Human Heterochromatin. *Mol. Biol. Cel.* 29, 2450–2457. doi:10.1091/mbc.E18-05-0331

Chalker, J. M., Bernardes, G. J. L., Lin, Y. A., and Davis, B. G. (2009). Chemical Modification of Proteins at Cysteine: Opportunities in Chemistry and Biology. *Chem. Asian J.* 4 (5), 630–640. doi:10.1002/asia.200800427

Chen, Q., Yang, R., Korolev, N., Liu, C. F., and Nordenskiöld, L. (2017). Regulation of Nucleosome Stacking and Chromatin Compaction by the Histone H4 N-Terminal Tail-H2a Acidic Patch Interaction. *J. Mol. Biol.* 429 (13), 2075–2092. doi:10.1016/j.jmb.2017.03.016

Cheutin, T. (2003). Maintenance of Stable Heterochromatin Domains by Dynamic HP1 Binding. *Science* 299 (5607), 721–725. doi:10.1126/science.1078572

Chodaparambil, J. V., Barbera, A. J., Lu, X., Kaye, K. M., Hansen, J. C., and Luger, K. (2007). A Charged and Contoured Surface on the Nucleosome Regulates Chromatin Compaction. *Nat. Struct. Mol. Biol.* 14 (11), 1105–1107. doi:10.1038/nsmb1334

Ciragan, A., Backlund, S. M., Mikula, K. M., Beyer, H. M., Samuli Ollila, O. H., and Iwaï, H. (2020). NMR Structure and Dynamics of TonB Investigated by Scar-Less Segmental Isotopic Labeling Using a Salt-Inducible Split Intein. *Front. Chem.* 8, 136. doi:10.3389/fchem.2020.00136

Dann, G. P., Liszczak, G. P., Bagert, J. D., Müller, M. M., Nguyen, U. T. T., Wojcik, F., et al. (2017). ISWI Chromatin Remodellers Sense Nucleosome Modifications to Determine Substrate Preference. *Nature* 548 (7669), 607–611. doi:10.1038/nature23671

Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W., and Richmond, T. J. (2002). Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9Å Resolution. *J. Mol. Biol.* 319 (5), 1097–1113. doi:10.1016/S0022-2836(02)00386-8

Dawson, P. E., and Kent, S. B. H. (2000). Synthesis of Native Proteins by Chemical Ligation. *Annu. Rev. Biochem.* 69 (1), 923–960. doi:10.1146/annurev.biochem.69.1.923

Debelouchina, G. T., and Muir, T. W. (2017). A Molecular Engineering Toolbox for the Structural Biologist. *Quart. Rev. Biophys.* 50, e7. doi:10.1017/S0033583517000051

Dorigo, B., Schalch, T., Bystricky, K., and Richmond, T. J. (2003). Chromatin Fiber Folding: Requirement for the Histone H4 N-Terminal Tail. *J. Mol. Biol.* 327 (1), 85–96. doi:10.1016/S0022-2836(03)00025-1

Dyer, P. N., Edayathumangalam, R. S., White, C. L., Bao, Y., Chakravarthy, S., Muthurajan, U. M., et al. (2003). "Reconstitution of Nucleosome Core Particles from Recombinant Histones and DNA," in *Methods in Enzymology* (New York, NY: Elsevier), 23–44. doi:10.1016/s0076-6879(03)75002-2

Farr, S. E., Woods, E. J., Joseph, J. A., Garaizar, A., and Collepardo-Guevara, R. (2021). Nucleosome Plasticity Is a Critical Element of Chromatin Liquid-Liquid Phase Separation and Multivalent Nucleosome Interactions. *Nat. Commun.* 12 (1), 2883. doi:10.1038/s41467-021-23090-3

Fenley, A. T., Anandakrishnan, R., Kidane, Y. H., and Onufriev, A. V. (2018). Modulation of Nucleosomal DNA Accessibility via Charge-Altering post-translational Modifications in Histone Core. *Epigenetics & Chromatin* 11 (1), 11. doi:10.1186/s13072-018-0181-5

Fierz, B., and Muir, T. W. (2012). Chromatin as an Expansive Canvas for Chemical Biology. *Nat. Chem. Biol.* 8 (5), 417–427. doi:10.1038/nchembio.938

Fierz, B., and Poirier, M. G. (2019). Biophysics of Chromatin Dynamics. *Annu. Rev. Biophys.* 48 (1), 321–345. doi:10.1146/annurev-biophys-070317-032847

Fragai, M., Luchinat, C., Parigi, G., and Ravera, E. (2013). Practical Considerations over Spectral Quality in Solid State NMR Spectroscopy of Soluble Proteins. *J. Biomol. NMR* 57 (2), 155–166. doi:10.1007/s10858-013-9776-0

Frederick, K. K., Michaelis, V. K., Caporini, M. A., Andreas, L. B., Debelouchina, G. T., Griffin, R. G., et al. (2017). Combining DNP NMR with Segmental and Specific Labeling to Study a Yeast Prion Protein Strain that Is Not Parallel In-Register. *Proc. Natl. Acad. Sci. USA* 114 (14), 3642–3647. doi:10.1073/pnas.1619051114

Fyodorov, D. V., and Kadonaga, J. T. (2003). "Chromatin Assembly *In Vitro* with Purified Recombinant ACF and NAP-1," in *Methods in Enzymology* (New York, NY: Elsevier), 499–515. doi:10.1016/s0076-6879(03)71037-4

Gao, M., Nadaud, P. S., Bernier, M. W., North, J. A., Hammel, P. C., Poirier, M. G., et al. (2013). Histone H3 and H4 N-Terminal Tails in Nucleosome Arrays at Cellular Concentrations Probed by Magic Angle Spinning NMR Spectroscopy. *J. Am. Chem. Soc.* 135 (41), 15278–15281. doi:10.1021/ja407526s

Gibson, B. A., Doolittle, L. K., Schneider, M. W. G., Jensen, L. E., Gamarra, N., Henry, L., et al. (2019). Organization of Chromatin by Intrinsic and Regulated Phase Separation. *Cell* 179 (2), 470–484.e421. doi:10.1016/j.cell.2019.08.037

Gupta, S., and Tycko, R. (2018). Segmental Isotopic Labeling of HIV-1 Capsid Protein Assemblies for Solid State NMR. *J. Biomol. NMR* 70 (2), 103–114. doi:10.1007/s10858-017-0162-1

Hackeng, T. M., Griffin, J. H., and Dawson, P. E. (1999). Protein Synthesis by Native Chemical Ligation: Expanded Scope by Using Straightforward

Methodology. *Proc. Natl. Acad. Sci.* 96 (18), 10068–10073. doi:10.1073/pnas.96.18.10068

Hancock, R. (2018). Crowding, Entropic Forces, and Confinement: Crucial Factors for Structures and Functions in the Cell Nucleus. *Biochem. Mosc.* 83 (4), 326–337. doi:10.1134/S0006297918040041

Hartmann, S. R., and Hahn, E. L. (1962). Nuclear Double Resonance in the Rotating Frame. *Phys. Rev.* 128(5), 2042–2053. doi:10.1103/PhysRev.128.2042

Hassan, A., Quinn, C. M., Struppe, J., Sergeyev, I. V., Zhang, C., Guo, C., et al. (2020). Sensitivity Boosts by the CPMAS CryoProbe for Challenging Biological Assemblies. *J. Magn. Reson.* 311, 106680. doi:10.1016/j.jmr.2019.106680

He, S., Wu, Z., Tian, Y., Yu, Z., Yu, J., Wang, X., et al. (2020). Structure of Nucleosome-Bound Human BAF Complex. *Science* 367 (6480), 875–881. doi:10.1126/science.aaz9761

Hihara, S., Pack, C.-G., Kaizu, K., Tani, T., Hanafusa, T., Nozaki, T., et al. (2012). Local Nucleosome Dynamics Facilitate Chromatin Accessibility in Living Mammalian Cells. *Cel Rep.* 2 (6), 1645–1656. doi:10.1016/j.celrep.2012.11.008

Hsieh, T.-H. S., Weiner, A., Lajoie, B., Dekker, J., Friedman, N., and Rando, O. J. (2015). Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. *Cell* 162 (1), 108–119. doi:10.1016/j.cell.2015.05.048

Huynh, V. A. T., Robinson, P. J. J., and Rhodes, D. (2005). A Method for the *In Vitro* Reconstitution of a Defined "30nm" Chromatin Fibre Containing Stoichiometric Amounts of the Linker Histone. *J. Mol. Biol.* 345 (5), 957–968. doi:10.1016/j.jmb.2004.10.075

Imai, R., Nozaki, T., Tani, T., Kaizu, K., Hibino, K., Ide, S., et al. (2017). Density Imaging of Heterochromatin in Live Cells Using Orientation-Independent-DIC Microscopy. *MBoC* 28 (23), 3349–3359. doi:10.1091/mbc.e17-06-0359

Iwai, H., and Züger, S. (2007). Protein Ligation: Applications in NMR Studies of Proteins. *Biotechnol. Genet. Eng. Rev.* 24 (1), 129–146. doi:10.1080/02648725.2007.10648096

Jang, S., and Song, J.-J. (2019). The Big Picture of Chromatin Biology by Cryo-EM. *Curr. Opin. Struct. Biol.* 58, 76–87. doi:10.1016/j.sbi.2019.05.017

Kalashnikova, A. A., Porter-Goff, M. E., Muthurajan, U. M., Luger, K., and Hansen, J. C. (2013). The Role of the Nucleosome Acidic Patch in Modulating Higher Order Chromatin Structure. *J. R. Soc. Interf.* 10 (82), 20121022. doi:10.1098/rsif.2012.1022

Kantidze, O. L., and Razin, S. V. (2020). Weak Interactions in Higher-Order Chromatin Organization. *Nucleic Acids Res.* 48 (9), 4614–4626. doi:10.1093/nar/gkaa261

Kato, H., Gruschus, J., Ghirlando, R., Tjandra, N., and Bai, Y. (2009). Characterization of the N-Terminal Tail Domain of Histone H3 in Condensed Nucleosome Arrays by Hydrogen Exchange and NMR. *J. Am. Chem. Soc.* 131 (42), 15104–15105. doi:10.1021/ja9070078

Kato, H., van Ingen, H., Zhou, B.-R., Feng, H., Bustin, M., Kay, L. E., et al. (2011). Architecture of the High Mobility Group Nucleosomal Protein 2-nucleosome Complex as Revealed by Methyl-Based NMR. *Proc. Natl. Acad. Sci.* 108 (30), 12283–12288. doi:10.1073/pnas.1105848108

Khanna, N., Zhang, Y., Lucas, J. S., Dudko, O. K., and Murre, C. (2019). Chromosome Dynamics Near the Sol-Gel Phase Transition Dictate the Timing of Remote Genomic Interactions. *Nat. Commun.* 10 (1), 2771. doi:10.1038/s41467-019-10628-9

Kilic, S., Felekyan, S., Doroshenko, O., Boichenko, I., Dimura, M., Vardanyan, H., et al. (2018). Single-molecule FRET Reveals Multiscale Chromatin Dynamics Modulated by HP1α. *Nat. Commun.* 9 (1), 235. doi:10.1038/s41467-017-02619-5

Kim, K., and Guck, J. (2020). The Relative Densities of Cytoplasm and Nuclear Compartments Are Robust against Strong Perturbation. *Biophysical J.* 119 (10), 1946–1957. doi:10.1016/j.bpj.2020.08.044

Kitevski-LeBlanc, J. L., Yuwen, T., Dyer, P. N., Rudolph, J., Luger, K., and Kay, L. E. (2018). Investigating the Dynamics of Destabilized Nucleosomes Using Methyl-TROSY NMR. *J. Am. Chem. Soc.* 140 (14), 4774–4777. doi:10.1021/jacs.8b00931

Kouzarides, T. (2007). Chromatin Modifications and Their Function. *Cell* 128 (4), 693–705. doi:10.1016/j.cell.2007.02.005

Krietenstein, N., Abraham, S., Venev, S. V., Abdennur, N., Gibcus, J., Hsieh, T.-H. S., et al. (2020). Ultrastructural Details of Mammalian Chromosome Architecture. *Mol. Cel* 78 (3), 554–565.e7. doi:10.1016/j.molcel.2020.03.003

Kwon, B., Tietze, D., White, P. B., Liao, S. Y., and Hong, M. (2015). Chemical Ligation of the Influenza M2 Protein for Solid-State NMR Characterization of the Cytoplasmic Domain. *Protein Sci.* 24 (7), 1087–1099. doi:10.1002/pro.2690

Lang, K., and Chin, J. W. (2014). Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins. *Chem. Rev.* 114 (9), 4764–4806. doi:10.1021/cr400355w

Larson, A. G., Elnatan, D., Keenen, M. M., Trnka, M. J., Johnston, J. B., Burlingame, A. L., et al. (2017). Liquid Droplet Formation by HP1α Suggests a Role for Phase Separation in Heterochromatin. *Nature* 547 (7662), 236–240. doi:10.1038/nature22822

le Paige, U. B., Xiang, S., Hendrix, M. M. R. M., Zhang, Y., Folkers, G. E., Weingarth, M., et al. (2021). Characterization of Nucleosome Sediments for Protein Interaction Studies by Solid-State NMR Spectroscopy. *Magn. Reson.* 2 (1), 187–202. doi:10.5194/mr-2-187-2021

Lewandowski, J. R. (2013). Advances in Solid-State Relaxation Methodology for Probing Site-specific Protein Dynamics. *Acc. Chem. Res.* 46 (9), 2018–2027. doi:10.1021/ar300334g

Lilly Thankamony, A. S., Wittmann, J. J., Kaushik, M., and Corzilius, B. (2017). Dynamic Nuclear Polarization for Sensitivity Enhancement in Modern Solid-State NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* 102-103, 120–195. doi:10.1016/j.pnmrs.2017.06.002

Lim, B. J., Ackermann, B. E., and Debelouchina, G. T. (2020). Targetable Tetrazine-Based Dynamic Nuclear Polarization Agents for Biological Systems. *ChemBioChem* 21 (9), 1315–1319. doi:10.1002/cbic.201900609

Liu, C. C., and Schultz, P. G. (2010). Adding New Chemistries to the Genetic Code. *Annu. Rev. Biochem.* 79 (1), 413–444. doi:10.1146/annurev.biochem.052308.105824

Lowary, P. T., and Widom, J. (1998). New DNA Sequence Rules for High Affinity Binding to Histone Octamer and Sequence-Directed Nucleosome Positioning. *J. Mol. Biol.* 276 (1), 19–42. doi:10.1006/jmbi.1997.1494

Lu, X., Simon, M. D., Chodaparambil, J. V., Hansen, J. C., Shokat, K. M., and Luger, K. (2008). The Effect of H3K79 Dimethylation and H4K20 Trimethylation on Nucleosome and Chromatin Structure. *Nat. Struct. Mol. Biol.* 15 (10), 1122–1124. doi:10.1038/nsmb.1489

Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997). Crystal Structure of the Nucleosome Core Particle at 2.8 Å Resolution. *Nature* 389 (6648), 251–260. doi:10.1038/38444

Luger, K., Dechassa, M. L., and Tremethick, D. J. (2012). New Insights into Nucleosome and Chromatin Structure: an Ordered State or a Disordered Affair? *Nat. Rev. Mol. Cel Biol* 13 (7), 436–447. doi:10.1038/nrm3382

Machida, S., Takizawa, Y., Ishimaru, M., Sugita, Y., Sekine, S., Nakayama, J.-i., et al. (2018). Structural Basis of Heterochromatin Formation by Human HP1. *Mol. Cel* 69 (3), 385–397.e8. doi:10.1016/j.molcel.2017.12.011

Maeshima, K., Matsuda, T., Shindo, Y., Imamura, H., Tamura, S., Imai, R., et al. (2018). A Transient Rise in Free Mg2+ Ions Released from ATP-Mg Hydrolysis Contributes to Mitotic Chromosome Condensation. *Curr. Biol.* 28 (3), 444–451.e6. doi:10.1016/j.cub.2017.12.035

Maeshima, K., Ide, S., and Babokhov, M. (2019). Dynamic Chromatin Organization without the 30-nm Fiber. *Curr. Opin. Cel Biol.* 58, 95–104. doi:10.1016/j.ceb.2019.02.003

Mandal, A., Boatz, J. C., Wheeler, T. B., and van der Wel, P. C. A. (2017). On the Use of Ultracentrifugal Devices for Routine Sample Preparation in Biomolecular Magic-Angle-Spinning NMR. *J. Biomol. NMR* 67 (3), 165–178. doi:10.1007/s10858-017-0089-6

Mashtalir, N., Suzuki, H., Farrell, D. P., Sankar, A., Luo, J., Filipovski, M., et al. (2020). A Structural Model of the Endogenous Human BAF Complex Informs Disease Mechanisms. *Cell* 183 (3), 802–817.e24. doi:10.1016/j.cell.2020.09.051

Matlahov, I., and van der Wel, P. C. A. (2018). Hidden Motions and Motion-Induced Invisibility: Dynamics-Based Spectral Editing in Solid-State NMR. *Methods* 148, 123–135. doi:10.1016/j.ymeth.2018.04.015

McGinty, R. K., and Tan, S. (2021). Principles of Nucleosome Recognition by Chromatin Factors and Enzymes. *Curr. Opin. Struct. Biol.* 71, 16–26. doi:10.1016/j.sbi.2021.05.006

McGinty, R. K., Makde, R. D., and Tan, S. (2016). "Preparation, Crystallization, and Structure Determination of Chromatin Enzyme/Nucleosome Complexes," in *Methods in Enzymology* (New York, NY: Elsevier), 43–65. doi:10.1016/bs.mie.2016.01.003

Min, J., and Liu, K. (2021). Structures of Chromatin Modulators in Complex with Nucleosome. *Curr. Opin. Chem. Biol.* 63, 105–114. doi:10.1016/j.cbpa.2021.02.018

Mishra, L. N., Pepenella, S., Rogge, R., Hansen, J. C., and Hayes, J. J. (2016). Acetylation Mimics within a Single Nucleosome Alter Local DNA Accessibility in Compacted Nucleosome Arrays. *Sci. Rep.* 6 (1), 34808. doi:10.1038/srep34808

Morris, G. A., and Freeman, R. (1979). Enhancement of Nuclear Magnetic Resonance Signals by Polarization Transfer. *J. Am. Chem. Soc.* 101(3), 760–762. doi:10.1021/ja00497a058

Morrison, E. A., Baweja, L., Poirier, M. G., Wereszczynski, J., and Musselman, C. A. (2021). Nucleosome Composition Regulates the Histone H3 Tail Conformational Ensemble and Accessibility. *Nucleic Acids Res.* 49 (8), 4750–4767. doi:10.1093/nar/gkab246

Müller, M. M., and Muir, T. W. (2015). Histones: At the Crossroads of Peptide and Protein Chemistry. *Chem. Rev.* 115 (6), 2296–2349. doi:10.1021/cr5003529

Muir, T. W., Sondhi, D., and Cole, P. A. (1998). Expressed Protein Ligation: A General Method for Protein Engineering. *Proc. Natl. Acad. Sci.* 95 (12), 6705–6710. doi:10.1073/pnas.95.12.6705

Müller, M. M., Fierz, B., Bittova, L., Liszczak, G., and Muir, T. W. (2016). A Two-State Activation Mechanism Controls the Histone Methyltransferase Suv39h1. *Nat. Chem. Biol.* 12 (3), 188–193. doi:10.1038/nchembio.2008

Munowitz, M., Aue, W. P., and Griffin, R. G. (1982). Two-dimensional Separation of Dipolar and Scaled Isotropic Chemical Shift Interactions in Magic Angle NMR Spectra. *J. Chem. Phys.* 77 (4), 1686–1689. doi:10.1063/1.444064

Neumann, H., Peak-Chew, S. Y., and Chin, J. W. (2008). Genetically Encoding Nε-Acetyllysine in Recombinant Proteins. *Nat. Chem. Biol.* 4 (4), 232–234. doi:10.1038/nchembio.73

Nguyen, U. T. T., Bittova, L., Müller, M. M., Fierz, B., David, Y., Houck-Loomis, B., et al. (2014). Accelerated Chromatin Biochemistry Using DNA-Barcoded Nucleosome Libraries. *Nat. Methods* 11 (8), 834–840. doi:10.1038/nmeth.3022

Nozaki, T., Imai, R., Tanbo, M., Nagashima, R., Tamura, S., Tani, T., et al. (2017). Dynamic Organization of Chromatin Domains Revealed by Super-resolution Live-Cell Imaging. *Mol. Cel* 67 (2), 282–293.e7. doi:10.1016/j.molcel.2017.06.018

Ohno, M., Ando, T., Priest, D. G., Kumar, V., Yoshida, Y., and Taniguchi, Y. (2019). Sub-nucleosomal Genome Structure Reveals Distinct Nucleosome Folding Motifs. *Cell* 176 (3), 520–534.e25. doi:10.1016/j.cell.2018.12.014

Otterstrom, J., Castells-Garcia, A., Vicario, C., Gomez-Garcia, P. A., Cosma, M. P., and Lakadamyali, M. (2019). Super-resolution Microscopy Reveals How Histone Tail Acetylation Affects DNA Compaction within Nucleosomes *In Vivo*. *Nucleic Acids Res.* 47 (16), 8470–8484. doi:10.1093/nar/gkz593

Ou, H. D., Phan, S., Deerinck, T. J., Thor, A., Ellisman, M. H., and O'Shea, C. C. (2017). ChromEMT: Visualizing 3D Chromatin Structure and Compaction in Interphase and Mitotic Cells. *Science* 357 (6349), eaag0025. doi:10.1126/science.aag0025

Peng, Y., Li, S., Landsman, D., and Panchenko, A. R. (2021). Histone Tails as Signaling Antennas of Chromatin. *Curr. Opin. Struct. Biol.* 67, 153–160. doi:10.1016/j.sbi.2020.10.018

Pines, A., Gibby, M. G., and Waugh, J. S. (1973). Proton-Enhanced NMR of Dilute Spins in Solids. *J. Chem. Phys.* 59(2), 569–590. doi:10.1063/1.1680061

Pinto, F., Thornton, E. L., and Wang, B. (2020). An Expanded Library of Orthogonal Split Inteins Enables Modular Multi-Peptide Assemblies. *Nat. Commun.* 11 (1). doi:10.1038/s41467-020-15272-2

Pirman, N. L., Barber, K. W., Aerni, H. R., Ma, N. J., Haimovich, A. D., Rogulina, S., et al. (2015). A Flexible Codon in Genomically Recoded *Escherichia coli* Permits Programmable Protein Phosphorylation. *Nat. Commun.* 6, 8130. doi:10.1038/ncomms9130

Poepsel, S., Kasinath, V., and Nogales, E. (2018). Cryo-EM Structures of PRC2 Simultaneously Engaged with Two Functionally Distinct Nucleosomes. *Nat. Struct. Mol. Biol.* 25 (2), 154–162. doi:10.1038/s41594-018-0023-y

Poirier, M. G., Oh, E., Tims, H. S., and Widom, J. (2009). Dynamics and Function of Compact Nucleosome Arrays. *Nat. Struct. Mol. Biol.* 16 (9), 938–944. doi:10.1038/nsmb.1650

Prakash, K., and Fournier, D. (2017). Histone Code and Higher-Order Chromatin Folding: A Hypothesis. *Genomics Comput. Biol.* 3 (2), 41. doi:10.18547/gcb.2017.vol3.iss2.e41

Quinn, C. M., and Polenova, T. (2017). Structural Biology of Supramolecular Assemblies by Magic-Angle Spinning NMR Spectroscopy. *Quart. Rev. Biophys.* 50, e1. doi:10.1017/S0033583516000159

Rabdano, S. O., Shannon, M. D., Izmailov, S. A., Gonzalez Salguero, N., Zandian, M., Purusottam, R. N., et al. (2021). Histone H4 Tails in Nucleosomes: a Fuzzy Interaction with DNA. *Angew. Chem. Int. Ed.* 60 (12), 6480–6487. doi:10.1002/anie.202012046

Rao, S. S. P., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* 159 (7), 1665–1680. doi:10.1016/j.cell.2014.11.021

Ricci, M. A., Manzo, C., García-Parajo, M. F., Lakadamyali, M., and Cosma, M. P. (2015). Chromatin Fibers Are Formed by Heterogeneous Groups of Nucleosomes *In Vivo*. *Cell* 160 (6), 1145–1158. doi:10.1016/j.cell.2015.01.054

Risca, V. I., Denny, S. K., Straight, A. F., and Greenleaf, W. J. (2017). Variable Chromatin Structure Revealed by *In Situ* Spatially Correlated DNA Cleavage Mapping. *Nature* 541 (7636), 237–241. doi:10.1038/nature20781

Sanulli, S., Trnka, M. J., Dharmarajan, V., Tibble, R. W., Pascal, B. D., Burlingame, A. L., et al. (2019). HP1 Reshapes Nucleosome Core to Promote Phase Separation of Heterochromatin. *Nature* 575 (7782), 390–394. doi:10.1038/s41586-019-1669-2

Schubeis, T., Lührs, T., and Ritter, C. (2015). Unambiguous Assignment of Short- and Long-Range Structural Restraints by Solid-State NMR Spectroscopy with Segmental Isotope Labeling. *ChemBioChem* 16 (1), 51–54. doi:10.1002/cbic.201402446

Shah, N. H., and Muir, T. W. (2014). Inteins: Nature's Gift to Protein Chemists. *Chem. Sci.* 5 (2), 446–461. doi:10.1039/C3SC52951G

Shi, X., Prasanna, C., Nagashima, T., Yamazaki, T., Pervushin, K., and Nordenskiöld, L. (2018). Structure and Dynamics in the Nucleosome Revealed by Solid-State NMR. *Angew. Chem. Int. Ed.* 57 (31), 9734–9738. doi:10.1002/anie.201804707

Shi, X., Prasanna, C., Pervushin, K., and Nordenskiöld, L. (2020a). Solid-state NMR 13C, 15N Assignments of Human Histone H3 in the Nucleosome Core Particle. *Biomol. NMR Assign* 14 (1), 99–104. doi:10.1007/s12104-020-09927-w

Shi, X., Prasanna, C., Soman, A., Pervushin, K., and Nordenskiöld, L. (2020b). Dynamic Networks Observed in the Nucleosome Core Particles Couple the Histone Globular Domains with DNA. *Commun. Biol.* 3 (1), 639. doi:10.1038/s42003-020-01369-3

Shoaib, M., Chen, Q., Shi, X., Nair, N., Prasanna, C., Yang, R., et al. (2021). Histone H4 Lysine 20 Mono-Methylation Directly Facilitates Chromatin Openness and Promotes Transcription of Housekeeping Genes. *Nat. Commun.* 12 (1), 4800. doi:10.1038/s41467-021-25051-2

Shogren-Knaak, M. (2006). Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions. *Science* 311 (5762), 844–847. doi:10.1126/science.1124000

Simon, M. D., Chu, F., Racki, L. R., de la Cruz, C. C., Burlingame, A. L., Panning, B., et al. (2007). The Site-specific Installation of Methyl-Lysine Analogs into Recombinant Histones. *Cell* 128 (5), 1003–1012. doi:10.1016/j.cell.2006.12.041

Simpson, R. T., Thoma, F., and Brubaker, J. M. (1985). Chromatin Reconstituted from Tandemly Repeated Cloned DNA Fragments and Core Histones: A Model System for Study of Higher Order Structure. *Cell* 42 (3), 799–808. doi:10.1016/0092-8674(85)90276-4

Sinha, K. K., Gross, J. D., and Narlikar, G. J. (2017). Distortion of Histone Octamer Core Promotes Nucleosome Mobilization by a Chromatin Remodeler. *Science* 355 (6322), eaaa3761. doi:10.1126/science.aaa3761

Soman, A., Liew, C. W., Teo, H. L., Berezhnoy, N. V., Olieric, V., Korolev, N., et al. (2020). The Human Telomeric Nucleosome Displays Distinct Structural and Dynamic Properties. *Nucleic Acids Res.* 48 (10), 5383–5396. doi:10.1093/nar/gkaa289

Stevens, A. J., Brown, Z. Z., Shah, N. H., Sekar, G., Cowburn, D., and Muir, T. W. (2016). Design of a Split Intein with Exceptional Protein Splicing Activity. *J. Am. Chem. Soc.* 138 (7), 2162–2165. doi:10.1021/jacs.5b13528

Strickfaden, H., Tolsma, T. O., Sharma, A., Underhill, D. A., Hansen, J. C., and Hendzel, M. J. (2020). Condensed Chromatin Behaves like a Solid on the Mesoscale *In Vitro* and in Living Cells. *Cell* 183 (7), 1772–1784.e13. doi:10.1016/j.cell.2020.11.027

Strom, A. R., Biggs, R. J., Banigan, E. J., Wang, X., Chiu, K., Herman, C., et al. (2021). HP1α Is a Chromatin Crosslinker that Controls Nuclear and Mitotic Chromosome Mechanics. *Elife* 10, e63972. doi:10.7554/eLife.63972

Struppe, J., Quinn, C. M., Lu, M., Wang, M., Hou, G., Lu, X., et al. (2017). Expanding the Horizons for Structural Analysis of Fully Protonated Protein Assemblies by NMR Spectroscopy at MAS Frequencies above 100 kHz. *Solid State. Nucl. Magn. Reson.* 87, 117–125. doi:10.1016/j.ssnmr.2017.07.001

Stützer, A., Liokatis, S., Kiesel, A., Schwarzer, D., Sprangers, R., Söding, J., et al. (2016). Modulations of DNA Contacts by Linker Histones and Post-translational Modifications Determine the Mobility and Modifiability of Nucleosomal H3 Tails. *Mol. Cel* 61 (2), 247–259. doi:10.1016/j.molcel.2015.12.015

Su, J.-H., Zheng, P., Kinrot, S. S., Bintu, B., and Zhuang, X. (2020). Genome-Scale Imaging of the 3D Organization and Transcriptional Activity of Chromatin. *Cell* 182 (6), 1641–1659.e26. doi:10.1016/j.cell.2020.07.032

Teles, K., Fernandes, V., Silva, I., Leite, M., Grisolia, C., Lobbia, V. R., et al. (2020). Nucleosome Binding Peptide Presents Laudable Biophysical and *In Vivo* Effects. *Biomed. Pharmacother.* 121, 109678. doi:10.1016/j.biopha.2019.109678

van der Wel, P. C. A. (2018). New Applications of Solid-State NMR in Structural Biology. *Emerging Top. Life Sci.* 2 (1), 57–67. doi:10.1042/ETLS20170088

van Emmerik, C. L., and van Ingen, H. (2019). Unspinning Chromatin: Revealing the Dynamic Nucleosome Landscape by NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* 110, 1–19. doi:10.1016/j.pnmrs.2019.01.002

Wang, Z. A., and Liu, W. R. (2017). Proteins with Site-Specific Lysine Methylation. *Chem. Eur. J.* 23 (49), 11732–11737. doi:10.1002/chem.201701655

Weber, A. (2019). Evidence for and against Liquid-Liquid Phase Separation in the Nucleus. *ncRNA* 5 (4), 50. doi:10.3390/ncrna5040050

Wiegand, T., Cadalbert, R., von Schroetter, C., Allain, F. H.-T., and Meier, B. H. (2018). Segmental Isotope Labelling and Solid-State NMR of a 12 × 59 kDa Motor Protein: Identification of Structural Variability. *J. Biomol. NMR* 71 (4), 237–245. doi:10.1007/s10858-018-0196-z

Wiegand, T., Lacabanne, D., Torosyan, A., Boudet, J., Cadalbert, R., Allain, F. H.-T., et al. (2020). Sedimentation Yields Long-Term Stable Protein Samples as Shown by Solid-State NMR. *Front. Mol. Biosci.* 7, 17. doi:10.3389/fmolb.2020.00017

Woodcock, C. L. F., Safer, J. P., and Stanchfield, J. E. (1976). Structural Repeating Units in Chromatin. *Exp. Cel Res.* 97 (1), 101–110. doi:10.1016/0014-4827(76)90659-5

Xiang, S., le Paige, U. B., Horn, V., Houben, K., Baldus, M., and van Ingen, H. (2018). Site-Specific Studies of Nucleosome Interactions by Solid-State NMR Spectroscopy. *Angew. Chem. Int. Ed.* 57 (17), 4571–4575. doi:10.1002/anie.201713158

Xu, J., Ma, H., Jin, J., Uttam, S., Fu, R., Huang, Y., et al. (2018). Super-Resolution Imaging of Higher-Order Chromatin Structures at Different Epigenomic States in Single Mammalian Cells. *Cel Rep.* 24 (4), 873–882. doi:10.1016/j.celrep.2018.06.085

Yan, L. Z., and Dawson, P. E. (2001). Synthesis of Peptides and Proteins without Cysteine Residues by Native Chemical Ligation Combined with Desulfurization. *J. Am. Chem. Soc.* 123 (4), 526–533. doi:10.1021/ja003265m

Yang, J. G., Madrid, T. S., Sevastopoulos, E., and Narlikar, G. J. (2006). The Chromatin-Remodeling Enzyme ACF Is an ATP-dependent DNA Length Sensor that Regulates Nucleosome Spacing. *Nat. Struct. Mol. Biol.* 13 (12), 1078–1083. doi:10.1038/nsmb1170

Zandian, M., Gonzalez Salguero, N., Shannon, M. D., Purusottam, R. N., Theint, T., Poirier, M. G., et al. (2021). Conformational Dynamics of Histone H3 Tails in Chromatin. *J. Phys. Chem. Lett.* 12, 6174–6181. doi:10.1021/acs.jpclett.1c01187

Zhou, J., Fan, J. Y., Rangasamy, D., and Tremethick, D. J. (2007). The Nucleosome Surface Regulates Chromatin Compaction and Couples it with Transcriptional Repression. *Nat. Struct. Mol. Biol.* 14 (11), 1070–1076. doi:10.1038/nsmb1323

Zhou, B.-R., Feng, H., Ghirlando, R., Kato, H., Gruschus, J., and Bai, Y. (2012). Histone H4 K16Q Mutation, an Acetylation Mimic, Causes Structural Disorder of its N-Terminal Basic Patch in the Nucleosome. *J. Mol. Biol.* 421 (1), 30–37. doi:10.1016/j.jmb.2012.04.032

Zhou, K., Gaullier, G., and Luger, K. (2019). Nucleosome Structure and Dynamics Are Coming of Age. *Nat. Struct. Mol. Biol.* 26 (1), 3–13. doi:10.1038/s41594-018-0166-x

Zoukimian, C., Meudal, H., De Waard, S., Ouares, K. A., Nicolas, S., Canepari, M., et al. (2019). Synthesis by Native Chemical Ligation and Characterization of the Scorpion Toxin AmmTx3. *Bioorg. Med. Chem.* 27 (1), 247–253. doi:10.1016/j.bmc.2018.12.009

Züger, S., and Iwai, H. (2005). Intein-based Biosynthetic Incorporation of Unlabeled Protein Tags into Isotopically Labeled Proteins for NMR Studies. *Nat. Biotechnol.* 23 (6), 736–740. doi:10.1038/nbt1097

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

REVIEW
published: 20 October 2021
doi: 10.3389/fmolb.2021.743181

Strategies for RNA Resonance Assignment by $^{13}C/^{15}N$- and ^{1}H-Detected Solid-State NMR Spectroscopy

*Philipp Innig Aguion and Alexander Marchanka**

Institute for Organic Chemistry and Centre of Biomolecular Drug Research (BMWZ), Leibniz University Hannover, Hanover, Germany

Magic angle spinning (MAS) solid-state NMR (ssNMR) is an established tool that can be applied to non-soluble or non-crystalline biomolecules of any size or complexity. The ssNMR method advances rapidly due to technical improvements and the development of advanced isotope labeling schemes. While ssNMR has shown significant progress in structural studies of proteins, the number of RNA studies remains limited due to ssNMR methodology that is still underdeveloped. Resonance assignment is the most critical and limiting step in the structure determination protocol that defines the feasibility of NMR studies. In this review, we summarize the recent progress in RNA resonance assignment methods and approaches for secondary structure determination by ssNMR. We critically discuss advantages and limitations of conventional ^{13}C- and ^{15}N-detected experiments and novel ^{1}H-detected methods, identify optimal regimes for RNA studies by ssNMR, and provide our view on future ssNMR studies of RNA in large RNP complexes.

Keywords: RNA, solid-state NMR, assignment, resonances, MAS

OPEN ACCESS

Edited by:
Amir Goldbourt,
Tel Aviv University, Israel

Reviewed by:
Ansgar B Siemer,
University of Southern California, United States
Shenlin Wang,
East China University of Science and Technology, China

****Correspondence:***
Alexander Marchanka
alexander.marchanka@oci.uni-hannover.de

Specialty section:
This article was submitted to Structural Biology, a section of the journal Frontiers in Molecular Biosciences

Received: *17 July 2021*
Accepted: *03 September 2021*
Published: *20 October 2021*

Citation:
Aguion PI and Marchanka A (2021) Strategies for RNA Resonance Assignment by $^{13}C/^{15}N$- and ^{1}H-Detected Solid-State NMR Spectroscopy. Front. Mol. Biosci. 8:743181. doi: 10.3389/fmolb.2021.743181

INTRODUCTION

In the last 2 decades, biomolecular solid-state NMR (ssNMR) spectroscopy has obtained a massive boost both from the progress in technical development, particularly with the advent of ultrafast magic angle spinning (MAS) probes, and from the introduction of novel isotope labeling techniques (Lu et al., 2010; Atreya, 2012; Marchanka et al., 2018a). New ssNMR studies of challenging biomolecular systems using cutting-edge technologies are being reported at an increased pace with most recent datasets having been acquired using spectrometers at the highest possible field strengths (1.2 GHz) (Callon et al., 2021; Nimerovsky et al., 2021) and under ultrafast MAS rates (111 kHz and, most recently, 150 kHz) (Penzel et al., 2019; Schledorn et al., 2020). ssNMR, in contrast to solution-state NMR, does not suffer from molecular weight (MW) limitations and therefore can be applied to various biomolecules, including membrane proteins, amyloid fibrils, and protein–protein assemblies (Castellani et al., 2002; Shi et al., 2011; Andreas et al., 2016; Quinn and Polenova, 2017). While structural ssNMR studies on proteins are well established, similar studies on free nucleic acids and nucleic acid parts of biomolecular complexes have been initiated significantly later and remain scarce due to as yet underdeveloped methodology and challenges arising from the significant spectral overlap of nucleic acid resonances (Marchanka and Carlomagno, 2014; Sreemantula and Marchanka, 2020). Nevertheless, a few impressive studies on nucleic acids have been performed in the last 2 decades. ssNMR studies of RNAs were pioneered by the Görlach

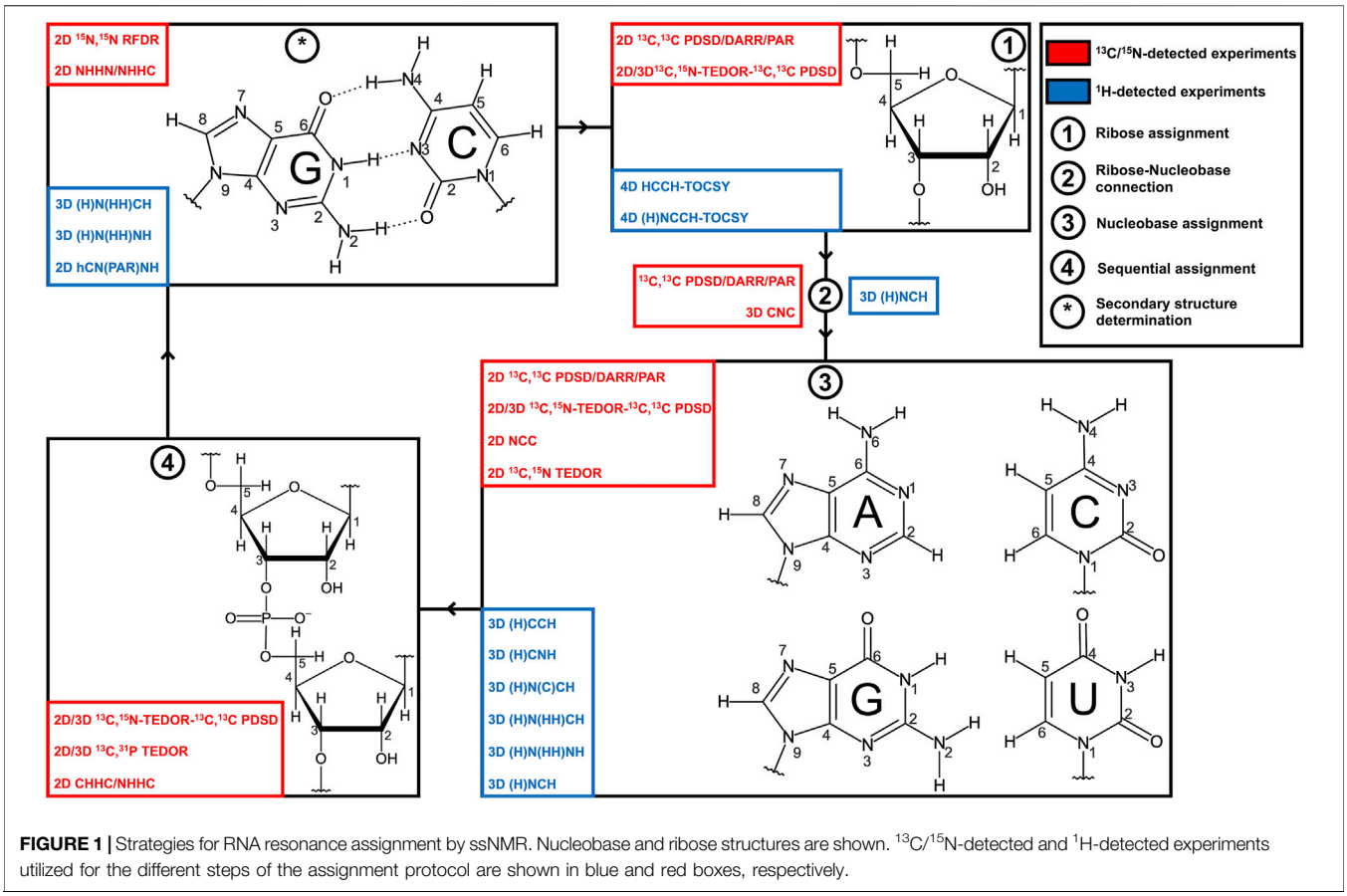


FIGURE 1 | Strategies for RNA resonance assignment by ssNMR. Nucleobase and ribose structures are shown. $^{13}C/^{15}N$-detected and ^{1}H-detected experiments utilized for the different steps of the assignment protocol are shown in blue and red boxes, respectively.

group, who have not only contributed to the methodological development but also provided important insights into the structure of the ~100 kDa $(CUG)_{97}$ RNA repeat involved in the neuromuscular disease myotonic dystrophy (Leppert et al., 2004; Riedel et al., 2005b; a; Riedel et al., 2006). The Drobny laboratory has used ssNMR to study the structure and dynamics of 29mer HIV TAR RNA bound to an 11mer Tat peptide using site-specific ^{19}F RNA labeling (Olsen et al., 2005; Huang et al., 2010; Olsen et al., 2010).

The first complete assignment of RNA resonances along with the first 3D structure of RNA in a protein–RNA complex established solely by ssNMR spectroscopy was obtained by the Carlomagno laboratory and was a major milestone in the development of ssNMR for RNAs (Marchanka et al., 2013; Marchanka et al., 2015). Our and Carlomagno's laboratories have also been active in the characterization of protein–RNA interfaces, and we have recently determined the structure of the same protein–RNA complex by a combination of paramagnetic relaxation enhancement (PRE)-aided ssNMR and chemical shift perturbation (CSP) analysis (Ahmed et al., 2020). Finally, the first studies on RNA at 40 kHz MAS (Yang et al., 2017; Zhao et al., 2019) and MAS ≥ 100 kHz (Marchanka et al., 2018b; Aguion et al., 2021) have been performed, reporting both assignment of resonances and identification of base pairs by sensitive ^{1}H-detected ssNMR spectroscopy. In the research field of phage viruses, the Goldbourt laboratory has obtained nucleotide-type assignment for very large native DNA (Sergeyev et al., 2011; Morag et al., 2014; Goldbourt, 2019) and has recently identified base pairs in native 1.2 MDa RNA from the bacteriophage MS2 (Lusky et al., 2021).

Structural determination of RNA by ssNMR comprises several steps that include sample preparation, spin system–specific (assignment of resonances within a nucleotide spin system) and site-specific/sequential (assignment of defined spin systems to a specific residue within the RNA) assignment of resonances, and the collection of distance and angular restraints which are then used in structural calculations (Marchanka and Carlomagno, 2019) (**Figure 1**). While in some studies, unambiguous assignment of RNA resonances is not necessary to obtain valuable structural information (Olsen et al., 2005; Huang et al., 2010, 2011; Yang et al., 2017; Lusky et al., 2021), in most cases, site-specific assignment of resonances is the main limiting and crucial step in structure determination by NMR.

In this review, we describe the methods for the resonance assignment of RNA by ssNMR and compare them with solution-state NMR approaches. We provide a comprehensive description of ssNMR experiments suitable for the spin system–specific assignment of riboses and nucleobases and identification of the base-pairing pattern in RNA. Furthermore, we briefly discuss ssNMR experiments for the sequential assignment of RNA. We examine conventional ^{13}C-detected and novel ^{1}H-detected ssNMR methods, critically assess their strengths

and limitations at different MAS frequencies, and discuss optimal MAS regimes for ssNMR studies of RNA.

Can RNA Resonances Always Be Assigned by ssNMR?

RNAs for NMR studies are typically prepared by *in vitro* transcription (Milligan et al., 1987; Milligan and Uhlenbeck, 1989) or chemical synthesis (Beaucage and Reese, 2009). While chemical synthesis can produce RNA with sophisticated site-specific labeling, this method is limited to RNA of ~70 nt in length and is not available in most laboratories. On the other hand, *in vitro* transcription can deliver RNA of any length labeled uniformly or selectively by nucleotide type and is potentially accessible to any laboratory. In this review, we will mostly discuss experiments suitable for the assignment of RNA that is either uniformly labeled or selectively labeled by nucleotide type.

Solution-state NMR studies of RNA have an intrinsic MW limit of ~40 kDa (120–150 nt) and larger RNA can be assigned only partially; advanced structural studies on large RNA use many differently labeled samples and sophisticated experiments (Lu et al., 2010; Keane et al., 2015; Brown et al., 2020). While ssNMR can, in principle, be applied to RNA of any size, the feasibility of ssNMR studies is determined by the quality of sample preparation, which has a direct impact on the spectral linewidth and therefore on spectral crowding. It is commonly accepted that ^{13}C linewidths smaller than 1 ppm are necessary to perform resonance assignment and obtain structural data on non–site-specific labeled samples. Due to the limited number of ssNMR RNA studies, statistics on the quality of different sample preparation techniques are very scarce; nevertheless, some patterns have been identified. Typical linewidths of lyophilized sample preparations are significantly larger than 1 ppm (Olsen et al., 2005; Huang et al., 2011) since insufficient hydration ultimately leads to large inhomogeneous broadening. As stated, linewidths greater than 1 ppm are not sufficient for the resonance assignment, which is, however, not necessary in studies utilizing site-specifically labeled RNAs (Olsen et al., 2005).

The most commonly used sample preparation method of RNA labeled uniformly or selectively by nucleotide type for both ^{13}C-detected and ^{1}H-detected ssNMR studies is micro (nano)-crystallization. This technique was developed for ssNMR studies of proteins (McDermott et al., 2000; Franks et al., 2005; Bertini et al., 2010) and has been successfully applied to study RNA (Huang et al., 2012; Marchanka et al., 2013; Yang et al., 2017). This method provides typical ^{13}C linewidths of 1 ppm (29mer HIV TAR RNA) (Huang et al., 2012) or even 0.5 ppm for the 26mer box C/D RNA in complex with L7Ae protein (Marchanka et al., 2013; Marchanka et al., 2015). The same complex shows an ^{1}H linewidth of separated C1′-H1′ resonances in the protonated ribose of 150 Hz (0.15 ppm) by ^{1}H-detection at 100 kHz MAS on a 1 GHz spectrometer, while the ^{1}H linewidth of imino resonances on an 850 MHz spectrometer estimates to 200–300 Hz (0.23–0.35 ppm) (Aguion et al., 2021). In the ^{1}H-detected imino spectrum of 26mer DIS-HIV-1 RNA acquired at 40 kHz MAS, the linewidths of the ^{1}H and ^{15}N dimensions are equal to 500 Hz (0.9 ppm) and 70 Hz (1 ppm), respectively (Yang et al., 2017).

A novel ethanol precipitation method introduced by the Wang group (Zhao et al., 2019) utilizing 75% D_2O/25% H_2O-based buffer yields a promising ^{15}N linewidth of 80 Hz (1.3 ppm) and 70 Hz (1.2 ppm) for imino resonances of 26mer DIS-HIV-1 RNA and 71mer RiboA71 domain of *add* adenine riboswitch, respectively. The ^{1}H linewidth of imino resonances in these two RNAs was generously estimated to 230 Hz (0.38 ppm) and 132 Hz (0.22 ppm), respectively. While RiboA71 showed a good chemical shift dispersion so that different spin systems could be identified and even tentatively assigned based on the known solution-state chemical shifts (Zhao et al., 2019), the chemical shift range for DIS-HIV-1 was very narrow and no identification of individual spin systems was possible. It can be speculated that for the isolated 26mer DIS-HIV-1 RNA, the tertiary structure and the local order are partially lost upon EtOH precipitation, while for the well-folded riboswitch RiboA71, the tertiary structure is preserved. Further investigations into this matter are required, but this approach undoubtedly added a new technique to the repertoire of ssNMR RNA sample preparations. Unfortunately, this method cannot be applied toward protein–RNA complexes due to the instant precipitation of most proteins under such conditions.

While both microcrystallization and ethanol precipitation methods demonstrate the general feasibility of RNA resonance assignment by ssNMR, the typically obtained linewidths are significantly worse than those in solution-state NMR (< 0.1 ppm for ^{1}H). It implies that nucleotide-type selective labeling is necessary to perform ^{13}C-detected ssNMR studies even on short (< 30 nt) RNA (Marchanka et al., 2018a), while ^{1}H-detection at MAS > 60 kHz allows us to study such RNA using a single uniformly labeled sample (Marchanka et al., 2018b). For larger RNA, site-specific or segmental labeling of short (< 30 nt) RNA stretches should be utilized to make both ^{13}C-detected and ^{1}H-detected ssNMR studies feasible (Marchanka et al., 2018a).

Is There Always a Need for Complete Resonance Assignment of RNA?

Solution-state NMR spectroscopy provides rapid information about the secondary structure of RNA and identifies canonical Watson–Crick (WC) and non-WC base pairs by observation of characteristic chemical shifts of immobilized amino (NH_2) and imino (NH) groups (Wacker et al., 2020). In solution-state NMR, site-specific assignment of ^{13}C resonances is typically not necessary to site-specifically assign imino resonances and identify the RNA secondary structure. 2D ^{1}H,^{1}H imino NOESY/3D ^{1}H,^{15}N HMQC-^{1}H,^{1}H NOESY (Nikonowicz and Pardi, 1993) coupled with HNN-COSY (Dingley and Grzesiek, 1998; Pervushin et al., 1998) and ^{1}H,^{15}N-TROSY (Favier and Brutscher, 2011) experiments provide imino assignment and allow imino–imino sequential walk for the nucleotides in the base-paired and/or stacked regions. However, if complete sequential assignment and, especially, determination of the three-dimensional structure of RNA are aimed at, full assignment of RNA resonances is necessary. This task

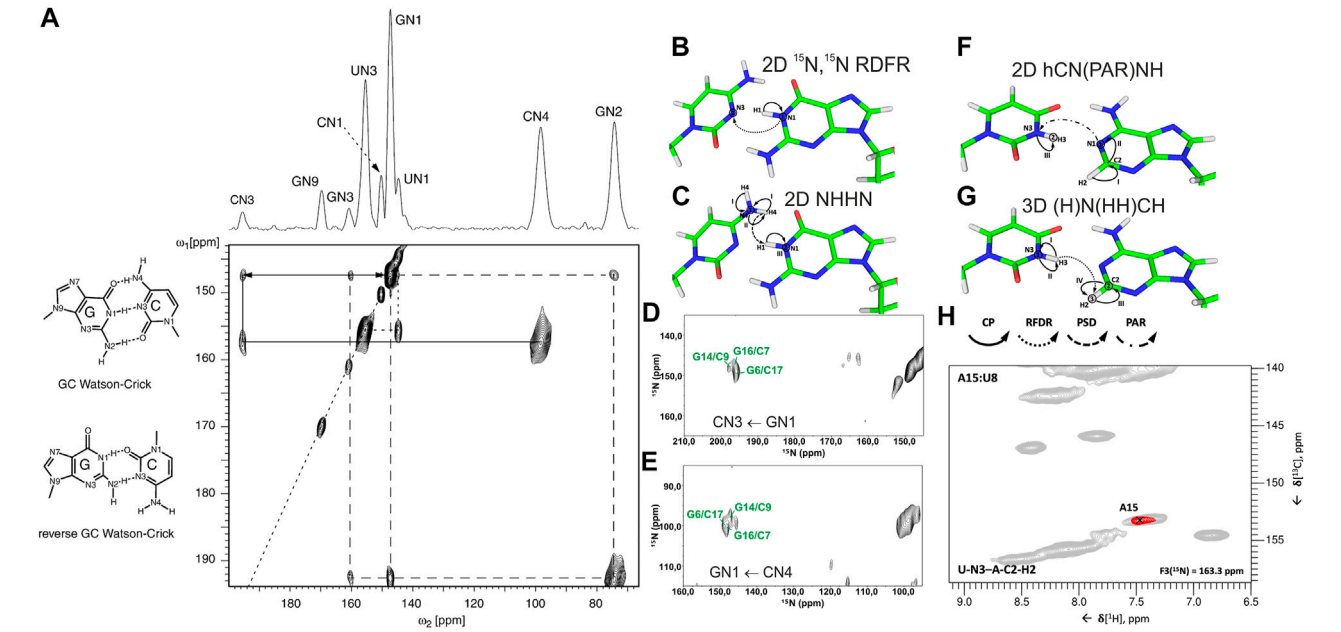


FIGURE 2 | Base pair identification and spin system–specific assignment of imino nitrogens by ssNMR. **(A)** Identification of WC G:C base pair in $(CUG)_{97}$ RNA repeat by $^{15}N,^{15}N$ RFDR experiment (Leppert et al., 2004). Magnetization transfer schemes for **(B)** $^{15}N,^{15}N$ recoupling (RFDR, PAR) experiment and **(C)** NHHN experiment shown on the example of WC G:C base pair. **(D)** 2D $^{15}N,^{15}N$ RFDR and **(E)** 2D NHHN spectra of G,C^{lab} 26mer box C/D RNA showing three assigned WC G:C base pairs (Marchanka et al., 2015). Magnetization transfer schemes for **(F)** 2D hCN(PAR)NH (Yang et al., 2017) and **(G)** 3D (H)N(HH)CH (Aguion et al., 2021) experiments **(H)** Representative 2D plane from the 3D (H)N(HH)CH spectrum shows WC base pair A15:U8 in 26mer Box C/D RNA. ^{15}N frequency of N3 nitrogen in uracil is indicated. In magnetization transfer schemes, solid arrows indicate CP transfers, dotted arrows indicate homonuclear RFDR recoupling, the dashed arrow indicates PSD transfer, and the dash-dotted arrow indicates PAR transfer; Roman numbers indicate CP transfers; Arabic numbers correspond to the spectral dimensions (t_1–t_3). Figure **(A)** is reproduced from J. Leppert, C. R. Urbinati, S. Häfner, O. Ohlenschläger, M. S. Swanson, M. Görlach, and R. Ramachandran, Identification of NH...N hydrogen bonds by magic angle spinning solid state NMR in a double stranded RNA associated with myotonic dystrophy. Nucleic Acids Res., 2004, 32, 3, 1,177–1,183, by permission of Oxford University Press. Spectrum in **(D)** is adapted from (Marchanka et al., 2015). Spectrum in **(E)** is reprinted from (Marchanka and Carlomagno, 2019) with permission from Elsevier.

comprises several steps and includes spin system–specific assignment of all ribose atoms, determination of ribose–base connections, and assignment of base resonances, followed by sequential assignment *via* ^{13}C-edited/filtered $^{1}H,^{1}H$ NOEs (Zwahlen et al., 1997; Breeze, 2000). Detailed description of solution-state NMR methods for the RNA assignment can be obtained from the classic work by the Schwalbe group (Fürtig et al., 2003) and from a few recent reviews (Scott and Hennig, 2008; Barnwal et al., 2017).

Since dipolar couplings are preserved in solid-state NMR, they can be utilized to provide rapid insights into the secondary structure of RNA by direct observation of base–base correlations. In contrast to solution-state NMR, ssNMR spectra not only show resonances from the base-paired nucleotides but also from any other immobilized nucleotides, so that observation of amino or imino resonances is not necessarily indicative of a base pairing.

In conventional $^{13}C/^{15}N$-detected ssNMR spectroscopy at MAS < 20 kHz, base pair information can be obtained directly by measuring ^{15}N–^{15}N correlations between base-paired nucleotides, either through direct dipolar transfers *via* radiofrequency-driven dipolar recoupling (RFDR) (Bennett et al., 1992) or proton-assisted recoupling (PAR) (Lewandowski et al., 2009), or *via* spin diffusion (SD)-based experiments, for example, proton-driven SD (PDSD) (Szeverenyi et al., 1982) and dipolar-assisted rotational resonance (DARR) (Takegoshi et al., 2001), or *via* proton spin diffusion (PSD) NHHN/NHHC experiments (Lange et al., 2002; Riedel et al., 2005a; Herbst et al., 2008). The Görlach laboratory has directly observed canonical WC G:C base pairs in $(CUG)_{97}$ RNA (**Figure 2A**), employing both $^{15}N,^{15}N$ RFDR (Leppert et al., 2004) (**Figure 2B**) and NHHN (Riedel et al., 2005a) experiments (**Figure 2C**). In our study on 26mer box C/D RNA, we have acquired both RFDR and NHHN spectra that were sufficient for the detection of WC G:C (**Figures 2D,E**) and A:U base pairs. Site-specific assignment *via* imino–imino sequential walk was not possible in ^{15}N-detected spectra due to poor ^{15}N chemical shift dispersion and typical ^{15}N linewidths ≥ 1 ppm. Recently, the Goldbourt laboratory has employed $^{15}N,^{15}N$ PDSD and $^{15}N,^{15}N$ PDSD·RFDR experiments to identify the presence of WC G:C and wobble G:U base pairs in native 1 MDa-sized bacteriophage MS2 RNA and obtained nucleotide-type–specific assignments (Lusky et al., 2021). In all ^{15}N-detected experiments described above, after the initial ^{1}H-^{15}N cross-polarization (CP), magnetization is evolved on ^{15}N during t_1. Following that, ^{15}N magnetization is spread to nearby nitrogens, either directly by various ^{15}N–^{15}N

recoupling schemes (**Figure 2B**) or indirectly through protons *via* the N→H-PSD-H→N scheme (**Figure 2C**). Finally, the ^{15}N magnetization is recorded during t_2.

^{1}H-detected ssNMR on RNA amino and imino groups is possible in two different regimes depending on the MAS frequency used. Since imino and amino protons are exchangeable, high level of deuteration will reduce the network of ^{1}H, ^{1}H dipolar couplings and therefore make ^{1}H resonances observable even at MAS frequencies of ~20 kHz. The Reif and Carlomagno groups used this approach to acquire an ^{1}H-^{15}N dipolar-based CP-HSQC (Zhou et al., 2007) spectrum of deuterated 26mer box C/D RNA in complex with L7Ae protein at 24 kHz MAS in 90% D_2O buffer (Asami et al., 2013). Although the recorded spectrum showed disperse imino resonances of several nucleotides, sequence-specific assignment *via* sequential walk has not been attempted due to low sensitivity. Recently, the Wang laboratory (Yang et al., 2017) has acquired proton-detected ^{15}N,^{15}N PAR correlations at 40 kHz MAS to obtain information about WC G:C and A:U base pairs. In their hCN(PAR)NH experiment, after initial CP transfer from ^{1}H to ^{13}C, a specific ^{13}C-^{15}N CP step is used to transfer magnetization to nitrogen atoms, whose chemical shifts are recorded in t_1. Magnetization transfer across the base pair is achieved by ^{15}N–^{15}N PAR transfer with a contact time of 7 ms. A final CP transfer brings the magnetization back to ^{1}H for detection (**Figure 2F**). Their approach was successful as they were able to confirm the presence of WC G:C and A:U base pairs in the 26mer DIS-HIV-1. Unfortunately, the low resolution in the ^{1}H dimension (~500 Hz for imino protons) did not allow the identification of different spin systems and thus the unambiguous assignment of hydrogen bonds. In our recent study at 100 kHz MAS, in view of good linewidths in the proton dimension (200–300 Hz), spin system–specific detection of base pairs was possible (Aguion et al., 2021), allowing rapid identification of almost all base pairs present in 26mer box C/D RNA. While the 2D/3D (H)N(HH)NH experiment detects WC G:C and non-WC U:U base pairs, the 2D/3D (H)N(HH)CH experiment identifies WC A:U and non-WC G:A base pairs. In the 2D/3D (H)N(HH)CH experiment (**Figure 2G**), after initial ^{1}H-^{15}N CP transfer, ^{15}N chemical shifts are evolved during t_1. After the second CP has transferred magnetization back to the protons, a ^{1}H,^{1}H RFDR of 0.48–0.96 ms spreads magnetization to all nearby protons within a distance of 3–4 Å. The third CP transfers magnetization to ^{13}C for an optional evolution (in 3D experiment). Finally, a short read-out CP step transfers magnetization back to the protons for detection during t_2 (t_3 in 3D). In the 2D/3D (H)N(HH)NH experiment, the third CP step transfers magnetization to the directly attached nitrogens, where their chemical shifts evolve during t_2 (in the optional 3D experiment). The final short ^{15}N-^{1}H CP read-out step transfers magnetization back to the protons for detection during t_2 (t_3 in 3D). Despite efficient identification of both WC and non-WC base pairs (**Figure 2H**), sequential imino–imino walk as performed in solution-state NMR was not feasible due to 1) low sensitivity and 2) strong signal overlap.

Presented case studies show that ssNMR allows rapid identification of the type and number of base pairs present in the RNA. However, sequence-specific assignment of base pairs and therefore RNA secondary structure determination based on amino and imino fingerprints alone are not possible and, hence, complete assignment of RNA resonances is necessary. The above will require their correlation with the well-resolved nucleobase carbons C6/C8 (C6-H6/C8-H8 groups) and then with the ribose C1′ (C1′-H1′ groups). Finally, nucleotides should be connected sequentially through ^{1}H–^{1}H and/or ^{13}C–^{13}C correlations (**Figure 1**).

In the following sections, we will discuss in detail experimental strategies for the resonance assignment of different RNA moieties by ssNMR.

Ribose Assignment

Spin system–specific assignment of ribose resonances by conventional ^{13}C-detected ssNMR at MAS frequencies < 20 kHz can be achieved using a multitude of correlation schemes provided a satisfactory spectral linewidth is obtained. In the pioneering study by the Görlach group, mostly intra-ribose and partially ribose–base correlations were obtained using symmetry-based adiabatic ZQ recoupling experiments (Riedel et al., 2004). However, due to the low chemical shift resolution arising from limitations in sample preparation, spin system–specific assignment was not possible despite only three different nucleotides being present in the $(CUG)_{97}$ RNA. The Drobny group (Huang et al., 2012) acquired ^{13}C,^{13}C PDSD (Szeverenyi et al., 1982) experiments on the selectively uracil-labeled TAR RNA. Also, there, very narrow carbon chemical shift dispersion together with poor resolution impeded any spin system–specific resonance assignments. In our study on 26mer box C/D RNA, we have exploited 2D ^{13}C,^{13}C PDSD (**Figure 3A**) to correlate intra-ribose resonances and even to obtain ribose–base correlations (Marchanka et al., 2013). While 100 ms PDSD mixing was sufficient to obtain a full set of intra-ribose correlations (**Figure 3B**), 500 ms mixing additionally provided not only intra-base correlations but also an almost complete set of ribose–base and several inter-nucleotide correlations (s. below). Despite good ^{13}C linewidths of 0.5 ppm and usage of nucleotide-type selective labeling (Marchanka et al., 2013; Marchanka et al., 2018a), homonuclear ^{13}C,^{13}C spectra yielded the assignment of less than half of the nucleotides of 26mer box C/D RNA. The assignment process is hampered particularly by significant spectral crowding of C2′/C3′ carbons, and furthermore, the lack of a proton dimension does not help to lift the ambiguity.

A MAS regime of 20 kHz < ω_R < 60 kHz is less suitable for the ribose assignment in RNA labeled uniformly or selectively by nucleotide type. First, since ribose protons are not exchangeable, acquisition of NMR spectra in deuterated buffer does not bear any advantage. Coherence lifetimes of ribose protons are unfavorable at MAS frequencies < 60 kHz (Marchanka et al., 2018b), so that ^{13}C-detected ssNMR must be applied. While 3.2 mm probes (maximum MAS frequency = 24 kHz) have optimal ^{13}C sensitivity, the smaller rotor size in 2.5 and 1.7 mm probes attenuates ^{13}C sensitivity due to a smaller sample volume. Second, pure SD-based experiments (PDSD

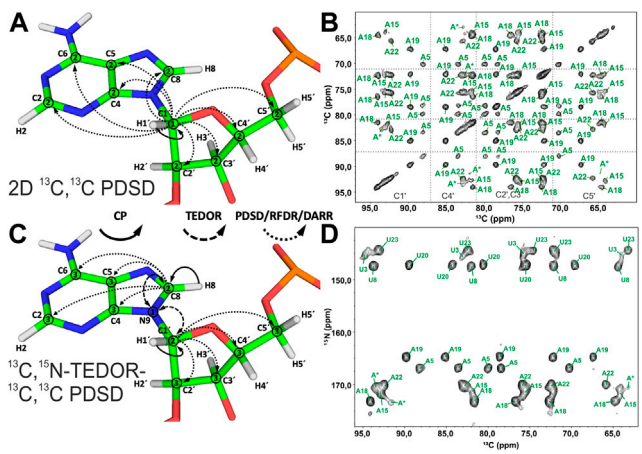


FIGURE 3 | Ribose assignment by ^{13}C-detected ssNMR. **(A)** Magnetization transfer in homonuclear ^{13}C,^{13}C recoupling experiment (PDSD, DARR, and PAR) for the spin system–specific assignment of carbon resonances in nucleotides shown on the example of an adenosine. **(B)** Zoom in on the ribose region of ^{13}C,^{13}C PDSD spectra of the A^{lab}-26mer box C/D RNA measured at the mixing time of 100 ms. **(C)** Magnetization transfer in heteronuclear ^{13}C,^{15}N-TEDOR-^{13}C,^{13}C PDSD experiment shown on the example of an adenosine. **(D)** Ribose region of the ^{13}C,^{15}N-TEDOR-^{13}C,^{13}C PDSD spectrum of the A,U^{lab}-26mer box C/D RNA measured at TEDOR and PDSD mixing times of 3.2 and 100 ms, respectively. In magnetization transfer schemes, solid arrows indicate CP transfers, dashed arrows indicated TEDOR transfer, and dotted arrows indicate homonuclear ^{13}C,^{13}C recoupling; Arabic numbers correspond to the spectral dimensions (t_1–t_3). The spectrum in **(B)** is adapted with permission from (Marchanka et al., 2013), © John Wiley and Sons, 2013. The spectrum in **(D)** is adapted from (Marchanka et al., 2015).

and DARR) cannot work well at MAS > 20 kHz; therefore, advanced recoupling schemes must be utilized to correlate ribose resonances in this regime. These mixing schemes may include finite-pulse (fp) RFDR (Bennett et al., 1992; Nishiyama et al., 2014), R-symmetry–driven SD (Hou et al., 2011), mixed rotational and rotary resonance (MIRROR) (Scholz et al., 2008), phase-alternated recoupling irradiation (PARIS) (Weingarth et al., 2009), combined R2-symmetry–driven sequences (CORD) (Hou et al., 2013), and many others. A recent review by the Hou group provides a comprehensive description of different recoupling techniques at fast MAS (Ji et al., 2021). Many first-order homonuclear dipolar recoupling sequences (e.g., fpRFDR) suffer from dipolar truncation (Bayro et al., 2009), which may impair observation of long-range correlations, whereas second-order sequences (SD-based, RF-assisted SD, and PAR) are free from this effect (Ji et al., 2021). Regardless of the chosen type of homonuclear recoupling, all ^{13}C,^{13}C correlation experiments have a similar form. After an initial short CP transfer from ^{1}H to ^{13}C, magnetization is evolved on the starting ^{13}C during t_1. Subsequently, carbons are connected by one of the recoupling schemes. Finally, ^{13}C chemical shifts are recorded during t_2 (**Figure 3A**).

Since N1 and N9 nitrogens in pyrimidines and purines, respectively, have very distinct chemical shifts, the ^{15}N dimension can be added to improve assignment by separation of ribose resonances of different nucleotide types. Optimal separation can be achieved by acquisition of, for example, ^{13}C,^{15}N TEDOR (Jaroniec et al., 2002) or ^{13}C,^{15}N TEDOR-^{13}C,^{13}C PDSD (Riedel et al., 2005b; Daviso et al., 2013) experiments. In the TEDOR-PDSD experiment, the ^{13}C magnetization (C1′) is prepared *via* a short ^{1}H,^{13}C CP step. During a short ^{13}C,^{15}N TEDOR transfer, the magnetization is propagated to nearby nitrogens (N1 and N9 in pyrimidines and purines, respectively). After t_1 evolution on ^{15}N, the magnetization is transferred back to ^{13}C. A subsequent ^{13}C,^{13}C PDSD or DARR step spreads the magnetization into ^{13}C spins of ribose (C2′-C5′), whose chemical shifts are recorded during t_2 (t_3 in 3D) (**Figure 3C**). An optional ^{13}C evolution step before PDSD (DARR) yields the 3D TEDOR-PDSD experiment which improves resolution at the cost of sensitivity. In addition to improving ribose assignment (**Figure 3D**), TEDOR-PDSD provides ribose–base connections and improves the assignment of nucleobase carbons (*vide infra*).

^{1}H-detected ssNMR at MAS frequencies > 60 kHz and in particularly ≥ 100 kHz opens new avenues for structural characterization of biomolecules, significantly improving resolution and increasing the sensitivity per unit of the sample. Such probeheads operate with significantly smaller rotors that ultimately reduce the sample volume/number of spins packed into the ssNMR rotor and therefore attenuate sensitivity. Due to the optimized coil sensitivity, increased fill factor, and narrowed lines, overall sensitivity is not reduced as a cube of the rotor diameter (Schledorn et al., 2020); nevertheless, exclusively ^{1}H detection with the sensitivity increased by a factor of $(\gamma_H/\gamma_C)^{3/2} = 8$ can compensate for the smaller rotor size. In our experience, a MAS regime where $\omega_R \geq 100$ kHz is optimal for RNA studies by ^{1}H-detected ssNMR. Coherence lifetimes of both H1′ and H2′-H5′ protons increase significantly at MAS frequencies above 60 kHz and reach 4.2 and 1.7 ms, respectively, at 109 kHz MAS (Marchanka et al., 2018b). In addition to many dipolar mixing schemes that can be used at MAS ≥ 100 kHz, scalar ^{13}C couplings can be utilized to correlate all carbons in the ribose with each other in a manner typically utilized in solution-state NMR (Hu et al., 1998). J-coupling–based correlation spectroscopy becomes possible due to the long ^{13}C $T_{1\rho}$ relaxation times of 50 ms at ≥ 100 kHz MAS (Marchanka et al., 2018b). In our study, we have implemented a 4D HCCH-TOCSY experiment utilizing low-power 20 kHz WALTZ-16 (Shaka et al., 1983) mixing of a length of 25 ms, which allows us to fully correlate all CH groups in the ribose ring (**Figures 4A,B**). In the 4D HCCH-TOCSY, after t_1 evolution on the starting proton, magnetization is transferred by a short CP to directly attached ^{13}C. After t_2 evolution on this starting carbon, magnetization is transferred to all carbons in the ribose by WALTZ-16 mixing. After evolution during t_3 on the final carbon, magnetization is transferred by a short CP to ^{1}H for detection during t_4 (**Figure 4A**). This experiment has been acquired utilizing non-uniform sampling (NUS) (Paramasivam et al., 2012; Sergeyev et al., 2017) and required 68 h of measurement time. To explore the feasibility of dipolar coupling–based transfer in the ribose at 100 kHz MAS, we have acquired ^{13}C,^{13}C fpRFDR spectra with 8 and 16 ms of mixing time. A full set of intra-ribose correlations has been obtained from both 3D (H)CCH spectra acquired with high sensitivity using uniform sampling within 40 h.

Despite good resolution in ^{1}H-detected 4D HCCH-TOCSY or 3D (H)CCH-fpRFDR spectra, three nucleotides in the helical

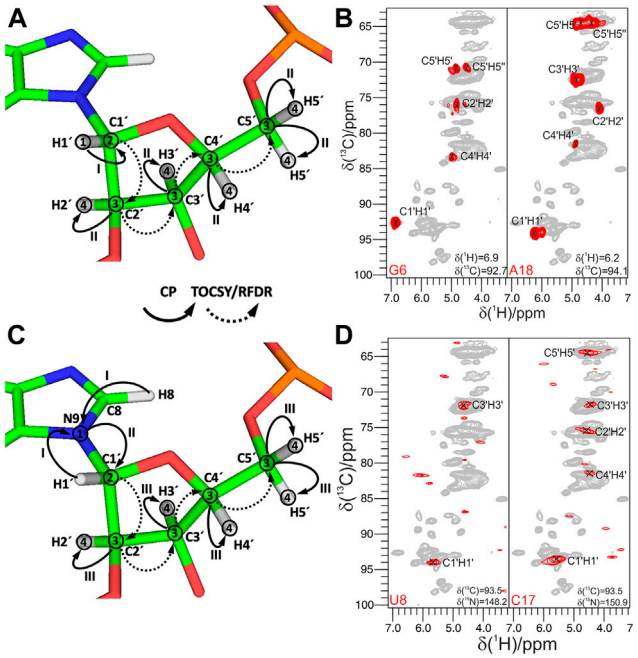


FIGURE 4 | Ribose assignment by ^{1}H-detected ssNMR. **(A,C)** Magnetization transfer for **(A)** 4D HCCH-TOCSY and **(C)** 4D (H)NCCH-TOCSY experiments. Solid arrows indicate CP transfers and dotted arrows indicate homonuclear ^{13}C,^{13}C TOCSY/RFDR recoupling; Arabic numbers indicate spectral dimensions (t_1–t_4); Roman numbers indicate CP transfers. Assignment of ribose spin systems with **(B)** 4D HCCH-TOCSY and **(D)** 4D (H)NCCH-TOCSY experiments. In **(B)** representative 2D planes from the 4D experiment show the spin systems of G6 and A18. ^{1}H and ^{13}C frequencies are indicated in each panel. In **(D)** representative 2D planes from the 4D experiment show the spin systems of U8 and C17. ^{13}C and ^{15}N frequencies are indicated in each panel. For reference, the red contours of either HCCH-TOCSY or 4D (H)NCCH-TOCSY spectra are overlaid onto 2D CP-HSQC spectra (in gray). The HCCH-TOCSY spectrum was recorded on an 800 MHz spectrometer at a MAS frequency of 100 kHz, while the (H) NCCH-TOCSY spectrum was recorded on a 1 GHz spectrometer and at 100 kHz MAS. Figures **(B)** and **(D)** are reproduced from (Marchanka et al., 2018b) with permission from the Royal Society of Chemistry.

regions were not assigned due to low dispersion of the C1′ and H1′ chemical shifts (Marchanka et al., 2018b). Similar to ^{13}C-detected experiments, the N1/N9 dimension can be introduced to resolve spectral crowding at the price of sensitivity. In the 4D (H)NCCH-TOCSY experiment (**Figure 4C**), after a long ^{1}H-^{15}N CP transfer, magnetization is evolved on N1 and N9 during t_1. Subsequently, magnetization is transferred by a specific ^{13}C-^{15}N CP step to the C1′ carbon in the ribose. From here on, the magnetization path follows one of the HCCH-TOCSY experiments and delivers a set of well-resolved N1/N9-C1′-CX′-HX′ correlations after 100 h of measurement time, albeit with low sensitivity (**Figure 4D**).

Ribose–Nucleobase Connection

The next step of the resonance assignment protocol is the connection of riboses to the nucleobases, which can, in principle, be obtained in ssNMR by long-range carbon–carbon correlations utilizing previously discussed homonuclear recoupling schemes, for example, PDSD (Szeverenyi et al., 1982), DARR (Takegoshi et al., 2001), RFDR/fpRFDR (Bennett et al., 1992), PAR (De Paëpe et al., 2008), and others (Ji et al., 2021) (**Figure 5A**). Due to the two-bond distance between ribose C1′ and nucleobase C2/C6 and C4/C8 carbons, dipolar truncation (Bayro et al., 2009) can impair the efficiency of first-order recoupling schemes, so second-order dipolar recoupling (e.g., PDSD, DARR, PAR, and CORD) should be preferred depending on the MAS frequency. While in our studies, we have used the PDSD scheme at MAS < 20 kHz (Marchanka et al., 2013) and have employed fpRFDR at 100 kHz, various recoupling schemes (Ji et al., 2021) can be utilized for this purpose. Their efficacies toward RNA have to be evaluated in future studies.

As discussed previously in the ribose assignment section, spectral resolution in ^{13}C (and ^{1}H) dimensions might not be enough for the spin system assignment using homonuclear recoupling schemes. As for ribose assignment experiments, acquisition of ^{15}N-edited spectra resolves nucleotides by their N1/N9 chemical shifts and facilitates unambiguous ribose–base correlation.

In ^{13}C-detected ssNMR spectroscopy, ribose and base resonances can be connected by the CNC-type experiment (Franks et al., 2005; Schuetz et al., 2010) (**Figure 5B**), which is similar to the HCNCH (Sklenar et al., 1993) experiment in liquids. After initial CP from H1′ to C1′, magnetization is evolved on C1′ during t_1. Next, a long SPECIFIC-CP step (Baldus et al., 1998) transfers magnetization to N1 and N9 in pyrimidines and purines, respectively, where it evolves during t_2. Finally, during a last SPECIFIC-CP step, magnetization is transferred to C2/C6 and C4/C8 in pyrimidines and purines, respectively, for the detection during t_3 (Marchanka et al., 2013). ^{13}C detection at MAS < 20 kHz allows straightforward acquisition of ^{13}C signals of both protonated and non-protonated carbons and provides well-separated C1′-N1-C2/C4 and C1′-N9-C4/C8 chemical shifts for pyrimidines and purines, respectively (**Figure 5C**). Due to the very distinct chemical shifts of N1 and N9 nitrogens, just two differently double nucleotide-type selectively labeled samples are sufficient to obtain such correlations for all nucleotides in RNA. This experiment can be acquired in the ribose-to-base or base-to-ribose direction. In our experience, ribose-to-base transfer was more efficient. Moreover, C1′ → N1/N9 → C2/C4 correlations were more intense than C1′ → N1/N9 → C6/C8 due to a shorter coherence lifetime caused by the increased dipolar relaxation in the latter case. Due to the two low-γ specific CP transfers and the intrinsically low sensitivity of ^{13}C detection, a single good-quality 3D CNC spectrum required more than 120 h of experimental time on a 700 MHz Bruker spectrometer. However, this experiment provides assignment for C4 and C2 carbons in purines and pyrimidines, respectively, whose assigned chemical shifts are very rarely reported in the BMRB database (Ulrich et al., 2007) (s. below). The same experiment can be acquired at a MAS regime of 20 kHz < ω_R < 60 kHz; however, its sensitivity will be compromised due to the necessity of ^{13}C-detection at decreased sample volume (see previous section).

Ribose–nucleobase correlations can be obtained by ^{1}H-detected experiments at MAS > 60 kHz on a single

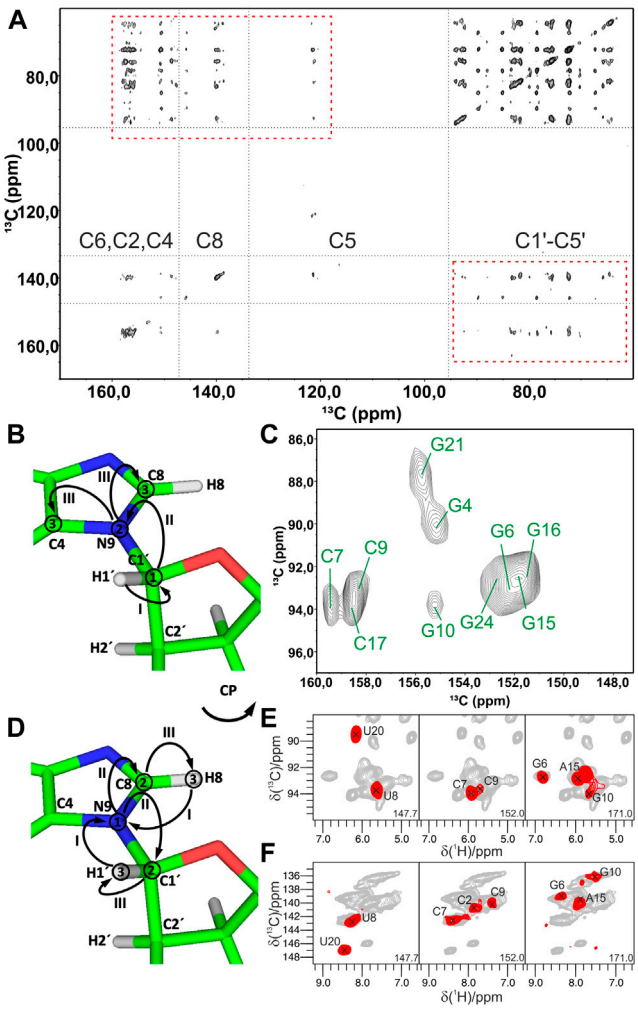


FIGURE 5 | Experiments for the ribose–nucleobase correlations. **(A)** $^{13}C,^{13}C$ PDSD spectrum of the A^{lab}-26mer box C/D RNA measured at 500 ms PDSD mixing time. Ribose–base cross-peaks are highlighted by rectangles. **(B,C)** ^{13}C-detected 3D CNC experiment. **(B)** Magnetization transfer in CNC experiment shown on the example of a purine; **(C)** 2D C_{ribose}–C_{base} projection of the 3D CNC spectrum of G,C^{lab} 26mer box C/D RNA. Spectra in figures (A,C) were acquired on a 700 MHz spectrometer at 13 kHz MAS. **(D–F)** ^{1}H-detected 3D (H)NCH experiment. **(D)** Magnetization transfer shown on the example of a purine. **(E,F)** Representative ^{13}C–^{1}H cross-sections from the (H)NCH spectra show either ribose N1/N9–C1′–H1′ **(E)** or base N1–C6–H6/N9–C8–H8 **(F)** correlations. ^{15}N frequencies are indicated in each panel. (H)NCH spectra were recorded on a 1 GHz spectrometer at 111 kHz MAS. In magnetization transfer schemes, solid arrows indicate CP transfers; Arabic numbers correspond to the spectral dimensions (t_1–t_3); Roman numbers indicate CP transfers. Figures **(A,C)** are adapted with permission from (Marchanka et al., 2013), © John Wiley and Sons, 2013. Figures **(E,F)** are reproduced from (Marchanka et al., 2018b) with permission from the Royal Society of Chemistry.

uniformly labeled $^{13}C,^{15}N$ RNA sample (Marchanka et al., 2018b). In the ^{1}H-detected (H)NCH experiment, after long ^{1}H-^{15}N CP transfer and evolution on ^{15}N during t_1, either ribose- or base-tuned band-selective CP transfers the magnetization either to C1′ or to C6/C8 for the evolution during t_2. A final short CP transfers magnetization to directly bound H1′ or H6/H8 protons for the detection during t_3 (**Figure 5D**). This experiment provides a set of either ribose-specific N1/N9–C1′–H1′ or base-specific N1–C6–H6/N9–C8–H8 correlations in pyrimidines/purines, respectively (**Figures 5E,F**). Due to the high sensitivity of ^{1}H detection, good quality ribose- and base-specific spectra have been obtained in 12 and 19 h, respectively (Marchanka et al., 2018b). While the (H)NCH experiment provides unambiguous correlation of ribose and base CH groups with their corresponding N1/N9 nitrogens, ribose–base correlations obtained by these two experiments might be ambiguous, since they share only N1/N9 chemical shifts. The narrow chemical shift range of N1/N9 nitrogens is exacerbated by intrinsically worse linewidths in ssNMR compared to solution-state. To lift the ambiguity, an additional ^{1}H evolution period can be added before the first ^{1}H-^{15}N CP step, yielding the 4D HNCH experiment. In such a spectrum, chemical shifts in the indirect proton dimension (H6 or H8) and N1/N9 represent the nucleobase, while N1/N9, C1′ and H1′ chemical shifts correspond to the ribose.

Nucleobase Assignment

After resonances of riboses are assigned and correlated with the nucleobase C6-H6/C8-H8 groups, assignment of resonances in the nucleobase is performed. This is not a trivial task, particularly in ^{1}H-detected NMR spectroscopy, due to the low density of protons in RNA nucleobases. The majority of reported assigned carbons and nitrogen RNA chemical shifts in the BMRB database belong to either protonated nuclei or nuclei directly attached to the protonated carbons. The fraction of assigned chemical shifts for C4 and C5 carbons in purines is very low and estimates to only 1.8% (72/3,921) and 2.0% (79/3,921), respectively, of all assigned base carbon chemical shifts in the BMRB database. Sensitive dipolar coupling–based transfer in ssNMR provides an unprecedented opportunity for the assignment of these otherwise not-easy-to-access carbons.

C2/C6 and C4/C8 carbons in pyrimidines and purines, respectively, are assigned by the 3D CNC experiment acquired at MAS < 20 kHz, as described in the section above. H6-C6 and H8-C8 groups are assigned by the (H)NCH experiment at ultrafast MAS.

Strategies for the assignment of the remaining carbons and nitrogens in the nucleobase depend strongly on if either ^{13}C- or ^{1}H-detection is utilized.

At MAS frequencies < 20 kHz, using ^{13}C-detection, $^{13}C,^{13}C$ recoupling schemes discussed above for the ribose assignment can be used to obtain carbon assignment in the nucleobase. 500 ms PDSD recoupling (**Figure 5A**) or 8–16 ms PAR (**Figure 6A**) were found to be sufficient to obtain many, although overlapping, correlations in the nucleobases of RNA selectively labeled by nucleotide type. As for ribose assignment, inclusion of the nitrogen dimension can reduce spectral crowding. The previously described $^{13}C,^{15}N$ TEDOR-$^{13}C,^{13}C$ PDSD experiment (Riedel et al., 2005b; Daviso et al., 2013) can provide assignment of carbon resonances in nucleobases by connecting their chemical shifts to the chemical shifts of assigned carbons in the nucleobase (C6/C8) and the ribose (C1′) (**Figure 6B**).

The NCC experiment (Pauli et al., 2001; Igumenova et al., 2004; Franks et al., 2005) is employed to connect amino and imino nitrogens with nucleobase and (partially) ribose carbons (Marchanka et al., 2013). After an initial short CP from ^{1}H to ^{15}N, ^{15}N chemical shifts evolve during t_1 and yield the frequencies of N6 (A), N4 (C), N1, N2 (G), and N3 (U). Following this, the magnetization is transferred by SPECIFIC-CP to the nearby carbons, N6→C6 (A), N4→C4 (C), N1→C2/C6 and N2→C2 (G), and N3→C2/C4 (U). Here, magnetization can be optionally evolved during t_2 or spread directly by PDSD to all base carbons and C1′ for the detection during t_2 or t_3 (in the 3D version) (**Figures 6C,D**). Needless to say, diverse ^{13}C,^{13}C recoupling schemes can be utilized instead of PDSD. The NCC experiment can, in theory, provide unambiguous assignment for all protonated nitrogens and all carbons in the nucleobase if protonated nitrogens can be uniquely correlated with well-resolved carbons in the nucleobase (C2/C6, C4/C8) or C1′. However, the poor chemical shift dispersion of ^{15}N amino and imino resonances usually renders assignment inconclusive, especially in helical regions.

Assignment of the remaining non-protonated nitrogens in the base is important for characterization of non-WC base pairs, especially G:A and U:U. Assignment of N7 nitrogen in adenosines and guanosines can be readily obtained from the ^{13}C,^{15}N TEDOR (Jaroniec et al., 2002) experiment that provides (H8)C8-N7 correlations. Assignment of N1 and N3 nitrogens in adenosines is obtained from the same-type TEDOR experiment by acquisition of (H2)C2-N1/N3 correlations (**Figure 6E**). To obtain site-specific assignment of these N1/N3 nitrogens, C2 carbons have to be correlated with C6/C8 carbons by any of the ^{13}C,^{13}C correlation schemes discussed above.

The assignment of non-protonated nitrogens (N3) in both cytidines and guanosines is the most challenging task, due to the absence of any protons in their close vicinity. It could, in principle, be obtained from a modified NCC experiment, where a long initial ^{1}H-^{15}N CP transfers the magnetization from the distant (r = 2.4 Å) amino protons H41/H42 (C) or H21/H22 (G) to the N3 nitrogen. ^{15}N chemical shifts evolve during t_1, and then magnetization is transferred to the C2 and C4 carbons, where it is recorded during t_2 (**Figure 6F**). It is uncertain if 1) $T_{1\rho}$ during the H41/H42→N3 and H21/H22→N3 transfer is long enough to allow for the efficient long-range magnetization transfer and 2) resolution of ^{15}N and ^{13}C resonances are good enough to allow for the spin system–specific assignment. However, N3 of cytidine can be easily assigned indirectly through base-paired guanosine in the WC G:C base pair. Here, the cross-strand N–N distance of ~2.0 Å allows the straightforward connection of N3 in cytidine with N1 in guanosine using, for example, ^{15}N,^{15}N RFDR or ^{15}N,^{15}N PAR correlations (**Figure 2B**).

While ^{13}C-detected ssNMR allows almost complete assignment of carbons and nitrogens in the nucleobases using only two different types of experiments, namely, CNC and NCC, ^{1}H-detected ssNMR on RNA nucleobases is more challenging. Despite the significant advantages of 1) additional spectral dimensions and 2) higher sensitivity, ^{1}H detection obviously requires that the magnetization transfer pathway ends on a proton. As is known from solution-state NMR, many different types of experiments are necessary to assign RNA nucleobases (Fürtig et al., 2003).

Very recently, we have addressed this challenge and have published the set of ssNMR experiments that allow assignment of nucleobase resonances (Aguion et al., 2021).

In the first step, all nucleobase carbons are correlated with the previously assigned C6-H6 (pyrimidines) or C8-H8 (purines) groups using ^{13}C,^{13}C fpRFDR recoupling. This step is accomplished by the 3D (H)CCH experiment, which starts with a long ^{1}H-^{13}C CP step to gain sufficient ^{13}C magnetization on both protonated and non-protonated carbons. After t_1 evolution on ^{13}C, a phase-cycled selective inversion pulse cancels the signals of the ribose ring. Following this, 8 ms-long ^{13}C,^{13}C fpRFDR recoupling transfers magnetization to nearby carbons, whose chemical shifts are recorded during t_2. Finally, the ^{13}C magnetization is transferred by a short read-out CP to directly attached protons for the detection during t_3 (**Figure 7A**). This experiment correlates all nucleobase carbons with the protonated C5-H5 and C6-H6 groups in pyrimidines, C8-H8 groups in purines, and C2-H2 groups in adenosines (**Figure 7B**).

In the second step, ^{15}N-^{1}H imino and amino resonances are correlated with the assigned carbons and C-H groups using two different experiments. The 3D (H)CNH experiment (**Figure 7C**) starts with a long ^{1}H-^{13}C CP step to transfer proton magnetization to non-protonated carbons. After t_1 evolution on ^{13}C, magnetization is transferred to directly attached amino or imino nitrogens *via* a ^{13}C,^{15}N CP step. After t_2 evolution on nitrogens, a phase-cycled selective inversion pulse selects either amino or imino magnetization that is subsequently transferred by a short read-out CP to ^{1}H for the detection during t_3. This experiment yields C6-H6-H61/H62 correlations for A, C4-H4-H41/H42 for C, C2/C6-N1-H1 and C2-N2-H21/H22 for G, and C2/C4-N3-H3 for U (**Figure 7D**). The correlation of C2/C4 resonances in uridines and C2/C6 resonances in guanosines with imino groups is sufficient for site-specific assignment of imino resonances if chemical shift dispersion exists in at least one of the correlated carbons. Amino groups, on the other hand, have only one adjacent carbon (C6 in A, C4 in C, and C2 in G) with a very narrow chemical shift range, which often prevents unambiguous assignment of amino resonances. This is aggravated by the poor resolution of amino resonances compared to imino resonances. To resolve the ambiguity in these cases, we have combined the two experiments reported above to develop 3D 1) (H)N(C)CH and 2) H(NC)CH experiments. In experiment 1) (**Figure 7E**), after an initial short ^{1}H-^{15}N CP step, the chemical shifts of amino and imino nitrogens are recorded during t_1. Subsequently, magnetization is transferred to directly attached carbons by a ^{15}N, ^{13}C CP step. The phase-cycled selective inversion pulse selects nucleobase carbon magnetization at ~ 160 ppm, which is then spread to all nearby carbons by a 14 ms RDFR mixing step. Following this, ^{13}C magnetization is evolved during t_2 and finally transferred from protonated carbons by a short read-out CP to directly attached protons for the detection during t_3 (**Figure 7E**). Despite the low sensitivity of this experiment due to the modest efficiency of both low-γ ^{13}C,^{15}N CP and ^{13}C,^{13}C RFDR transfers, it provides important information by correlating imino and amino nitrogens directly with the well-resolved C6-H6 and C5-H5 groups in pyrimidines, C8-H8 groups in pyrimidines, and C2-H2 groups in adenosines (**Figure 7F**). In

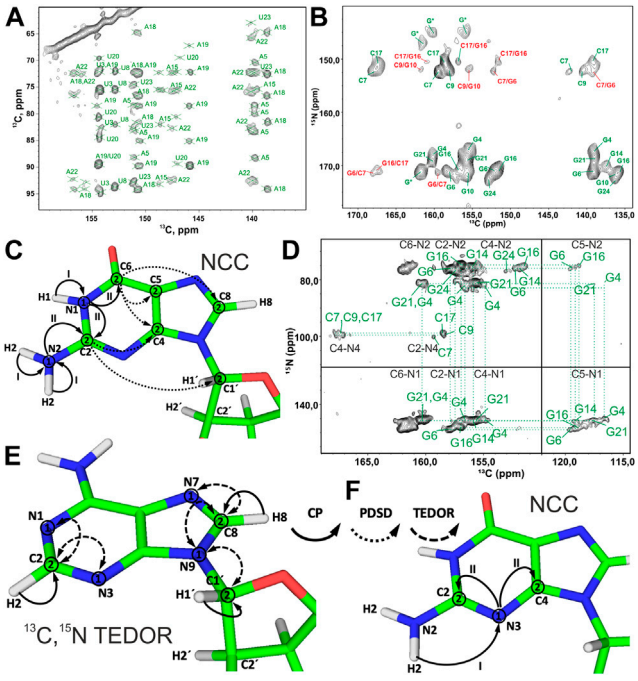


FIGURE 6 | Nucleobase assignment by ^{13}C-detected ssNMR. **(A)** Zoom in on the ribose–base region of the ^{13}C,^{13}C PAR spectrum of the A,U^{lab}-26mer box C/D RNA acquired at 10 ms PAR mixing time. **(B)** Zoom in on the base region of the ^{13}C,^{15}N-TEDOR-^{13}C,^{13}C PDSD spectrum of the G,C^{lab} 26mer box C/D RNA. Intra- and inter-nucleotide correlations are labeled in green and red, respectively. **(C,D)** NCC experiment for the assignment of nitrogen and carbon resonances in the nucleobases. **(C)** Magnetization transfer in NCC experiment shown on the example of a guanosine. **(D)** 2D NCC spectrum of G,C^{lab} 26mer box C/D RNA acquired at 150 ms PDSD mixing time. Magnetization transfer in **(E)** ^{13}C,^{15}N TEDOR experiment for the assignment of N7 nitrogens in purines and N1/N3 nitrogens in adenosines and **(F)** in the modified NCC experiment for the assignment of N3 nitrogens in cytidines and guanosines. In magnetization transfer schemes, solid arrows indicate CP transfers, dotted arrows indicate PDSD transfers, and dashed arrows indicate TEDOR transfers; Arabic numbers correspond to spectral dimensions (t_1–t_2); Roman numbers indicate CP transfers. Spectra in figures **(A,B)** and **(D)** were acquired on a 700 MHz spectrometer at 16 kHz MAS and 13 kHz MAS, respectively. The spectrum in **(B)** is adapted from (Marchanka et al., 2015). The spectrum in **(D)** is adapted with permission from (Marchanka et al., 2013) © John Wiley and Sons, 2013.

experiment 2), chemical shifts of amino and imino protons are evolved instead of nitrogens and should provide better spin system separation due to better resolution in the proton dimension. Unfortunately, its sensitivity was poor due to the short coherence lifetimes of imino (~4 ms) and especially amino (~1 ms) protons at 100 kHz MAS. Relaxation behavior of RNA amino/imino groups could be improved by 1) increasing the MAS frequency and 2) performing experiments in partial deuterated buffer (~25–50%), whereas the overall sensitivity of the experiment can be improved by utilizing optimized ^{13}C,^{13}C recoupling schemes.

While ^{13}C,^{13}C recoupling-based experiments were successfully utilized for the assignment of most nucleobase resonances, some nuclei can be more efficiently assigned through ^{1}H,^{1}H recoupling schemes. Thus, amino groups of guanosines were assigned through correlation with well-resolved imino groups through the 3D (H)N(HH)NH experiment described above (**Figures 7G,H**). Similarly, amino groups of cytidines were assigned through the previously described 3D (H)N(HH)CH experiment (**Figure 7I**). This experiment provides N4–C5–H5 correlations for all four cytidines in the structured region of 26mer box C/D RNA (**Figure 7J**) and, moreover, provides assignment of WC A:U (**Figures 2G,H**) and non-WC G:A base pairs.

In the third step, assignment of non-protonated nitrogen resonances is performed. In ^{1}H-detected ssNMR, it can be obtained from a modified version of the 3D (H)NCH experiment described above (**Figure 7K**). For effective excitation of N1, N3, and N7 resonances, the ^{15}N carrier frequency should be shifted to a higher ppm (e.g., ~190 ppm). Moreover, a selective ^{13}C refocusing pulse applied after t_2 eliminates unwanted signals of the ribose ring. The experiment correlates N7 resonances with protonated C8-H8 groups in purines and N1 and N3 resonances with protonated C2-H2 groups in adenosines. N3 resonances in guanosines, however, do not have any adjacent CH group and can therefore not be assigned in this experiment. The same applies to N3 resonances in cytidines. As in ^{13}C-detected ssNMR, assignment of these resonances remains challenging, but N3 resonances of cytidines can likewise be obtained indirectly through base-paired guanosines. Here, the close distance of cytidine N3 and H1 resonances in the base-paired guanosine (r = 2.0 Å) can be exploited in a simple 2D ^{1}H-^{15}N-HSQC experiment with long-range ^{1}H-^{15}N/^{15}N-^{1}H CP transfer times of 8 ms (Aguion et al., 2021).

At intermediate MAS rates, nucleobase assignment can be performed either by acquisition of ^{13}C-detected NC-type experiments (Marchanka et al., 2013; Zhao et al., 2019), taking into account attenuated ^{13}C sensitivity, or using ^{1}H detection of amino/imino protons employing 3D (H)CNH and 3D (H)N(HH)NH experiments. ^{1}H-detection of nucleobase H6/H8/H2 protons would be unfavorable due to their short coherence lifetimes at MAS < 60 kHz (Marchanka et al., 2018b).

Sequential Assignment

After spin system–specific assignment is achieved, nucleotides are sequentially connected with each other to obtain site-specific assignments. In solution-state NMR, this task is accomplished by measurement of either through-space internucleotide H6/H8(i)-H2′,H1′(i–1) NOEs or through detection of overlapped C_{ribose}–P correlations (Fürtig et al., 2003). A similar approach has been utilized in ssNMR; however, due to the preservation of direct dipolar couplings, accessible distances are significantly longer and are not limited to ^{1}H,^{1}H restraints. The first ever sequential RNA correlations by ssNMR were obtained by the Görlach group. They have utilized a CHHC-type experiment (Lange et al., 2002) and have measured sequential H2′,H3′(i-1)-H6/H8(i) contacts in the $(CUG)_{97}$ repeat (**Figures 8A,B**) (Riedel et al., 2006). In our ^{13}C- and ^{15}N-detected study on 26mer box C/D RNA, we have utilized ^{13}C,^{13}C and ^{13}C,^{31}P correlations to obtain sequential assignments (Marchanka et al., 2015). Due to severe resonance overlap, eight different double nucleotide-type selective labeled samples were prepared and an ^{15}N-editing TEDOR step was coupled to the long-range ^{13}C,^{13}C PDSD recoupling (**Figure 8C**). Long-range correlations up to 9 Å were obtained at 700 ms of PDSD mixing

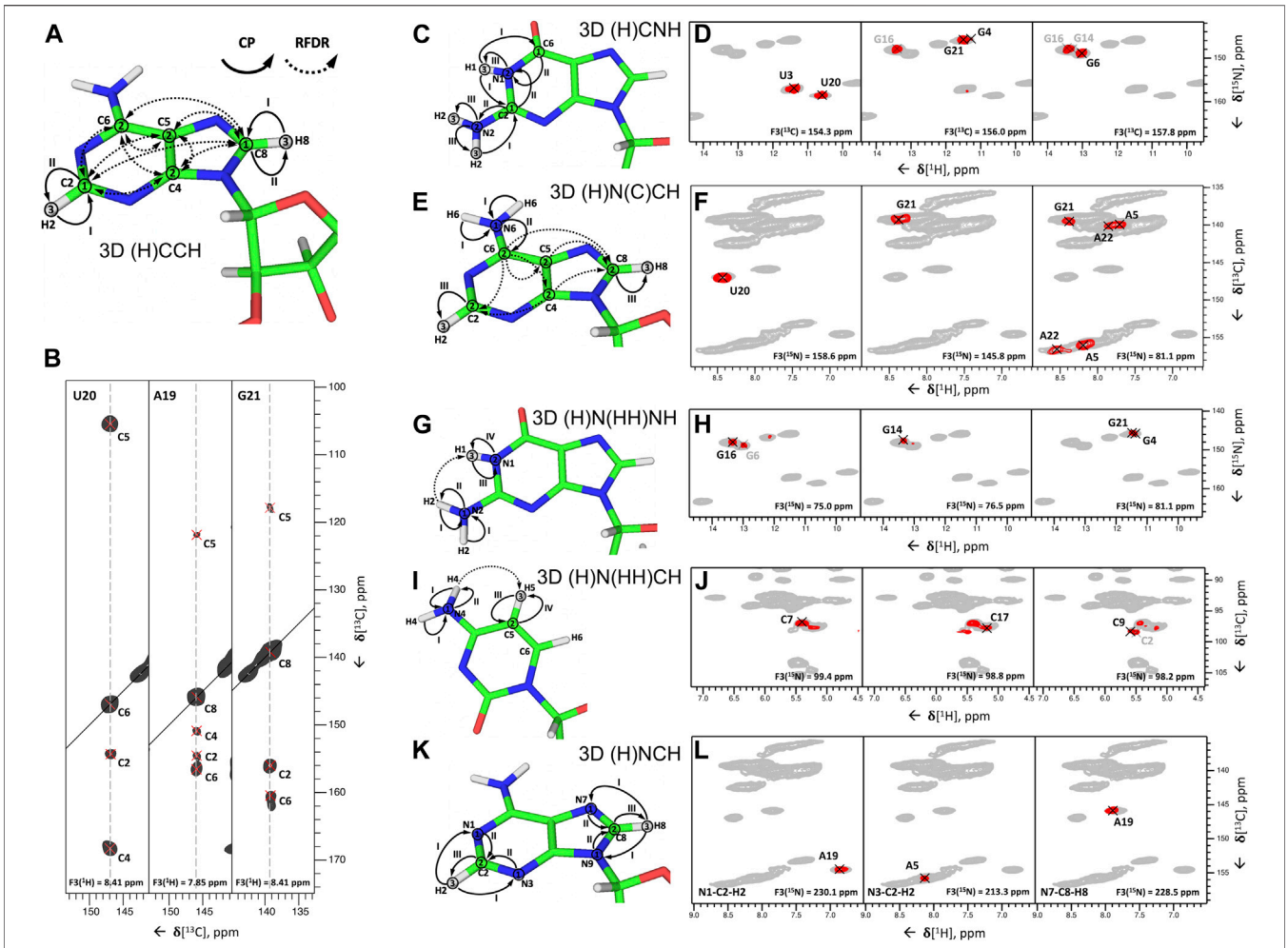


FIGURE 7 | Nucleobase assignment by ^{1}H-detected ssNMR. **(A,C,E,G,I,K)** Magnetization transfers for **(A)** 3D (H)CCH, **(C)** 3D (H)CNH, **(E)** 3D (H)N(C)CH, **(G)** 3D (H)N(HH)NH, **(I)** 3D (H)N(HH)CH experiments, and **(K)** 3D(H)NCH experiments, shown on the example of adenosine **(A,E,K)**, guanosine **(C,G),** and cytidine **(I)**. Solid arrows indicate CP transfers and dotted arrows indicated RFDR transfers; Arabic numbers indicate spectral dimensions (t_1–t_3); Roman numbers indicate CP transfers. **(B,D,F,H,J,L)** Assignment of base spin systems in the 26mer Box C/D RNA with **(B)** 3D (H)CCH, **(D)** 3D (H)CNH, **(F)** 3D (H)N(C)CH, **(H)** 3D (H)N(HH)NH, **(J)** 3D (H)N(HH)CH, and **(L)** (H)NCH experiments. In **(B)** representative 2D ^{13}C–^{13}C planes extracted from the 3D (H)CCH spectrum show correlations of all intra-nucleotide nucleobase carbons for the nucleotides U20, A19, and G21. In **(D)** representative 2D planes extracted from the 3D (H)CNH experiment show correlations of imino nitrogens with distinct nucleobase carbon resonances for the nucleotides U3, U20, G4, G16, G21, G6, and G14. In **(F)** representative 2D planes extracted from the 3D (H)N(C)CH experiment show correlations of imino and amino nitrogens with protonated nucleobase carbons for the nucleotides U20, G21, A5, and A22. In **(H)** representative 2D planes extracted from the 3D (H)N(HH)NH experiment show correlations of amino and imino nitrogens of guanosines for the nucleotides G16, G6, G14, G21, and G4. In **(J)** representative 2D planes extracted from the 3D (H)N(HH)CH experiment show correlations of amino nitrogens with protonated carbons of cytidines for the nucleotides C7, C17, C9, and C2. In **(L)** representative 2D planes extracted from the 3D (H)NCH experiment show correlations of non-protonated nitrogens with protonated carbons for the nucleotides A19 and A5. ^{13}C or ^{15}N frequencies are indicated in each panel. Peaks labeled in gray indicate that the peak maximum is not in the plane shown here. For reference in panels **(D,F,H,J,L)**, the red contours of 3D H(C)NH and (H)N(HH)NH spectra **(D,H)** are overlaid onto 2D ^{1}H-^{15}N CP-HSQC spectra (in gray), while red contours of 3D (H)N(C)CH, (H)N(HH)CH and (H)NCH spectra are overlaid onto 2D ^{1}H-^{13}C CP-HSQC spectra (in gray), tailored either for the C2-H2/C6-H6/C8-H8 **(F,L)** or the ribose/C5-H5 spectral regions **(J)**. All spectra were recorded on an 850 MHz spectrometer at a MAS frequency of 100 kHz (Aguion et al., 2021). All spectra are adapted from (Aguion et al., 2021).

time. While ^{13}C,^{31}P correlations acquired by the ^{13}C,^{31}P TEDOR experiment (**Figure 8D**) provided important sequential contacts in the non-canonical region of RNA (kink-turn) and corroborated sequential assignments obtained by ^{13}C,^{13}C correlation experiments, their value for correlating nucleotides in the helical regions was limited due to strong resonance overlap. Additional ^{1}H,^{1}H sequential contacts have been obtained from CHHC and NHHC experiments acquired at 200 us of mixing time (**Figure 8E**). ^{13}C,^{31}P and CHHC/NHHC correlation experiments were essential to distinguish sequential contacts from the long-distance inter-strand correlations.

While ^{1}H-detected ssNMR allows spin system–specific assignment of both riboses and nucleobases from single, uniformly ^{13}C,^{15}N labeled RNA samples (Marchanka et al., 2018b; Aguion et al., 2021), no

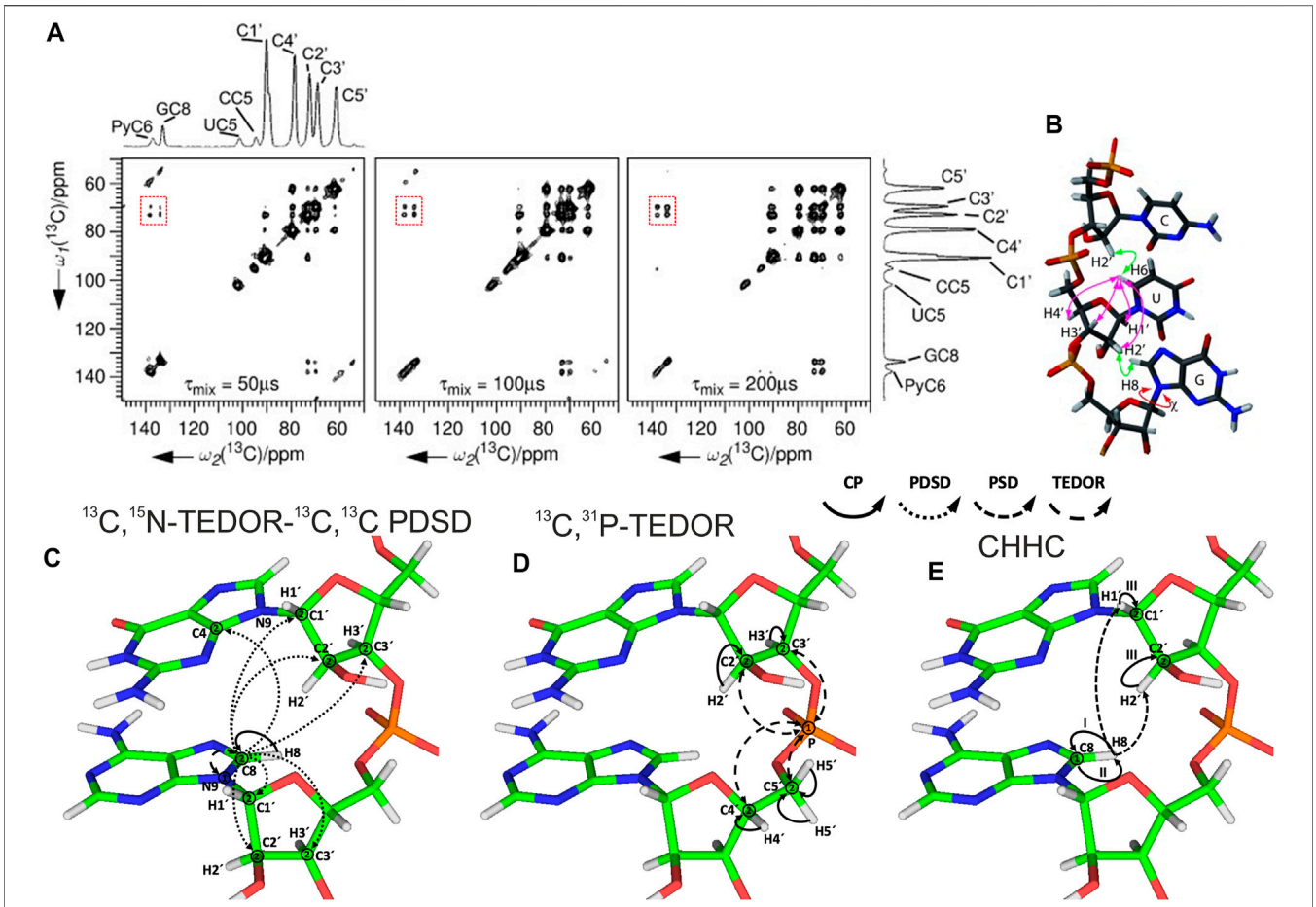


FIGURE 8 | Sequential assignment of RNA by ssNMR. **(A)** CHHC spectra of $(CUG)_{97}$ RNA acquired at different mixing times. The assignments of the different resonances are indicated on the spectral projection shown. Sequential correlations are highlighted. **(B)** Intra- and internucleotide ^1H–^1H correlations in a double-stranded A-form helical RNA (Riedel et al., 2006). **(C–E)** Magnetization transfer in **(C)** $^{13}C,^{15}N$-TEDOR-$^{13}C,^{13}C$ PDSD, **(D)** $^{13}C,^{31}P$ TEDOR, and **(E)** CHHC experiments shown on the example of sequential guanosine and adenosine (Marchanka et al., 2015). Solid arrows indicate CP transfers, dotted arrows indicate homonuclear $^{13}C,^{13}C$ PDSD transfer, densely dashed arrows indicate PSD transfer, and dashed arrows indicated TEDOR transfer; Arabic numbers correspond to the spectral dimensions (t_1–t_2); Roman numbers indicate CP transfers. Panels **(A)** and **(B)** are reprinted with permission from (Riedel et al., 2006) © John Wiley and Sons, 2006.

sequence-specific assignment of RNA by ^1H-detected ssNMR has been reported to date. This objective can be achieved by utilizing, for example, $^1H,^1H$ recoupling coupled with ^{13}C-editing either on one (3D) or both sides (4D) to overcome resonance overlap. Furthermore, ^{13}C-^{31}P and ^1H-^{31}P correlation experiments may be beneficial at fast MAS due to effective averaging of CSA. Our group is currently working in this direction, and we hope to present procedures for the sequential assignment and measurement of structural restraints based on ^1H-detected ssNMR in the near future.

DISCUSSION AND OUTLOOK

Feasibility of ssNMR approaches for RNA studies largely depends on the quality of sample preparation. While nano/microcrystalline crystallization of rigid RNAs or RNA in RNP complexes (Marchanka et al., 2013) and EtOH precipitation of rigid well-folded RNAs (Zhao et al., 2019) yield well-dispersed spectra, these methods might be less suitable for small RNAs without a well-defined structure, as they provide spectra with narrow chemical shift dispersion (Huang et al., 2012; Yang et al., 2017). This can be an indication of partial loss of the tertiary structure upon PEG or EtOH precipitation, which may limit the advantages of ssNMR methods. Complementary methods of sample preparation, for example, sedimentation of the dissolved sample directly into the ssNMR rotor by ultracentrifugation (Bertini et al., 2011; Wiegand et al., 2020), should be evaluated for their feasibility toward RNA.

Access to narrow RNA resonances makes ssNMR assignment feasible both by $^{13}C/^{15}N$-detected and ^1H-detected experiments. While conventional ^{13}C- and ^{15}N-detected studies at MAS < 20 kHz require preparation of many different samples, ^1H-detected ssNMR at MAS ≥ 100 kHz allows straightforward assignment of small (< 30 nt) RNA using a single uniformly $^{13}C,^{15}N$-labeled sample. We find the intermediate MAS regime of 20 kHz < ω_R < 60 kHz suboptimal for RNA studies due to the decreased sensitivity from using ^{13}C-detection and the broad ^1H lines caused by the short coherence lifetime from using ^1H-detection. Coherence lifetimes of RNA protons increase nearly linearly in the

20–110 kHz MAS range (Marchanka et al., 2018b), and recent protein studies indicate that such a linear regime continues up to at least 150 kHz MAS (Schledorn et al., 2020). ^{1}H-detected ssNMR studies at MAS frequencies beyond 110 kHz will further improve spectral quality due to better effective averaging of dipolar ^{1}H–^{1}H interactions and improved homogeneous linewidth.

Most of the reported ssNMR RNA studies were performed on protonated, nucleotide-type selective or uniformly ^{13}C,^{15}N labeled RNA. High hydrogen density in the RNA ribose leads to crowded H2′-H5′/H5″ resonances, which are additionally broadened by strong ^{1}H–^{1}H dipolar couplings. As protein deuteration is beneficial for ^{1}H-detected ssNMR studies (Andreas et al., 2015), particularly at MAS < 100 kHz (Cala-De Paepe et al., 2017), selective ribose deuteration as implemented by the Williamson group (Tolbert and Williamson, 1996; Tolbert and Williamson, 1997) can be advantageous for the spectral quality of RNA ribose resonances in ssNMR due to reduction of their spectral overlap and dilution of the ^{1}H–^{1}H couplings network. Furthermore, sparse labeling of ribose carbons can remove spectral overlap in C2′/C3′ carbons and reduce the dense network of ^{13}C,^{13}C dipolar- and J-couplings, consequently improving linewidths (Tolbert and Williamson, 1997; Davis et al., 2005) and eliminating dipolar truncation (Bayro et al., 2009). In addition to selective labeling of ribose, specific ^{13}C and ^{2}H labeling of nucleobases can be beneficial for ssNMR studies. Selective nucleobase deuteration can reduce spectral crowding and improve linewidth of H6 pyrimidine resonances by specific deuteration of H5 protons. Furthermore, atom-specific ^{13}C/^{15}N labeling of nucleobases, though tedious to synthesize, can remove resonance overlap and facilitate sequential assignment and acquisition of long-range distance restraints (Wunderlich et al., 2012; Juen et al., 2016). Increasing commercial availability of phosphoramidites with different labeling patterns together with the availability of RNA chemical synthesis machinery will allow wide usage of ssNMR RNA studies on atom-specific labeled RNA in the near future.

So far, the presented studies were mostly limited to small RNAs < 30 nt in length with only a few examples of larger RNA, for example, 72 nt long guide RNA in box C/D complex (Marchanka et al., 2018a) or 71 nt long riboA71 (Zhao et al., 2019). While ssNMR can, in principle, be applied to RNA of any size, ssNMR spectra of longer RNA will ultimately have extreme spectral crowding even if nucleotide-type selective labeling is utilized. For such RNAs, nucleotide-type selective labeling should be coupled with segmental labeling (Tzakos et al., 2007; Nelissen et al., 2008; Duss et al., 2010), which will reduce spectral crowding and make complete assignment and structural studies of RNA feasible. This strategy can report on short isotope-labeled RNA stretches exclusively in large RNA or protein–RNA complexes, thereby providing valuable structural information that may be out of reach for other structural biology techniques.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

PIA was supported by the Deutsche Forschungsgemeinschaft (DFG grant CA294/21-1). AM was supported by the Deutsche Forschungsgemeinschaft (DFG grant MA5157/3-1).

ACKNOWLEDGMENTS

We thank Dr. Megha Karanth for the critical reading of the manuscript.

REFERENCES

Ahmed, M., Marchanka, A., and Carlomagno, T. (2020). Structure of a Protein-RNA Complex by Solid-State NMR Spectroscopy. *Angew. Chem. Int. Ed.* 59, 6866–6873. doi:10.1002/anie.201915465

Aguion, P. I., Kirkpatrick, J., Carlomagno, T., and Marchanka, A. (2021). Identification of RNA Base Pairs and Complete Assignment of Nucleobase Resonances by 1H-Detected Solid-State NMR Spectroscopy at 100 kHz MAS. *Angew. Chem. Int. Ed.* doi:10.1002/anie.202107263

Andreas, L. B., Jaudzems, K., Stanek, J., Lalli, D., Bertarello, A., Le Marchand, T., et al. (2016). Structure of Fully Protonated Proteins by Proton-Detected Magic-Angle Spinning NMR. *Proc. Natl. Acad. Sci. USA.* 113, 9187–9192. doi:10.1073/pnas.1602248113

Andreas, L. B., Le Marchand, T., Jaudzems, K., and Pintacuda, G. (2015). High-Resolution Proton-Detected NMR of Proteins at Very Fast MAS. *J. Magn. Reson.* 253, 36–49. doi:10.1016/j.jmr.2015.01.003

Asami, S., Rakwalska-Bange, M., Carlomagno, T., and Reif, B. (2013). Protein-RNA Interfaces Probed by1H-Detected MAS Solid-State NMR Spectroscopy. *Angew. Chem. Int. Ed.* 52, 2345–2349. doi:10.1002/anie.201208024

Atreya, H. S. (2012). *Isotope Labeling in Biomolecular NMR*. Dordrecht, New York: Springer.

Baldus, M., Petkova, A. T., Herzfeld, J., and Griffin, R. G. (1998). Cross Polarization in the Tilted Frame: Assignment and Spectral Simplification in Heteronuclear Spin Systems. *Mol. Phys.* 95, 1197–1207. doi:10.1080/00268979809483251

Barnwal, R. P., Yang, F., and Varani, G. (2017). Applications of NMR to Structure Determination of RNAs Large and Small. *Arch. Biochem. Biophys.* 628, 42–56. doi:10.1016/j.abb.2017.06.003

Bayro, M. J., Huber, M., Ramachandran, R., Davenport, T. C., Meier, B. H., Ernst, M., et al. (2009). Dipolar Truncation in Magic-Angle Spinning NMR Recoupling Experiments. *J. Chem. Phys.* 130, 114506. doi:10.1063/1.3089370

Beaucage, S. L., and Reese, C. B. (2009). Recent Advances in the Chemical Synthesis of RNA. *Curr. Protoc. Nucleic Acid Chem.* 16, 2. doi:10.1002/0471142700.nc0216s38

Bennett, A. E., Griffin, R. G., Ok, J. H., and Vega, S. (1992). Chemical Shift Correlation Spectroscopy in Rotating Solids: Radio Frequency-Driven Dipolar Recoupling and Longitudinal Exchange. *J. Chem. Phys.* 96, 8624–8627. doi:10.1063/1.462267

Bertini, I., Bhaumik, A., De Paëpe, G., Griffin, R. G., Lelli, M., Lewandowski, J. R., et al. (2010). High-Resolution Solid-State NMR Structure of a 17.6 kDa Protein. *J. Am. Chem. Soc.* 132, 1032–1040. doi:10.1021/ja906426p

Bertini, I., Luchinat, C., Parigi, G., Ravera, E., Reif, B., and Turano, P. (2011). Solid-State NMR of Proteins Sedimented by Ultracentrifugation. *Proc. Natl. Acad. Sci.* 108, 10396–10399. doi:10.1073/pnas.1103854108

Breeze, A. L. (2000). Isotope-Filtered NMR Methods for the Study of Biomolecular Structure and Interactions. *Prog. Nucl. Magn. Reson. Spectrosc.* 36, 323–372. doi:10.1016/S0079-6565(00)00020-0

Brown, J. D., Kharytonchyk, S., Chaudry, I., Iyer, A. S., Carter, H., Becker, G., et al. (2020). Structural Basis for Transcriptional Start Site Control of HIV-1 RNA Fate. *Science* 368, 413–417. doi:10.1126/science.aaz7959

Cala-De Paepe, D., Stanek, J., Jaudzems, K., Tars, K., Andreas, L. B., and Pintacuda, G. (2017). Is Protein Deuteration Beneficial for Proton Detected Solid-State NMR at and Above 100 kHz Magic-Angle Spinning?. *Solid State. Nucl. Magn. Reson.* 87, 126–136. doi:10.1016/j.ssnmr.2017.07.004

Callon, M., Malär, A. A., Pfister, S., Římal, V., Weber, M. E., Wiegand, T., et al. (2021). Biomolecular Solid-State NMR Spectroscopy at 1200 MHz: the Gain in Resolution. *J. Biomol. NMR.* 75, 255, 272. doi:10.1007/s10858-021-00373-x

Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K., and Oschkinat, H. (2002). Structure of a Protein Determined by Solid-State Magic-Angle-Spinning NMR Spectroscopy. *Nature* 420, 98–102. doi:10.1038/nature01070

Davis, J. H., Tonelli, M., Scott, L. G., Jaeger, L., Williamson, J. R., and Butcher, S. E. (2005). RNA Helical Packing in Solution: NMR Structure of a 30kDa GAAA Tetraloop-Receptor Complex. *J. Mol. Biol.* 351, 371–382. doi:10.1016/j.jmb.2005.05.069

Daviso, E., Eddy, M. T., Andreas, L. B., Griffin, R. G., and Herzfeld, J. (2013). Efficient Resonance Assignment of Proteins in MAS NMR by Simultaneous Intra- and Inter-Residue 3D Correlation Spectroscopy. *J. Biomol. NMR.* 55, 257–265. doi:10.1007/s10858-013-9707-0

De Paëpe, G., Lewandowski, J. R., Loquet, A., Böckmann, A., and Griffin, R. G. (2008). Proton Assisted Recoupling and Protein Structure Determination. *J. Chem. Phys.* 129, 245101. doi:10.1063/1.3036928

Dingley, A. J., and Grzesiek, S. (1998). Direct Observation of Hydrogen Bonds in Nucleic Acid Base Pairs by Internucleotide2JNNCouplings. *J. Am. Chem. Soc.* 120, 8293–8297. doi:10.1021/ja981513x

Duss, O., Maris, C., von Schroetter, C., and Allain, F. H. (2010). A Fast, Efficient and Sequence-Independent Method for Flexible Multiple Segmental Isotope Labeling of RNA Using Ribozyme and RNase H Cleavage. *Nucleic Acids Res.* 38, e188. doi:10.1093/nar/gkq756

Favier, A., and Brutscher, B. (2011). Recovering Lost Magnetization: Polarization Enhancement in Biomolecular NMR. *J. Biomol. NMR.* 49, 9–15. doi:10.1007/s10858-010-9461-5

Franks, W. T., Zhou, D. H., Wylie, B. J., Money, B. G., Graesser, D. T., Frericks, H. L., et al. (2005). Magic-Angle Spinning Solid-State NMR Spectroscopy of the β1 Immunoglobulin Binding Domain of Protein G (GB1): 15N and13C Chemical Shift Assignments and Conformational Analysis. *J. Am. Chem. Soc.* 127, 12291–12305. doi:10.1021/ja044497e

Fürtig, B., Richter, C., Wöhnert, J., and Schwalbe, H. (2003). NMR Spectroscopy of RNA. *ChemBioChem.* 4, 936–962. doi:10.1002/cbic.200300700

Goldbourt, A. (2019). Structural Characterization of Bacteriophage Viruses by NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* 114-115, 192–210. doi:10.1016/j.pnmrs.2019.06.004

Herbst, C., Riedel, K., Ihle, Y., Leppert, J., Ohlenschläger, O., Görlach, M., et al. (2008). MAS Solid State NMR of RNAs With Multiple Receivers. *J. Biomol. NMR.* 41, 121–125. doi:10.1007/s10858-008-9247-1

Hou, G., Yan, S., Sun, S., Han, Y., Byeon, I.-J. L., Ahn, J., et al. (2011). Spin Diffusion Driven by R-Symmetry Sequences: Applications to Homonuclear Correlation Spectroscopy in MAS NMR of Biological and Organic Solids. *J. Am. Chem. Soc.* 133, 3943–3953. doi:10.1021/ja108650x

Hou, G., Yan, S., Trébosc, J., Amoureux, J.-P., and Polenova, T. (2013). Broadband Homonuclear Correlation Spectroscopy Driven by Combined R2nv Sequences Under Fast Magic Angle Spinning for NMR Structural Analysis of Organic and Biological Solids. *J. Magn. Reson.* 232, 18–30. doi:10.1016/j.jmr.2013.04.009

Hu, W., Kakalis, L. T., Jiang, L., Jiang, F., Ye, X., and Majumdar, A. (1998). 3D HCCH-COSY-TOCSY Experiment for the Assignment of Ribose and Amino Acid Side Chains in C-13 Labeled RNA and Protein. *J. Biomol. NMR.* 12, 559–564. doi:10.1023/a:1008365301124

Huang, W., Bardaro, M. F., Varani, G., and Drobny, G. P. (2012). Preparation of RNA Samples With Narrow Line Widths for Solid State NMR Investigations. *J. Magn. Reson.* 223, 51–54. doi:10.1016/j.jmr.2012.07.018

Huang, W., Varani, G., and Drobny, G. P. (2010). 13C/15N–19F Intermolecular REDOR NMR Study of the Interaction of TAR RNA With Tat Peptides. *J. Am. Chem. Soc.* 132, 17643–17645. doi:10.1021/ja1051439

Huang, W., Varani, G., and Drobny, G. P. (2011). Interactions of Protein Side Chains With RNA Defined With REDOR Solid State NMR. *J. Biomol. NMR.* 51, 347–356. doi:10.1007/s10858-011-9573-6

Igumenova, T. I., Wand, A. J., and McDermott, A. E. (2004). Assignment of the Backbone Resonances for Microcrystalline Ubiquitin. *J. Am. Chem. Soc.* 126, 5323–5331. doi:10.1021/ja030546w

Jaroniec, C. P., Filip, C., and Griffin, R. G. (2002). 3D TEDOR NMR Experiments for the Simultaneous Measurement of Multiple Carbon−Nitrogen Distances in Uniformly 13C,15N-Labeled Solids. *J. Am. Chem. Soc.* 124, 10728–10742. doi:10.1021/ja026385y

Ji, Y., Liang, L., Bao, X., and Hou, G. (2021). Recent Progress in Dipolar Recoupling Techniques Under Fast MAS in Solid-State NMR Spectroscopy. *Solid State. Nucl. Magn. Reson.* 112, 101711. doi:10.1016/j.ssnmr.2020.101711

Juen, M. A., Wunderlich, C. H., Nussbaumer, F., Tollinger, M., Kontaxis, G., Konrat, R., et al. (2016). Excited States of Nucleic Acids Probed by Proton Relaxation Dispersion NMR Spectroscopy. *Angew. Chem. Int. Ed.* 55, 12008–12012. doi:10.1002/anie.201605870

Keane, S. C., Heng, X., Lu, K., Kharytonchyk, S., Ramakrishnan, V., Carter, G., et al. (2015). Structure of the HIV-1 RNA Packaging Signal. *Science* 348, 917–921. doi:10.1126/science.aaa9266

Lange, A., Luca, S., and Baldus, M. (2002). Structural Constraints From Proton-Mediated Rare-Spin Correlation Spectroscopy in Rotating Solids†. *J. Am. Chem. Soc.* 124, 9704–9705. doi:10.1021/ja026691b

Leppert, J., Urbinati, C. R., Häfner, S., Ohlenschläger, O., Swanson, M. S., Görlach, M., et al. (2004). Identification of NH...N Hydrogen Bonds by Magic Angle Spinning Solid State NMR in a Double-Stranded RNA Associated With Myotonic Dystrophy. *Nucleic Acids Res.* 32, 1177–1183. doi:10.1093/nar/gkh288

Lewandowski, J. R., Paëpe, G. D., Eddy, M. T., and Griffin, R. G. (2009). 15N−15N Proton Assisted Recoupling in Magic Angle Spinning NMR. *J. Am. Chem. Soc.* 131, 5769–5776. doi:10.1021/ja806578y

Lu, K., Miyazaki, Y., and Summers, M. F. (2010). Isotope Labeling Strategies for NMR Studies of RNA. *J. Biomol. NMR.* 46, 113–125. doi:10.1007/s10858-009-9375-2

Lusky, O. S., Meir, M., and Goldbourt, A. (2021). Characterizing Hydrogen Bonds in Intact RNA from MS2 Bacteriophage Using Solid State Magic Angle Spinning NMR. *bioRxiv.* [preprint]. 2021.2006.2002.446732. doi:10.1101/2021.06.02.446732

Marchanka, A., Simon, B., Althoff-Ospelt, G., and Carlomagno, T. (2015). RNA Structure Determination by Solid-State NMR Spectroscopy. *Nat. Commun.* 6, 7024. doi:10.1038/ncomms8024

Marchanka, A., and Carlomagno, T. (2014). Solid-State NMR and RNA Structure: A New Partnership? *eMagRes.* 3, 119–128. doi:10.1002/9780470034590.emrstm1358

Marchanka, A., and Carlomagno, T. (2019). "Solid-State NMR Spectroscopy of RNA," in *Chapter Nine - Solid-State NMR Spectroscopy of RNA," in Methods Enzymol*. Editor A.J. Wand (Academic Press), 333–371. doi:10.1016/bs.mie.2018.08.029

Marchanka, A., Kreutz, C., and Carlomagno, T. (2018a). Isotope Labeling for Studying RNA by Solid-State NMR Spectroscopy. *J. Biomol. NMR.* 71, 151–164. doi:10.1007/s10858-018-0180-7

Marchanka, A., Stanek, J., Pintacuda, G., and Carlomagno, T. (2018b). Rapid Access to RNA Resonances by Proton-Detected Solid-State NMR at >100 kHz MAS. *Chem. Commun.* 54, 8972–8975. doi:10.1039/C8CC04437F

Marchanka, A., Simon, B., and Carlomagno, T. (2013). A Suite of Solid-State NMR Experiments for RNA Intranucleotide Resonance Assignment in a 21 kDa Protein-RNA Complex. *Angew. Chem. Int. Ed.* 52, 9996–10001. doi:10.1002/anie.201304779

McDermott, A., Polenova, T., Bockmann, A., Zilm, K. W., Paulsen, E. K., Martin, R. W., et al. (2000). Partial NMR Assignments for Uniformly (13C, 15N)-Enriched BPTI in the Solid State. *J. Biomol. NMR.* 16, 209–219. doi:10.1023/A:1008391625633

Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987). Oligoribonucleotide Synthesis Using T7 RNA Polymerase and Synthetic DNA Templates. *Nucl. Acids Res.* 15, 8783–8798. doi:10.1093/nar/15.21.8783

Milligan, J. F., and Uhlenbeck, O. C. (1989). [5] Synthesis of Small RNAs Using T7 RNA Polymerase. *Methods Enzymol.* 180, 51–62. doi:10.1016/0076-6879(89)80091-6

Morag, O., Abramov, G., and Goldbourt, A. (2014). Complete Chemical Shift Assignment of the ssDNA in the Filamentous Bacteriophage Fd Reports on its Conformation and on its Interface With the Capsid Shell. *J. Am. Chem. Soc.* 136, 2292–2301. doi:10.1021/ja412178n

Nelissen, F. H., van Gammeren, A. J., Tessari, M., Girard, F. C., Heus, H. A., and Wijmenga, S. S. (2008). Multiple Segmental and Selective Isotope Labeling of Large RNA for NMR Structural Studies. *Nucleic Acids Res.* 36, e89. doi:10.1093/nar/gkn397

Nikonowicz, E. P., and Pardi, A. (1993). An Efficient Procedure for Assignment of the Proton, Carbon and Nitrogen Resonances in 13C/15N Labeled Nucleic Acids. *J. Mol. Biol.* 232, 1141–1156. doi:10.1006/jmbi.1993.1466

Nimerovsky, E., Movellan, K. T., Zhang, X. C., Forster, M. C., Najbauer, E., Xue, K., et al. (2021). Proton Detected Solid-State NMR of Membrane Proteins at 28 Tesla (1.2 GHz) and 100 kHz Magic-Angle Spinning. *Biomolecules* 11, 752. doi:10.3390/biom11050752

Nishiyama, Y., Zhang, R., and Ramamoorthy, A. (2014). Finite-Pulse Radio Frequency Driven Recoupling With Phase Cycling for 2D 1H/1H Correlation at Ultrafast MAS Frequencies. *J. Magn. Reson.* 243, 25–32. doi:10.1016/j.jmr.2014.03.004

Olsen, G. L., Bardaro, M. F., Echodu, D. C., Drobny, G. P., and Varani, G. (2010). Intermediate Rate Atomic Trajectories of RNA by Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 132, 303–308. doi:10.1021/ja907515s

Olsen, G. L., Edwards, T. E., Deka, P., Varani, G., Sigurdsson, S. T., and Drobny, G. P. (2005). Monitoring Tat Peptide Binding to TAR RNA by Solid-State 31P-19F REDOR NMR. *Nucleic Acids Res.* 33, 3447–3454. doi:10.1093/nar/gki626

Paramasivam, S., Suiter, C. L., Hou, G., Sun, S., Palmer, M., Hoch, J. C., et al. (2012). Enhanced Sensitivity by Nonuniform Sampling Enables Multidimensional MAS NMR Spectroscopy of Protein Assemblies. *J. Phys. Chem. B.* 116, 7416–7427. doi:10.1021/jp3032786

Pauli, J., Baldus, M., van Rossum, B., de Groot, H., and Oschkinat, H. (2001). Backbone and Side-Chain13C and15N Signal Assignments of the α-Spectrin SH3 Domain by Magic Angle Spinning Solid-State NMR at 17.6 Tesla. *ChemBioChem.* 2, 272–281. doi:10.1002/1439-7633(20010401)2:4<272:aid-cbic272>3.0.co;2-2

Penzel, S., Oss, A., Org, M.-L., Samoson, A., Böckmann, A., Ernst, M., et al. (2019). Spinning Faster: Protein NMR at MAS Frequencies up to 126 kHz. *J. Biomol. NMR.* 73, 19–29. doi:10.1007/s10858-018-0219-9

Pervushin, K., Ono, A., Fernandez, C., Szyperski, T., Kainosho, M., and Wuthrich, K. (1998). NMR Scalar Couplings Across Watson-Crick Base Pair Hydrogen Bonds in DNA Observed by Transverse Relaxation-Optimized Spectroscopy. *Proc. Natl. Acad. Sci.* 95, 14147–14151. doi:10.1073/pnas.95.24.14147

Quinn, C. M., and Polenova, T. (2017). Structural Biology of Supramolecular Assemblies by Magic-Angle Spinning NMR Spectroscopy. *Quart. Rev. Biophys.* 50, e1. doi:10.1017/S0033583516000159

Riedel, K., Herbst, C., Häfner, S., Leppert, J., Ohlenschläger, O., Swanson, M. S., et al. (2006). Constraints on the Structure of (CUG)97 RNA From Magic-Angle-Spinning Solid-State NMR Spectroscopy. *Angew. Chem. Int. Ed.* 45, 5620–5623. doi:10.1002/anie.200600769

Riedel, K., Leppert, J., Ohlenschläger, O., Görlach, M., and Ramachandran, R. (2005a). Characterisation of Hydrogen Bonding Networks in RNAs via Magic Angle Spinning Solid State NMR Spectroscopy. *J. Biomol. NMR.* 31, 331–336. doi:10.1007/s10858-005-1614-6

Riedel, K., Leppert, J. r., Ohlenschläger, O., Görlach, M., and Ramachandran, R. (2005b). TEDOR With Adiabatic Inversion Pulses: Resonance Assignments of 13C/15N Labelled RNAs. *J. Biomol. NMR.* 31, 49–57. doi:10.1007/s10858-004-6066-x

Riedel, K., Leppert, J. r., Häfner, S., Ohlenschläger, O., Görlach, M., and Ramachandran, R. (2004). Homonuclear Chemical Shift Correlation in Rotating Solids via RNvn Symmetry-Based Adiabatic RF Pulse Schemes. *J. Biomol. NMR.* 30, 389–395. doi:10.1007/s10858-004-4056-7

Schledorn, M., Malär, A. A., Torosyan, A., Penzel, S., Klose, D., Oss, A., et al. (2020). Protein NMR Spectroscopy at 150 kHz Magic-Angle Spinning Continues to Improve Resolution and Mass Sensitivity. *ChemBioChem.* 21, 2540–2548. doi:10.1002/cbic.202000341

Scholz, I., Huber, M., Manolikas, T., Meier, B. H., and Ernst, M. (2008). MIRROR Recoupling and its Application to Spin Diffusion Under Fast Magic-Angle Spinning. *Chem. Phys. Lett.* 460, 278–283. doi:10.1016/j.cplett.2008.05.058

Schuetz, A., Wasmer, C., Habenstein, B., Verel, R., Greenwald, J., Riek, R., et al. (2010). Protocols for the Sequential Solid-State NMR Spectroscopic Assignment of a Uniformly Labeled 25 kDa Protein: HET-S(1-227). *Chem. Eur. J. Chem. Bio.* 11, 1543–1551. doi:10.1002/cbic.201000124

Scott, L. G., and Hennig, M. (2008). "RNA Structure Determination by NMR," in *Bioinformatics: Data, Sequence Analysis and Evolution*. Editor J.M. Keith (Totowa, NJ: Humana Press), 29–61. doi:10.1007/978-1-60327-159-2_2

Sergeyev, I. V., Day, L. A., Goldbourt, A., and McDermott, A. E. (2011). Chemical Shifts for the Unusual DNA Structure in Pf1 Bacteriophage from Dynamic-Nuclear-Polarization-Enhanced Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 133, 20208–20217. doi:10.1021/ja2043062

Sergeyev, I. V., Itin, B., Rogawski, R., Day, L. A., and McDermott, A. E. (2017). Efficient Assignment and NMR Analysis of an Intact Virus Using Sequential Side-Chain Correlations and DNP Sensitization. *Proc. Natl. Acad. Sci. USA.* 114, 5171–5176. doi:10.1073/pnas.1701484114

Shaka, A. J., Keeler, J., Frenkiel, T., and Freeman, R. (1983). An Improved Sequence for Broadband Decoupling: WALTZ-16. *J. Magn. Reson. (1969).* 52, 335–338. doi:10.1016/0022-2364(83)90207-x

Shi, L., Kawamura, I., Jung, K.-H., Brown, L. S., and Ladizhansky, V. (2011). Conformation of a Seven-Helical Transmembrane Photosensor in the Lipid Environment. *Angew. Chem. Int. Ed.* 50, 1302–1305. doi:10.1002/anie.201004422

Sklenar, V., Rejante, M. R., Peterson, R. D., Wang, E., and Feigon, J. (1993). Two-Dimensional Triple-Resonance HCNCH experiment for Direct Correlation of Ribose H1' and Base H8, H6 Protons in 13C,15N-Labeled RNA Oligonucleotides. *J. Am. Chem. Soc.* 115, 12181–12182. doi:10.1021/ja00078a073

Sreemantula, A. K., and Marchanka, A. (2020). Solid-State NMR Spectroscopy for Characterization of RNA and RNP Complexes. *Biochem. Soc. Trans.* 48, 1077–1087. doi:10.1042/bst20191080

Szeverenyi, N. M., Sullivan, M. J., and Maciel, G. E. (1982). Observation of Spin Exchange by Two-Dimensional Fourier Transform 13C Cross Polarization-Magic-Angle Spinning. *J. Magn. Reson. (1969).* 47, 462–475. doi:10.1016/0022-2364(82)90213-x

Takegoshi, K., Nakamura, S., and Terao, T. (2001). Dipolar-Assisted Rotational Resonance in Magic-Angle Spinning NMR. *Chem. Phys. Lett.* 344, 631–637. doi:10.1016/s0009-2614(01)00791-6

Tolbert, T. J., and Williamson, J. R. (1997). Preparation of Specifically Deuterated and13C-Labeled RNA for NMR Studies Using Enzymatic Synthesis†. *J. Am. Chem. Soc.* 119, 12100–12108. doi:10.1021/ja9725054

Tolbert, T. J., and Williamson, J. R. (1996). Preparation of Specifically Deuterated RNA for NMR Studies Using a Combination of Chemical and Enzymatic Synthesis. *J. Am. Chem. Soc.* 118, 7929–7940. doi:10.1021/ja961274i

Tzakos, A. G., Easton, L. E., and Lukavsky, P. J. (2007). Preparation of Large RNA Oligonucleotides With Complementary Isotope-Labeled Segments for NMR Structural Studies. *Nat. Protoc.* 2, 2139–2147. doi:10.1038/nprot.2007.306

Ulrich, E. L., Akutsu, H., Doreleijers, J. F., Harano, Y., Ioannidis, Y. E., Lin, J., et al. (2007). BioMagResBank. *Nucleic Acids Res.* 36, D402–D408. doi:10.1093/nar/gkm957

Wacker, A., Weigand, J. E., Akabayov, S. R., Altincekic, N., Bains, J. K., Banijamali, E., et al. (2020). Secondary Structure Determination of Conserved SARS-CoV-2 RNA Elements by NMR Spectroscopy. *Nucleic Acids Res.* 48, 12415–12435. doi:10.1093/nar/gkaa1013

Weingarth, M., Demco, D. E., Bodenhausen, G., and Tekely, P. (2009). Improved Magnetization Transfer in Solid-State NMR With Fast Magic Angle Spinning. *Chem. Phys. Lett.* 469, 342–348. doi:10.1016/j.cplett.2008.12.084

Wiegand, T., Lacabanne, D., Torosyan, A., Boudet, J., Cadalbert, R., Allain, F. H.-T., et al. (2020). Sedimentation Yields Long-Term Stable Protein Samples as

Shown by Solid-State NMR. *Front. Mol. Biosci.* 7, 17. doi:10.3389/fmolb.2020.00017

Wunderlich, C. H., Spitzer, R., Santner, T., Fauster, K., Tollinger, M., and Kreutz, C. (2012). Synthesis of (6-13C)Pyrimidine Nucleotides as Spin-Labels for RNA Dynamics. *J. Am. Chem. Soc.* 134, 7558–7569. doi:10.1021/ja302148g

Yang, Y., Xiang, S., Liu, X., Pei, X., Wu, P., Gong, Q., et al. (2017). Proton-Detected Solid-State NMR Detects the Inter-Nucleotide Correlations and Architecture of Dimeric RNA in Microcrystals. *Chem. Commun.* 53, 12886–12889. doi:10.1039/C7CC07483B

Zhao, S., Yang, Y., Zhao, Y., Li, X., Xue, Y., and Wang, S. (2019). High-Resolution Solid-State NMR Spectroscopy of Hydrated Non-Crystallized RNA. *Chem. Commun.* 55, 13991–13994. doi:10.1039/C9CC06552K

Zhou, D. H., Shah, G., Cormos, M., Mullen, C., Sandoz, D., and Rienstra, C. M. (2007). Proton-Detected Solid-State NMR Spectroscopy of Fully Protonated Proteins at 40 kHz Magic-Angle Spinning. *J. Am. Chem. Soc.* 129, 11791–11801. doi:10.1021/ja073462m

Zwahlen, C., Legault, P., Vincent, S. J. F., Greenblatt, J., Konrat, R., and Kay, L. E. (1997). Methods for Measurement of Intermolecular NOEs by Multinuclear NMR Spectroscopy: Application to a Bacteriophage λ N-Peptide/BoxBRNA Complex. *J. Am. Chem. Soc.* 119, 6711–6721. doi:10.1021/ja970224q

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 21 October 2021
doi: 10.3389/fmolb.2021.743829

Biomolecular Perturbations in In-Cell Dynamic Nuclear Polarization Experiments

Sarah A. Overall and Alexander B. Barnes**

Laboratory of Physical Chemistry, ETH Zürich, Zürich, Switzerland

In-cell DNP is a growing application of NMR to the study of biomolecular structure and function within intact cells. An important unresolved question for in-cell DNP spectroscopy is the integrity of cellular samples under the cryogenic conditions of DNP. Despite the rich literature around cryopreservation of cells in the fields of stem cell/embryonic cell therapeutics, cell line preservation and in cryo-EM applications, the effect of cryopreservation procedures on DNP parameters is unclear. In this report we investigate cell survival and apoptosis in the presence of cryopreserving agents and DNP radicals. We also assess the effects of these reagents on cellular enhancements. We show that the DNP radical AMUPol has no effect on membrane permeability and does not induce apoptosis. Furthermore, the standard aqueous glass forming reagent, comprised of 60/30/10 d_8-glycerol/D_2O/H_2O (DNP juice), rapidly dehydrates cells and induces apoptosis prior to freezing, reducing structural integrity of the sample prior to DNP analysis. Preservation with d_6-DMSO at 10% v/v provided similar DNP enhancements per $\sqrt{}$unit time compared to glycerol preservation with superior maintenance of cell size and membrane integrity prior to freezing. DMSO preservation also greatly enhanced post-thaw survival of cells slow-frozen at 1°C/min. We therefore demonstrate that in-cell DNP-NMR studies should be done with d_6-DMSO as cryoprotectant and raise important considerations for the progression of in-cell DNP-NMR towards the goal of high quality structural studies.

Keywords: in-cell NMR, dynamic nuclear polarization, in-cell DNP-NMR, cryopreservation, apoptosis, cell viability

OPEN ACCESS

Edited by:
Józef Romuald Lewandowski, University of Warwick, United Kingdom

Reviewed by:
Henrike Heise, Heinrich Heine University of Düsseldorf, Germany
Moreno Lelli, University of Florence, Italy

****Correspondence:***
Sarah A. Overall
sarah.overall@phys.chem.ethz.ch
Alexander B. Barnes
alexander.barnes@phys.chem.ethz.ch

Specialty section:
This article was submitted to Structural Biology, a section of the journal Frontiers in Molecular Biosciences

Received: *19 July 2021*
Accepted: *06 October 2021*
Published: *21 October 2021*

Citation:
Overall SA and Barnes AB (2021) Biomolecular Perturbations in In-Cell Dynamic Nuclear Polarization Experiments. Front. Mol. Biosci. 8:743829. doi: 10.3389/fmolb.2021.743829

INTRODUCTION

In-cell NMR is the only technique that can provide structural, dynamic, and composition information of biomolecules in their native, intracellular context and at atomic resolution. However, cellular heterogeneity leads to spectral complexity and reduced sensitivity to spins of interest. These challenges are at the forefront of driving NMR method development. To date, in-cell NMR has yielded *de novo* in-cell structures of GB1 in eukaryotic cells (Muntener et al., 2016; Pan et al., 2016) and TTHA1718 in *E. coli* (Sakakibara et al., 2009), established interactions between α-synuclein and cellular chaperones (Burmann et al., 2020) and monitored structural maturation of SOD1 during protein folding *in situ* (Banci et al., 2013). Much of this work was pioneered by the work of Selenko and co-workers who demonstrated the transfer of high concentrations of uniformly labelled proteins into cells using electroporation and microinjection (Selenko et al., 2006). This circumvents some of the spectral complexity observed in in-cell NMR spectra. However, this requires very high concentrations of exogenously produced soluble proteins, which is not applicable to

membrane proteins, transient protein complexes, proteins which can't be concentrated or recombinantly produced. In addition, highly elevated cellular concentrations of the protein of interest are established, which for most proteins is non-physiological. The detection of physiologically relevant protein systems requires significant improvements in NMR sensitivity. Dynamic nuclear polarization (DNP) is one such technology. In-cell Dynamic Nuclear Polarization Nuclear Magnetic Resonance (DNP-NMR) holds great promise for physiologically relevant biomolecular structure determination, particularly for membrane bound proteins but also for the study of a wide range of cellular phenomena at the atomic level as demonstrated by in-cell NMR studies in both solution and solid state (Theillet et al., 2016; Thongsomboon et al., 2018; Overall et al., 2019; Luchinat et al., 2020). The sensitivity gains from DNP give access to many dilute biomolecular systems that are otherwise inaccessible to structural and compositional studies within the physiological context of an intact cell. The first demonstration of in-cell DNP within intact mammalian cells showed that significant enhancements could be achieved in the context of intact cells with a range of polarizing agents (Albert et al., 2019), with an added benefit that the DNP enhancement can be targeted to subcellular compartments. Subsequent studies used DNP-NMR to observe exogenously labelled ubiquitin electroporated into HeLa cells (Narasimhan et al., 2019), drug mediated expression of HIV particles (Overall et al., 2020) and antisense RNA drug complexes in HEK 293T cells (Schlagnitweit et al., 2019). However, this potential remains largely untapped. Optimization of sample preparation and an understanding of the cellular effects of DNP preparation are lacking.

The study of cryopreservation is a rich and active field of research that has extensively explored many different cryopreserving reagents and formulations since the birth of cryobiology in the 1950s. Cooling methods have also been explored to improve the storage of cells and tissue for research and therapeutic purposes (Raju et al., 2021) (Hornberger et al., 2019). Much of the reported work is centered around reducing intracellular ice crystal growth, considered to be the most damaging effect of cryogenic preservation (Mazur 2004). Vitrification, the formation of a non-crystalline amorphous glass, is an effective method for limiting ice crystal formation and maintaining high cell viability (Fahy and Wowk 2020). Vitrification is generally achieved through rapid cooling and provides the highest resolution of cellular structures by cryo-electron microscopy as well as being the method of choice for blastocyst and embryo preservation in IVF clinics (Sekhon et al., 2018) (Valjerdi et al., 2009). Thus, vitrification is an important method for the cryopreservation of mammalian cells. In reality, cooling rates required to vitrify cellular samples can be difficult or impossible to achieve for samples larger than a few nano liters (Berejnov et al., 2006). As a result, slow-cooling at 1°C/min is commonly used for the long-term storage of cell lines and primary cells for its ease of use (Baust et al., 2009).

Significant dehydration is observed among slow-cooled cryopreserved cells caused by the extraction of water from the intracellular compartment (Mazur 1984). Some have suggested that this dehydration event enhances cell survival by removing water that might otherwise nucleate ice crystals intracellularly (Meneghel et al., 2019). However, cryoprotecting media that maintains cell size by reducing dehydration proves much more effective at preserving cell viability (Huebinger et al., 2016). Molecular adaptations during slow-freezing or through interactions with cryoprotectants are thought to contribute to the ability of cells to withstand and recover from ice crystal growth (Mazur 2004; Meneghel et al., 2019). Dimethyl sulfoxide (DMSO) is an important and widely used cryopreserving agent for mammalian cells, which readily diffuses through cell membranes and appears to particularly enhance the adaptability of cell membranes to ice damage (Shi et al., 2001; Huebinger 2018) in addition to reducing lethal intracellular ice formation (Baust et al., 2009).

Despite this research, in-cell DNP studies have often utilized glycerol based preservation (60:30:10 d_8-glycerol:D_2O:H_2O) and liquid nitrogen flash-freezing for sample preparation. Flash-freezing, typically not adequate to vitrify large sample volumes, has been implemented for in-cell studies due to the reported instability of DNP radicals to the reducing environment of the intracellular compartment by us and others (Giotta and Wang 1972; Albert et al., 2019; McCoy et al., 2019). The use of DNP juice was established by Hall and co-workers, who showed that DNP of biomolecular samples could be greatly improved using high concentrations of glycerol to promote the formation of a homogenous glass (Hall et al., 1997). Ice crystal formation is associated with lower DNP enhancements due to increased paramagnetic quenching as the local concentration of radicals outside of ice crystals increases (Corzilius et al., 2014) further supporting the importance of glass forming reagents. Further studies revealed that heterogenous imperfections of the glassy matrix reduce polarization transfer, even at glycerol concentrations of <55% (Leavesley et al., 2018b). As a result, net DNP enhancements are maximized in homogenous glasses promoted by glycerol in aqueous samples.

Cellular samples are very different. High concentrations of any one molecule are often greatly detrimental to cellular homeostasis, even over short periods of time, creating osmotic imbalances that impose significant stress onto cells as well as other toxic effects (Baust et al., 2009). Furthermore, the distribution of cryopreserving reagents will inherently be heterogenous due to the heterogenous distribution of cellular structures with differing solvent accessibilities. This raises questions about the uniformity of in-cell enhancements and the appropriateness of glycerol based protecting reagents for in-cell studies. Here we perform a study of cellular integrity before and after DNP analysis and assess the effects of different cryoprotecting agents on DNP parameters and spectral quality of Jurkat T cells at natural abundance.

MATERIALS AND METHODS

Cell Culture

Jurkat T cells and a variant of the Jurkat T cell line JLat9.2 cells (containing a genomically integrated HIV genome that is basally inactive (Jordan et al., 2001)) were cultured at 37°C with 5% CO_2

atmosphere in unlabeled complete RPMI (2 mM L-Glutamine, 10% v/v Fetal Bovine Serum (FBS) (Gibco), 100 U/ml penicillin-100 μg/ml streptomycin (Gibco) and 10 mM sodium pyruvate (Gibco). Cells were counted using a hemocytometer and trypan blue (sigma-aldrich) staining by preparing a ½ dilution of 0.4% trypan blue with cells and observing under a light microscope. Blue cells were counted as dead and non-blue refractive cells were counted as viable.

Flow Cytometry

1×10^6 total cells were transferred to 3 ml flow cytometry tubes and washed with 1 ml FACS buffer (phosphate buffered saline (PBS), 1% Bovine serum albumin (BSA), 2 mM EDTA). Cells were then incubated with 1:100 diluted BioTracker NucView Caspase-3-405 (Biotium) for 15 min on ice then washed with 1 ml FACS buffer and resuspended in 500 μl FACS buffer prior to analysis. 1 min before analysis 10 μl of 10 μM propidium iodide (PI) was added and the sample analyzed on an LSR Fortessa flow cytometer (BD Biosciences). Data was analyzed with FlowJo software (treestar). All events were analyzed and reported. The only gating used was to remove doublets as shown in **Supplementary Figure S1**.

Annexin-V Staining

Annexin-V staining was done after caspase-3 staining. The cells were washed with Annexin-V binding buffer (0.01 M HEPES pH 7.4, 0.14 mM NaCl, 2.5 mM $CaCl_2$). Then resuspended in 100 μl Annexin-V binding buffer and 5 μl Annexin-V-488 (Sigma) added to each sample and incubated on ice for 10 min then diluted with 1 ml of Annexin-V binding buffer before immediately analyzing without washing.

Pre-Freeze Analysis

10×10^6 total cells were transferred to FACS tubes and washed with 2 ml FACS buffer by centrifugation. Cell pellets were then resuspended in 10 μl of cryopreservative as described in the sample preparation section but in the absence of radical. The cells were then incubated on ice for 10 min followed by washing with 1 ml of FACS buffer, resuspended in 100 μl of FACS buffer and stained for caspase-3 and Annexin-V as described above.

Post-Thaw Cell Culture

Frozen sapphire rotors containing cells were thawed in a 37°C water bath for no more than 5 s. Zirconia caps were immediately removed and the cells were collected by centrifugation upside down in a 15 ml conical tube for 1 min at 1,500 rpm at 4°C. The cells were immediately and gently resuspended in 3 ml prewarmed complete RPMI then transferred to six well plates and cultured for 24 h at 37°C prior to FACS analysis and trypan blue counting.

DNP-NMR Sample Preparation

Cell samples were prepared by washing 40×10^6 total JLat 9.2 or Jurkat T cells with 5 ml ice cold phosphate buffered saline (PBS) and gently pelleted at 1,500 rpm for 5 min at 4°C to remove culture media. The cell pellet (approximately 40 μl) was resuspended in 1 ml of deuterated PBS and incubated on ice for 10 min to allow exchange of intracellular water with D_2O. The cells were then pelleted again and resuspended in an equal volume of 2 fold concentrated cryopreserving reagent with DNP radicals. (20% DMSO with 20 mM AMUPol, 60/30/10 glycerol/D_2O/H_2O with 20 mM AMUPol or PBS with 20 mM AMUPol), to give a final AMUPol concentration of 10 mM (Sauvee et al., 2013). The cells were then incubated for the indicated times before packing into 3.2 mm sapphire cylindrical rotors (Bruker Biospin) using custom made Teflon filling tools (Overall et al., 2020) by centrifugation for 1 min at 1,500 rpm. Rotors were plunge frozen in liquid nitrogen for 10 min before capping with zirconia drive caps. Rotors were stored in liquid nitrogen prior to DNP analysis.

Slow-frozen cells were prepared as above but after filling, the rotors were immediately capped with zirconia drive caps and slow-cooled in a Mr. Frosty (Nalgene) at −80°C overnight before plunge freezing in liquid nitrogen for storage.

Solid-State DNP-NMR

Solid-state DNP-NMR was conducted on a 14 T Bruker DNP spectrometer operating at 600 MHz ^{1}H Larmor frequency and equipped with a second harmonic 365 GHz gyrotron and a 3.2 mm HX or HXY LTMAS probe. Samples were spun at 9 kHz with a sample temperature of ~104–108 K (microwaves off) and ~110–114 K (microwaves on). ^{13}C spectra were acquired using a cross polarization (CP) scheme with typical ^{1}H spin-locking amplitude of 100 kHz over a 1 ms tangential ramp centered at 50 kHz. Data was acquired under spinal64 ^{1}H decoupling at 100 kHz with 512 transients and recycle delay 1.26^*T_1. Enhancements were calculated as both:

The peak intensity ratio: $\varepsilon_{on/off} = \frac{Ion}{Ioff}$ and enhancement per $\sqrt{}$unit time: $(\frac{\varepsilon}{\sqrt{TB}})$. Effective enhancement per $\sqrt{}$unit time was calculated by: $(\frac{\varepsilon}{\sqrt{TB}}) \times \xi$

The quenching factor ξ was calculated as: $\xi = 1 - \frac{I}{I_0}$. Polarization build up times (T_B) were measured using a saturation recovery CP-sequence with 20 × 5 μs saturating pulses prior to the recovery delay.

Data Analysis

1D NMR data was processed in Topspin 4.07 using 50 Hz line broadening. Cellular components were assigned using data from the BMRB. T_B was calculated from peak intensities and fit to the equation:

$$Mz(t) = Mz\left(1 - e^{-\frac{t}{TB}}\right)$$

Only fits with an R^2 value of >0.97 are reported and used for further analysis. Statistical data was plotted and processed using GraphPad Prism 9. Statistical significance was determined using a parametric unpaired t test.

RESULTS

Pre-Freeze Viability of Cells Prepared for DNP

In order to characterize the effect DNP preparation has on cell integrity and its relationship to DNP enhancements and spectral

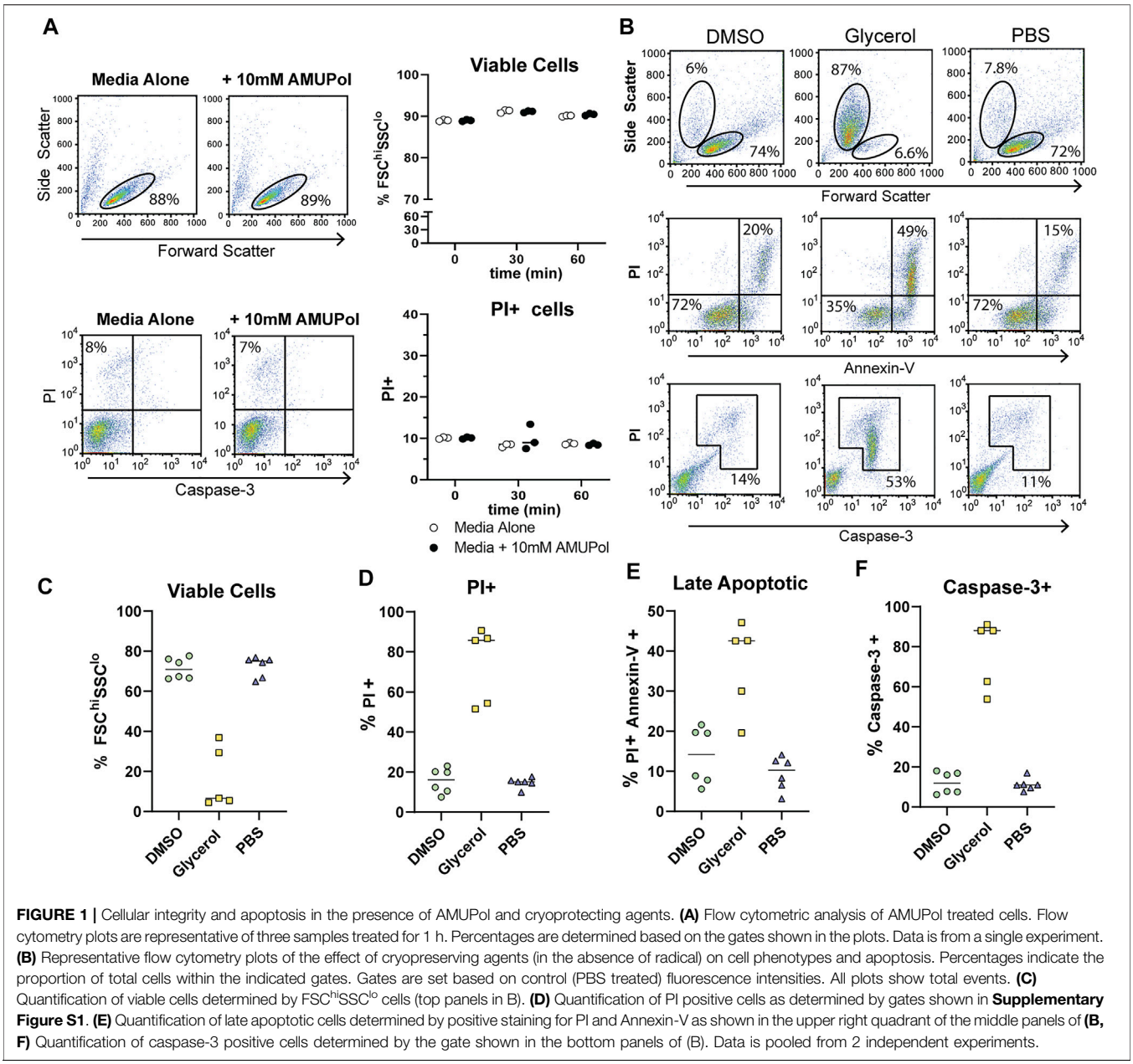


FIGURE 1 | Cellular integrity and apoptosis in the presence of AMUPol and cryoprotecting agents. **(A)** Flow cytometric analysis of AMUPol treated cells. Flow cytometry plots are representative of three samples treated for 1 h. Percentages are determined based on the gates shown in the plots. Data is from a single experiment. **(B)** Representative flow cytometry plots of the effect of cryopreserving agents (in the absence of radical) on cell phenotypes and apoptosis. Percentages indicate the proportion of total cells within the indicated gates. Gates are set based on control (PBS treated) fluorescence intensities. All plots show total events. **(C)** Quantification of viable cells determined by $FSC^{hi}SSC^{lo}$ cells (top panels in B). **(D)** Quantification of PI positive cells as determined by gates shown in **Supplementary Figure S1**. **(E)** Quantification of late apoptotic cells determined by positive staining for PI and Annexin-V as shown in the upper right quadrant of the middle panels of **(B, F)** Quantification of caspase-3 positive cells determined by the gate shown in the bottom panels of (B). Data is pooled from 2 independent experiments.

quality, we first assessed the cellular effects of DNP radicals and cryoprotecting reagents prior to cryogenic freezing. Comparison of the effects of AMUPol on cell size and viability by flow cytometry were carried out by incubating Jurkat T cells with 10 mM AMUPol for up to 1 h at 4°C. Cell viability was assessed using flow cytometry to measure forward light scatter (cell size), propidium iodide (PI) uptake (which reports on membrane permeability as the negatively charged PI can-not cross an intact plasma membrane) and caspase-3 cleavage (a marker of apoptosis). AMUPol had no effect on cell viability as determined by no change in the proportion of cells uptaking propidium iodide (PI) or exhibiting active caspase-3 cleavage (**Figure 1A**). Furthermore, there was no change in cell size [forward scatter intensity (FSC)] or cell density [side scatter intensity (SSC)], collectively indicating 10 mM AMUPol has no effect on membrane integrity or cellular homeostasis (**Figure 1A**).

The effects of cryopreserving media were then assessed after preparing cells for DNP in the absence of radical since the presence of 10 mM AMUPol had no effect on cell phenotypes. Cells were then incubated on ice for 10 min before staining for Annexin-V, caspase-3, and PI. Prior to the addition of cryoprotecting agent, cells were washed with PBS in D_2O which has been demonstrated to increase radical T_{2e} and is correlated with increased DNP enhancements (Martorana et al., 2014; Sauvee et al., 2013). Exposing cells to D_2O for short periods of time (within 24 h) has previously been shown to have no effect on cell morphology or viability (Martorana et al., 2014; Siegel et al., 1960). Incubation with 10% dimethyl sulfoxide

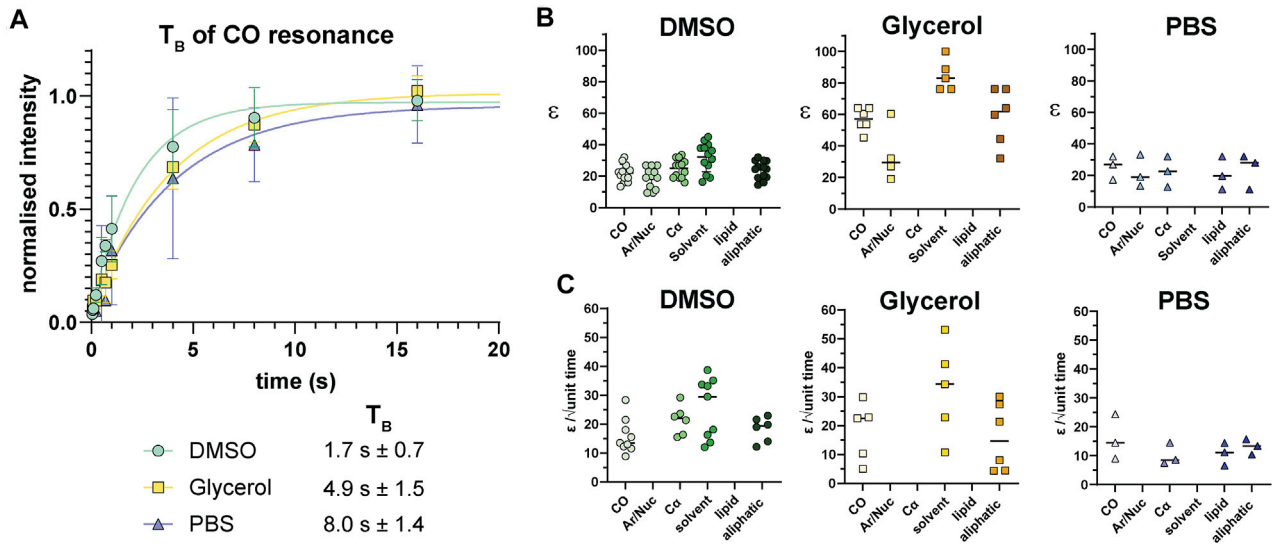


FIGURE 2 | ^{13}C enhancements of Jurkat T cells. **(A)** T_B build up curves for carbonyl resonances with different cryoprotecting media. Error bars indicate the SD. **(B)** ^{13}C enhancements (calculated as $I_{on}/I_{off} = \varepsilon_{on/off}$) in different cryoprotecting media across various cellular components as determined using data from the BMRB. **(C)** Enhancements calculated per √unit time using paired data from B and **Supplementary Table I**. Each symbol represents a single sample. Horizontal bars indicate the mean.

(DMSO) had no observable effect on cell size or density with comparable FSC vs SSC profiles to control cells, which were treated with PBS (**Figure 1B** top panels and **Figure 1C**). PI uptake was not increased (**Figure 1D**) but there was a slight increase in late apoptotic cells compared to PBS treated cells, increasing from an average of 11–14% but as high as 20% in some samples (**Figure 1B** middle panels and **Figure 1E**) indicative of apoptosis activation. However, this was not accompanied by caspase-3 cleavage (**Figure 1B** bottom panels and **Figure 1F**), which shows the apoptotic phenotype induced by DMSO to be reversible (Kroemer et al., 2009).

The effects of glycerol preservation were tested by adding equal volumes of DNP matrix (60/30/10 glycerol/D_2O/H_2O) to the cell pellet, resulting in a final glycerol concentration of 30% v/v. Cells incubated with the glycerol mixture underwent significant dehydration in which cell size decreased by 50% and cell density increased by 50% (**Figures 1B,C**). Furthermore, glycerol induced significant late-stage apoptosis with 42% of cells staining for Annexin-V concomitant with a large increase in PI uptake and caspase-3 cleavage by cells exhibiting a dehydrated profile (**Figures 1B–F**), indicative of extensive dehydration and apoptosis prior to freezing. This establishes a significantly altered cellular phenotype and possibly structure of glycerol treated cells, prior to DNP analysis. On the other hand, DMSO induced significantly fewer perturbations prior to freezing.

DNP Analysis of Cells Preserved With Different Cryoprotecting Formulations

Following analysis of pre-freeze viability, we prepared unlabeled Jurkat T cells for DNP in cryopreservation media and a final radical concentration of 10 mM AMUPol. Samples were then flash-frozen in liquid nitrogen and the effects of different cryopreserving reagents on DNP enhancements were assessed. Throughout this work, a single sample refers to a single rotor containing ~22 million cells. Comparison of polarization build up times (T_B) of carbonyl resonances revealed cells preserved with glycerol exhibited a 2.9-fold longer T_B value of 4.9 s ± 1.5 compared to DMSO preserved cells at 1.7 s ± 0.7 and 8 s ± 1.4 for PBS preserved cells. (**Figure 2A**; **Table 1**). Initial comparison of cellular enhancements (determined by $I_{on}/I_{off} = \varepsilon_{on/off}$) showed glycerol preservation provided 2.6-fold greater $\varepsilon_{on/off}$ of both solvent and cellular carbonyl (CO) signals compared to DMSO preserved cells (**Figure 2B**) and was mirrored in signal-to-noise ratios. However, taking into account T_B reveals glycerol preservation provides only a slightly larger enhancement per √unit time giving an average carbonyl enhancement of 18.1/√unit time ± 10 compared to 15.8/√unit time ± 6 for DMSO preserved cells and 16/unit time ± 7.8 for PBS preserved cells (**Figure 2C**). Interestingly, there was no correlation between T_B values and enhancements (**Supplementary Figure S2**) on a per sample basis, accounting for the spread in ε/√unit time. However, average values more closely resemble the linear relationship between $\varepsilon_{on/off}$ and T_B predicted by simulations and observed experimentally by others (Mentink-Vigier et al., 2017). This is further emphasized by the sample variation in $\varepsilon_{on/off}$ values observed for all cryopreservatives used. This could be reflective of heterogeneity in glass formation which may be a function of sample temperature history as glass homogeneity significantly contributes to DNP enhancements (Hall et al.,

TABLE 1 | T_B for different cellular resonances preserved under various preservation methods.

Resonance	Chemical shift	DMSO	Glycerol	PBS
Carbonyl (CO)	176 ppm	1.7 s ± 0.7	4.9 ± 1.5	8 s ± 1.4
Aromatic/Nucleic Acids	133 ppm	ND	ND	ND
Carbon α (Protein)	56 ppm	1.7 s ± 0.7	ND	6.6 s ± 0.6
Solvent	39.5 ppm (DMSO)	1.3 s ± 0.7	NA	NA
	65.2 ppm (Glycerol)	NA	3.2 s ± 1.0	NA
Lipid	34.5 ppm	ND	ND	6.6 s ± 0.6
Aliphatic	28 ppm	1.7 s ± 0.7	5 s ± 1.2	6.1 s ± 1.6
Number of samples		6	5	3

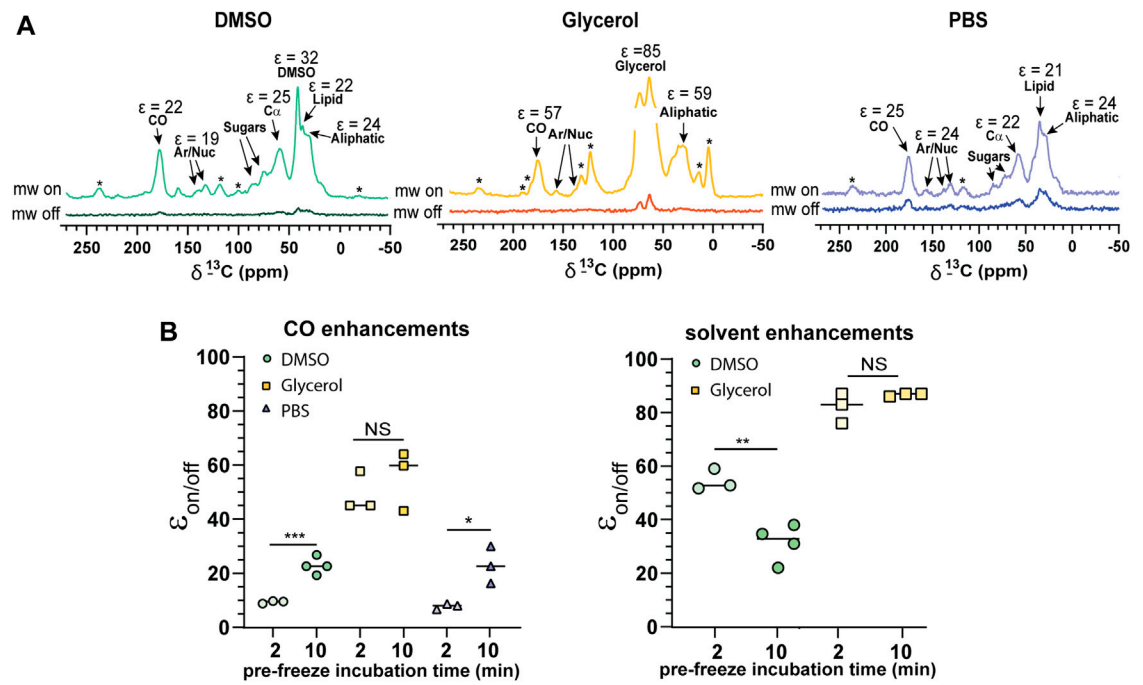


FIGURE 3 | Cellular penetration of AMUPol in Jurkat T cells. **(A)** Comparison of ^{1}H-^{13}C CP spectra of 10% DMSO, glycerol, and PBS protected Jurkat T cells. Cellular components and their average enhancements are indicated. Enhancements are given as peak intensity ratios as not all signals could have the ε/√unit time determined. * indicates spinning side bands. **(B)** Enhancements obtained after 2 and 10 min incubations with AMUPol prior to flash-freezing for CO and solvent peaks. Statistical significance was determined using a parametric *t*-test where * = $p < 0.05$, ** = $p < 0.01$ and *** = $p < 0.001$, NS = not significant. DNP experiments were carried out at 600 MHz ^{1}H larmor frequency at 9 kHz MAS.

1997; Corzilius et al., 2014; Leavesley et al., 2018b). For frozen samples, significant unintentional temperature changes could be experienced during insertion/ejection from the probe and during capping (in the case of flash-frozen samples). Especially when using sapphire rotors due to the high thermal conductivity of sapphire. It is also likely that the $\varepsilon_{on/off}$ values were additionally affected by variation in instrumentation performance such as microwave output and temperature regulation. We also measured quenching effects of AMUPol by acquiring in-cell spectra in the absence of radical. We found that in the presence of DMSO, 10 mM AMUPol results in a 50% reduction in the signal intensity of microwave off spectra, giving a quenching factor (ξ) of 0.5. In the presence of glycerol, the microwave off signal was quenched by 75% upon addition of AMUPol (**Supplementary Figure S3**), resulting in ξ of 0.25. If we take into account these measured quenching effects, then glycerol clearly performs poorly compared to 10% DMSO, giving an effective enhancement/√unit time of 4.5 s ± 2.5 and 7.9 s ± 3, respectively.

We observed relatively uniform enhancements over various cellular components from CO, Cα, aromatics and lipids after incubating cells with 10 mM AMUPol for 10 min prior to freezing (**Figure 3A**; **Supplementary Figure S4**; **Supplementary Table I**) suggesting an even distribution of AMUPol with respect to different cellular components. Due to the dominance of the glycerol peak and its spinning side bands we were not able to identify differences in the enhancement of sugars, or Cα

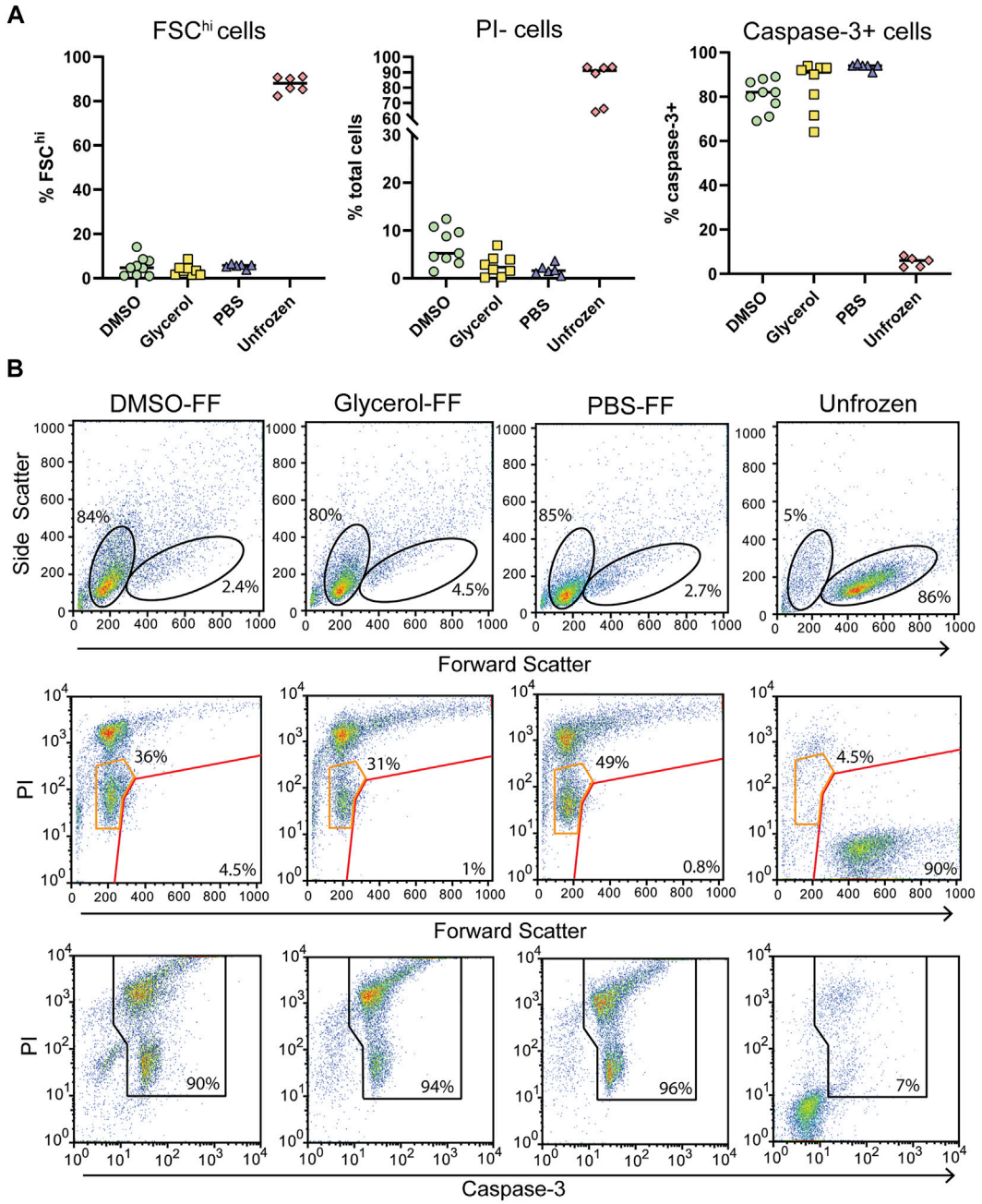


FIGURE 4 | Post DNP cellular integrity and apoptosis induction with flash-freezing (FF). **(A)** Post-thaw viability as determined by FSC fluorescence, PI, and capsase-3 staining. Each symbol represents a single rotor. **(B)** Representative flow cytometry plots of samples treated with various cryoprotecting reagents followed by flash-freezing (FF). Top panels show the gating strategy used to determine FSC^{hi} cells. Middle panels show PI uptake of FSC^{hi} cells (red gate) and FSC^{lo} cells (orange gate) cells Bottom panels show the gating strategy used to determine caspase-3+ cells.

components in glycerol preserved cells but observed no difference in the enhancements of carbonyl and aliphatic components (**Figure 3A**). The exception was solvent $\varepsilon_{on/off}$ values suggestive of increased radical concentration extracellularly. Consistent with this hypothesis, T_B values for solvent resonances was generally lower compared to cellular signals (**Table 1**). In addition, we also observed a time dependent component to the enhancements in DMSO treated cells but not glycerol treated cells. Incubating cells for 10 min prior to flash-freezing increased the enhancement of cellular signals by about 3-fold compared to those frozen after 2 min (**Figure 3B**). Concurrently, the solvent (DMSO) enhancement decreased from an average enhancement of $\varepsilon_{on/off} = 52$ after 2 min incubation to $\varepsilon_{on/off} = 32$ after 10 min, suggesting that diffusion of AMUPol into

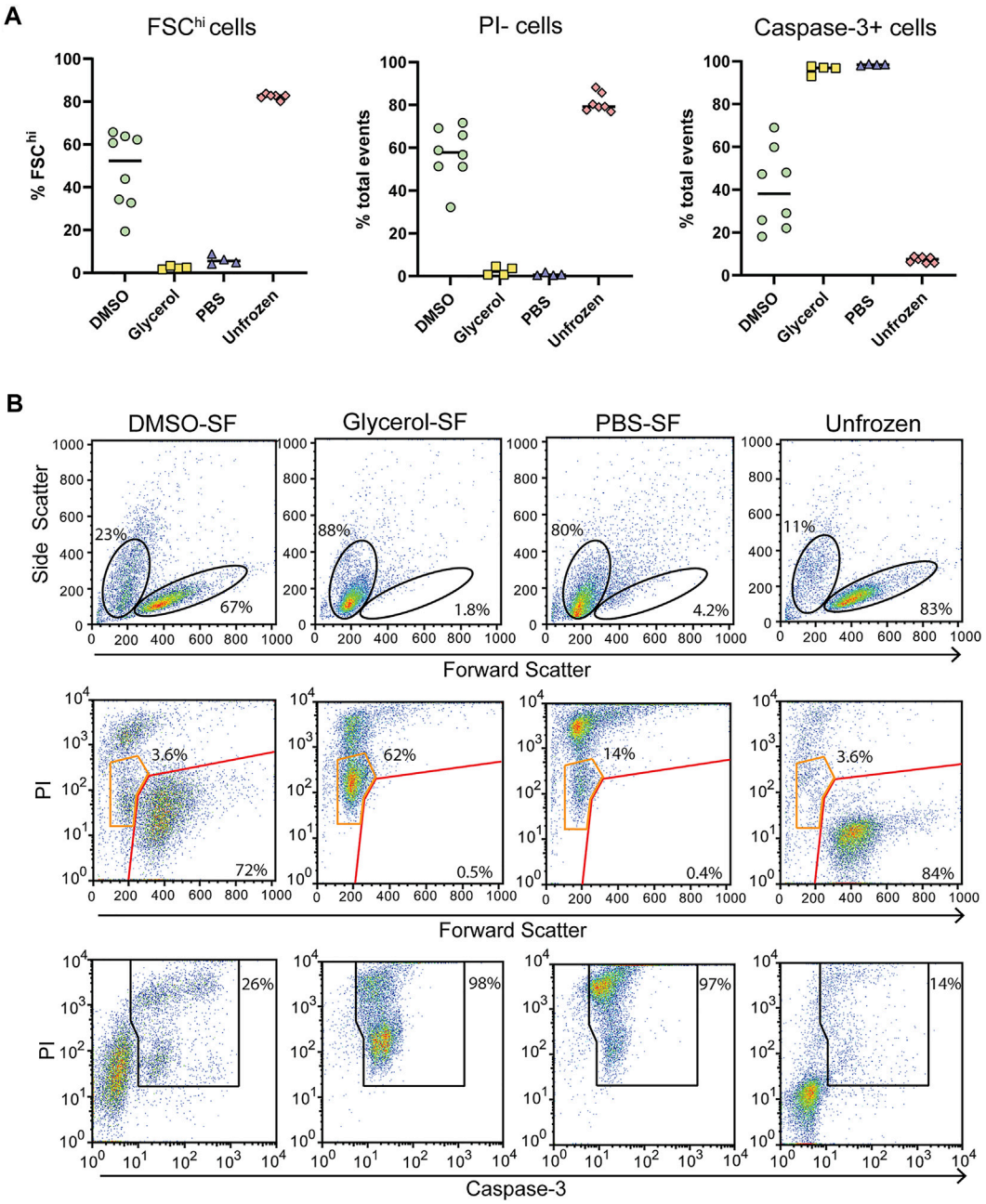


FIGURE 5 | Post thaw cellular integrity and apoptosis induction with slow-freezing (SF). **(A)** Quantification of cell viability by flow cytometry through FSC fluorescence, PI, and caspase-3 staining. Each symbol represents a single sample/rotor. **(B)** Representative flow cytometry plots show the gating strategy used to determine the percentages in (A). All plots show total events. Top panels show the gating strategy used to determine FSC^{hi} cells. Middle panels show PI uptake of FSC^{hi} cells (red gate) and FSC^{lo} cells (orange gate) cells. Bottom panels show the gating strategy used to determine caspase-3+ cells.

cells is delayed and solvent enhancements can be attributed to increased radical concentration outside the cell. This also implies that the concentration of AMUPol could be significantly increased to improve both $\varepsilon_{on/off}$ and $\varepsilon/\sqrt{\text{unit time}}$ values. The difference in time dependency of enhancements in DMSO prepared samples, compared to cells preserved with glycerol may be due to the effects of DMSO on membrane surfaces (Schrader et al., 2016). Accumulation of DMSO at the interfacial layer of the membrane may reduce membrane permeability to AMUPol. Certainly, the difference in dehydration effects of glycerol compared to DMSO would support this hypothesis. We did not observe an increase in enhancement after 10 min which may represent the time period required for the bulk cellular distribution of AMUPol (**Supplementary Figure S5**). This was also observed by Baldus and colleagues by

microscopy using PyPol, a fluorescent variant of AMUPol (Narasimhan et al., 2019) where cellular fluorescence with PyPol peaked after 10–15 min.

Post-DNP Cell Viability of Jurkat T Cells

Following in-cell DNP, sample rotors were thawed by rapid warming in a 37°C water bath for 5 s and the cells immediately placed in culture for 24 h and assessed for viability. We observed low cell survival of flash-frozen cells, indicating poor vitrification across the ~22 μL volume. Assessment of viability by trypan blue staining varied greatly between samples and also over-estimated the viability compared to the percentage of PI + cells detected by flow cytometry (**Figure 4A**; **Supplementary Figure S6**). Significant cell shrinkage and PI staining was observed with few to no viable PI- cells (**Figure 4B**). A significant population of $PI^{lo}FSC^{lo}$ cells were observed in frozen samples but not unfrozen samples (**Figure 4B** orange gates). These cells were predominantly caspase-3+, indicating that they are apoptotic and unlikely to be viable. Furthermore, viability was uncorrelated with cryoprotecting media as viability was equivalent to unprotected cells. Thus, the data clearly demonstrates poor cell viability of flash-frozen cells in 3.2 mm rotors. This would suggest significant intracellular ice crystal growth with this procedure.

For comparison, we assessed cells prepared the same way but that were control frozen with an approximate cooling rate of 1°C/min. Upon thawing only DMSO preserved cells benefited from slow-freezing with a significant increase in FSC^{hi} cells and reduced PI uptake (**Figures 5A,B**). Cell size and density was comparable to unfrozen cells and which generally correlates with increased viability. Glycerol and PBS preserved cells failed to recover from slow-freezing and exhibited equivalent FSC and SSC profiles to flash-frozen cells. In the case of glycerol preservation, PI fluorescence was lower, with most cells appearing as PI^{lo} compared to flash-frozen cells (**Figure 5B** orange gates). This was in contrast to PBS frozen cells which exhibited a predominantly PI^{hi} phenotype. Again, the viability of PI^{lo} cells is expected to be negligible at least 24 h post-thaw indicated by the low FSC fluorescence and high caspase-3 cleavage. However, the lower PI staining does indicate reduced membrane damage in the presence of glycerol, even if it is insufficient to preserve cell viability. Again, trypan blue staining over-estimated cell viability (**Supplementary Figure S6**). This could be due to cells faintly staining for trypan blue but remaining refractive under a light microscope. These cells likely take up PI and could form the PI^{lo} population we observed by flow cytometry. In addition, extreme dehydration caused by freezing results in a large reduction in cell size which could render cells too small to be recognized as cells by microscopy and are instead ignored as debris (both manually and with automated cell counters). This can be particularly challenging with a small cell type such as Jurkat T cells whose average size is 10 μm. These cells would shrink to <5 μm as indicated by flow cytometry, which is at or below the size limit for most automated cells counters and certainly likely to be difficult to distinguish by manual counting.

DNP Analysis of Slow-Frozen Jurkat T Cells

Slow-frozen cells were also analyzed by DNP-NMR. We observed no significant spectral differences between slow-frozen and flash-frozen cells in 1D spectra. This was expected given the 1D spectra represent an average chemical shift of all cellular components and thus reports on chemical composition more so than structural features. (**Figures 6A,B**). The relative intensity of the DMSO peak was reduced compared to cellular signals in slow-frozen samples (**Figure 6B**). We think this reflects the accumulation of DMSO at the interfacial region of membranes reducing its accessibility to AMUPol, reducing net enhancements of DMSO (Liao et al., 2016; Schrader et al., 2016). Enhancements by peak intensity ratios were lower in slow-frozen cells compared to flash-frozen cells, although this was highly variable (**Table 2**). The calculated T_B time was increased in slow-frozen cells by an order of magnitude and consistent between the two samples assessed, averaging 17.3 s ± 1.4 for carbonyl resonances (**Table 2**). The calculated $\varepsilon/\sqrt{\text{unit time}}$ was very low in slow-frozen cells, averaging $2.7/\sqrt{\text{unit time}} \pm 0.3$ compared to $12.7/\sqrt{\text{unit time}} \pm 0.8$ for flash-frozen cells. This data would be consistent with a lower radical concentration. It is unclear whether this is due to reduction of the radical during slow-freezing or a more even distribution of AMUPol resulting in reduced quenching effects but it is likely to be a combination of both.

Certainly, the DNP performance of slow-frozen Jurkat T cells was markedly poorer compared to flash-frozen Jurkat T cells, despite their improved viability and raises the possibility that AMUPol may not be the optimal radical for in-cell DNP-NMR studies at least in Jurkat T cells or that significantly higher AMUPol concentrations are required, though this must be balanced with paramagnetic quenching effects. Alternatively, a move towards improved vitrification methods may prove more valuable.

DISCUSSION

In-cell DNP NMR is an impactful new Frontier in biomolecular NMR. The promise of physiologically relevant biomolecular structures and interactions is heavily complicated by the heterogeneity and compositional complexity of cellular samples. This poses an extreme challenge to atomic scale analyses such as NMR. The strong dependence of sample quality on the quality of subsequent NMR data highlights the need for rigorous studies of cell preparation methods for in-cell DNP-NMR and especially given the added complication of cryogenic temperatures. The work presented here provides a significant contribution to this analysis and highlights key considerations when performing in-cell DNP-NMR towards improved *in situ* structural studies.

The goal of in-cell NMR is to capture native biomolecular conformations at the time of analysis. This is under the reasonable assumption that conformational relevance can be linked with cell viability. Cell viability after cryogenic freezing is affected in two independent ways. The first is damage done during freezing or by cryoprotectants (prior to analysis) and the second is damage done during warming

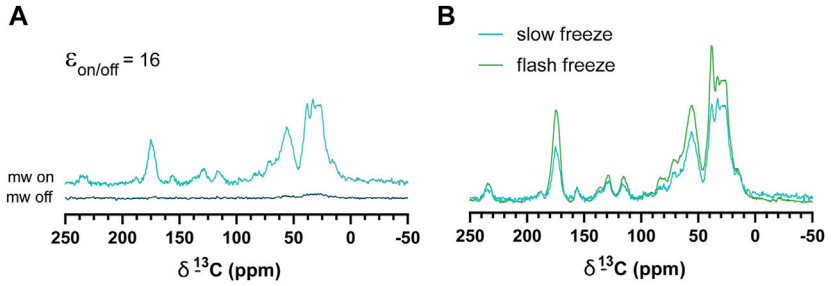


FIGURE 6 | DNP enhancements of slow-frozen cells. **(A)** ^{1}H-^{13}C CP spectra of slow-frozen Jurkat T cells preserved with 10% DMSO. The enhancement given is the peak intensity ratio of the carbonyl resonance. **(B)** Comparison of microwave on spectra from slow-frozen cells in 10% DMSO (dark teal line) to flash-frozen cells in 10% DMSO (green line). DNP experiments were carried out at 600 MHz ^{1}H larmor frequency at 9 kHz MAS.

TABLE 2 | Average T_B values and enhancements ($\varepsilon_{on/off}$) of slow-frozen Jurkat T cells compared to flash-frozen Jurkat T cells.

		Slow-frozen		Flash-frozen	
Resonance	**Chemical shift (ppm)**	**TB**	**εon/off**	**TB**	**εon/off**
Carbonyl (CO)	176	17.3 s ± 1.4	11 ± 6	1.75 s ± 0.2	17 ± 1.7
Aromatic/Nucleic Acids	133	17.1 s ± 1.1	ND	ND	ND
Carbon *a* (Cα)	56	17.7 s ± 0.4	12 ± 4	1.6 s ± 0.03	19 ± 3.3
Solvent (DMSO)	39.5	15.9 s ± 0.8	ND	1.1 s ± 0.1	27 ± 16
Lipid	34.5	15 s ± 0.6	ND	1.4 s ± 0.2	12 ± 1.2
Aliphatic	28	16.8 s ± 0.5	9 ± 2	1.5 s ± 0.2	20 ± 10
Number of Samples		2		3	

(after analysis). Of most relevance to in-cell DNP is damage done prior to or during freezing as this establishes the conformational landscape of the sample at the time of analysis. Our study of pre-freeze viability demonstrates 30% glycerol (mimicking the use of DNP matrix for sample preparation) to be highly detrimental to cell viability and membrane integrity. Cells became extremely dehydrated and apoptotic within 10 min of glycerol exposure and thus significant damage occurs prior to freezing. Under the conditions tested here, glycerol preservation significantly alters the molecular state of cells analyzed by DNP. We detected apoptosis by Annexin-V staining in which fluorophore labelled Annexin-V binds to phosphotidylserine on the outer leaflet of the plasma membrane (Lakshmanan and Batra 2013). Phosphotidylserine does not localize to the outer leaflet of healthy mammalian cells but begins to accumulate in the outer leaflet of apoptotic cells. The induction of apoptosis and alterations to the lipid composition of the plasma membrane should be an important consideration when choosing a DNP preparation method. This is a crucial consideration for the study of proteins with charge-charge interactions or lipid interactions, as this may create a population of proteins whose interactions are biologically irrelevant or only relevant to apoptotic processes.

Following freezing, we found survivability to be poor in flash-frozen cells. Viability was largely uncorrelated with the presence of cryopreserving reagents under flash-freezing methods. This clearly indicates a lack of vitrification of flash-frozen 3.2 mm rotors. Cooling rates of 10°C/s have been measured for liquid nitrogen flash-frozen 3.2 mm rotors (Hu et al., 2009). Estimated cooling rates required to obtain perfect vitrification are in excess of 10,000°C/s depending on the concentration of cryoprotectant and as high as 10^{6}°C/s for pure water (Berejnov et al., 2006), which is difficult to achieve with a 3.2 mm rotor of ~22 µl. High concentrations of protein or cryoprotectant can increase the glass transition temperature, which reduces the cooling rate required to achieve vitrification. Intracellular ice crystal growth is thought to be the greatest contributor to cell death by cryogenic damage and which occurs above the glass transition temperature (Mazur 2004). Studies using differential scanning calorimetry (DSC) suggest that liquid nitrogen freezing of samples above the intracellular glass transition results in poor cell survival, as was also observed in our study (Meneghel et al., 2019). The authors provide evidence for an intracellular transition of around −50°C (Meneghel et al., 2019). Thus, vitrification of the intracellular compartment requires fast cooling rates down to −50°C. On the other hand, significantly improved post-thaw cell viability was achieved with slow-freezing at the typical 1°C/min but only for DMSO preserved cells. Glycerol preserved cells did not survive slow or flash-freezing with recovery of viable cells comparable to cells frozen in the absence of cryoprotectant. Presumably, the poor viability of glycerol preserved cells in this case is due to the

dehydrating and apoptotic effects prior to cells entering the frozen state. Thus, for increased viability, in-cell DNP samples require slow-cooling methods and DMSO as cryoprotectant.

But is cell viability correlated with structural relevance? DMSO is known to have effects on membrane phase behavior and hydration, having been demonstrated to out complete water for hydrogen bonding of lipid head-groups, at least in model membrane systems (Schrader et al., 2015). Furthermore, room temperature studies of model membrane systems suggest that DMSO improves protein conformational homogeneity for structural studies, however this also implies non-physiological conformational selection (Liao et al., 2016) and could be due to the effects of DMSO on lipid chain melting temperature and interfacial hydration. Thus, while DMSO greatly improves cell viability, it may not necessarily lead to physiological structures. On the other hand, glycerol has not been documented to affect the biophysical properties of model membrane systems. Despite this, we observed compositional changes of the plasma membrane in the presence of glycerol in addition to severe dehydration effects. Dehydration is problematic for protein structural studies as the loss of water can greatly affect protein conformation due to the critical role played by solvent water in driving protein folding (Dill 1990; Wolkers et al., 2019). The dysregulation of electrolyte and small molecule concentrations should also be considered when performing in-cell structural studies and the effects of dehydration on ligand interactions and charge-charge interactions. These effects would be particularly problematic for drug interaction studies. Cellular dehydration is also significant in cells frozen slowly (Meneghel et al., 2019) as described earlier. This gives rise to freeze concentration of intracellular solutes and has been shown to enhance ROS production (Len et al., 2019), and activation of apoptotic pathways during the freezing process until the intracellular glass transition temperature is reached (Bissoyi and Pramanik 2014). These effects could be minimized with vitrification.

Vitrification may be achievable in smaller rotors such as 1.3 mm (2.5 µl volume) or 0.7 mm (0.56 µl volume) outer diameter rotors, particularly as higher spinning frequencies can now be reached under DNP conditions (Chaudhari et al., 2016; Berruyer et al., 2020). Given the widespread use of pulled straws of 1.7–0.8 mm in the vitrification of human embryonic stem cells and blastocytes in liquid nitrogen (with excellent post-thaw viability) this appears eminently possible (Chen et al., 2001). Thus, for high resolution structural studies, slow-cooling or cryoprotectants may not be suitable for the preservation of native membrane structures or transiently assembled signaling complexes for DNP analysis, despite improved viability post-DNP. It may be that like *in situ* cryo-EM, where the use of vitrification methods that obtain higher resolution images are favored over higher viability methods, so to structural studies by in-cell DNP-NMR might also have to balance the need for resolution and structural relevance with viability. Therefore, given the effects of cryopreserving reagents and slow-freezing on molecular structure and composition documented in this study and by others, it seems prudent to develop methods towards sample vitrification and cryoprotecting reagent free in-cell DNP-NMR.

We also observed distinctions between cryoprotectants in DNP parameters. Enhancements per √unit time were only slightly better in glycerol protected cells compared to DMSO preservation. This is attributable to the longer polarization build up time of glycerol preserved cells, despite exhibiting a 2.6-fold larger $\varepsilon_{on/off}$ compared to DMSO prepared cells. Shorter T_B times with d_6-DMSO is an interesting observation and it is unclear whether additional cell specific factors are contributing to such a short T_B time or whether the distribution of DMSO compared to glycerol influences this value or its effect on ice formation. The use of d_6-DMSO precludes the contribution of methyl mediated cross-relaxation to the short T_B time. The effect of fast polarization build up with lower enhancement factors in the presence of DMSO has also been reported in model membrane systems (Liao et al., 2016) and so it is unclear whether this is a property of AMUPol-DMSO interactions or a specific property of d_6-DMSO glasses. The variation in enhancement factors of glycerol mixtures <55% v/v in model systems are thought to be due to increased water-water interactions, presumably manifesting as increased ice formation (Leavesley et al., 2018b). Ice crystalization is also thought to increase paramagnetic quenching of nuclei by freeze concentrating DNP radicals outside of growing ice crystals which are extracting water from the surrounding space to feed crystal growth (Corzilius et al., 2014). The resulting increase in e-e-n polarization transfer at higher radical concentrations leads to reduced enhancement factors and shorter polarization build up times (Leavesley et al., 2018a). While the lower T_B time of DMSO preserved samples would be consistent with increased radical concentration, possibly due to ice formation, we would expect this phenomenon to be greater in PBS preserved cells in which ice crystal formation would be expected to be greatest. Since polarization build up times in PBS treated samples were consistently longer than that observed in DMSO preserved cells it suggests ice crystalization is not driving this phenomenon. The greatly increased T_B time of slow-frozen cells suggest that longer T_B times are reflective of lower cellular radical concentrations. While differences in ice formation can-not be ruled out, it seems much more likely that slow-frozen cells have lower AMUPol concentrations due to reduction of AMUPol by the intracellular environment and/or improved distribution of AMUPol throughout the cells. In line with this hypothesis, net enhancements were generally lower in slow-frozen cells, suggestive of reduced radical concentration, although the measured values were too variable to make a convincing conclusion and requires further investigation. Consistent with this hypothesis, solvent signals consistently exhibited shorter T_B times and higher enhancements in flash-frozen samples, which could be explained by a higher concentration of AMUPol and DMSO outside the cell compared to inside the cell. Although this

does not necessarily provide a clear demarcation between intracellular and extracellular signals. One might expect that T_B times of cellular signals to exhibit a wider range in values if this was the case. However, the ambiguity of intracellular verses extracellular signals makes such an assertion difficult. The T_B of extracellular glyco-proteins and plasma membrane signals could account for much of the cellular signal observed at earlier time points and longer relaxing intracellular signals may be hidden within the bulk build up curves.

CONCLUSION

The use of glycerol and its perturbation of cellular integrity and apoptosis induction makes glycerol not optimal for in-cell DNP. In addition, the enhancements per $\sqrt{}$unit time are only marginally greater than those achieved through using d_6-DMSO. Overall, we greatly favor the use of DMSO. A slow-freeze (1°C/min) protocol for the preparation of Jurkat T cells for DNP analysis is preferable when viability is a priority. However, the optimal cooling rate is likely dependent on cell type and intracellular localization of molecules of interest. Furthermore, the impact of cooling rates and cryoprotectants on protein/membrane structures remains to be determined. It is likely that the benefits of slow-freeze protocols will be diminished with smaller rotors if sufficiently fast cooling rates can be achieved with flash-freezing to reduce cellular dehydration through improved vitrification, potentially in the absence of cryoprotecting agents. These challenges highlight the need for ongoing studies in this area and the investigation of vitrification methods for in-cell DNP.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

SO and AB designed the study. SO conducted experiments and data analysis. SO and AB wrote the article.

FUNDING

This work was funded by internal ETH funding.

ACKNOWLEDGMENTS

We thank Shannon Rüeger for assistance with maintenance of Jurkat T cell cultures. We also thank Prof. Christophe Copéret for use of the 600 MHz DNP system, Alexander Yakimov and Zach Berkeson for assistance with this system. We also thank Dr. Nicholas Alaniva and Dr. Snædis Björgvinsdóttir for helpful critical feedback on the article.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.743829/full#supplementary-material

REFERENCES

Albert, B. J., Gao, C., Sesti, E. L., Saliba, E. P., Alaniva, N., Scott, F. J., et al. (2019). Dynamic Nuclear Polarization Nuclear Magnetic Resonance in Human Cells Using Fluorescent Polarizing Agents. *Biochemistry* 57 (31), 4741–4746. doi:10.1021/acs.biochem.8b00257.Dynamic

Banci, L., Barbieri, L. I., Bertini, I., Luchinat, E., Secci, E., Zhao, Y., et al. (2013). Atomic-Resolution Monitoring of Protein Maturation in Live Human Cells by NMR. *Nat. Chem. Biol.* 9, 297–299. doi:10.1038/nchembio.1202

Baust, J. G., Gao, D., and Baust, J. M. (2009). Cryopreservation. *Organogenesis* 5 (3), 90–96. doi:10.4161/org.5.3.10021

Berejnov, V., Husseini, N. S., Alsaied, O. A., and Thorne, R. E. (2006). Effects of Cryoprotectant Concentration and Cooling Rate on Vitrification of Aqueous Solutions. *J. Appl. Cryst.* 39, 244–251. doi:10.1107/S0021889806004717

Berruyer, P., Björgvinsdóttir, S., Bertarello, A., Stevanato, G., Rao, Y., Karthikeyan, G., et al. (2020). Dynamic Nuclear Polarization Enhancement of 200 at 21.15 T Enabled by 65 kHz Magic Angle Spinning. *J. Phys. Chem. Lett.* 11 (19), 8386–8391. doi:10.1021/acs.jpclett.0c02493

Bissoyi, A., and Pramanik, K. (2014). Role of the Apoptosis Pathway in Cryopreservation-Induced Cell Death in Mesenchymal Stem Cells Derived from Umbilical Cord Blood. *Biopreservation and Biobanking* 12 (4), 246–254. doi:10.1089/bio.2014.0005

Burmann, B. M., Gerez, J. A., Matečko-Burmann, I., Campioni, S., Kumari, P., Ghosh, D., et al. (2020). Regulation of α-synuclein by Chaperones in Mammalian cellsAlpha-Synuclein Regulation by Chaperones in Mammalian Cells. *Nature* 577 (7788), 127–132. doi:10.1038/s41586-019-1808-9

Chaudhari, S. R., Berruyer, P., Gajan, D., Reiter, C., Engelke, F., Silverio, D. L., et al. (2016). Dynamic Nuclear Polarization at 40 KHz Magic Angle Spinning. *Phys. Chem. Chem. Phys.* 18 (15), 10616–10622. doi:10.1039/c6cp00839a

Chen, S.-U., LienCheng, Y.-R., Cheng, Y.-Y., Chen, H.-F., Ho, H.-N., and Yang, Y.-S. (2001). Vitrification of Mouse Oocytes Using Closed Pulled Straws (CPS) Achieves a High Survival and Preserves Good Patterns of Meiotic Spindles, Compared with Conventional Straws, Open Pulled Straws (OPS) and Grids. *Hum. Reprod.* 16 (11), 2350–2356. doi:10.1093/humrep/16.11.2350

Corzilius, B., Andreas, L. B., Smith, A. A., Ni, Q. Z., and Griffin, R. G. (2014). Paramagnet Induced Signal Quenching in MAS-DNP Experiments in Frozen Homogeneous Solutions. *J. Magn. Reson.* 240, 113–123. doi:10.1016/j.jmr.2013.11.013

Dill, K. A. (1990). Dominant Forces in Protein Folding. *Biochemistry* 29 (31), 7133–7155. doi:10.1021/bi00483a001

Fahy, G. M., and Wowk, B. (2020). Principles of Ice-free Cryopreservation by Vitrification. *Methods Mol. Biol.* 2180, 27–97. doi:10.1007/978-1-0716-0783-1_2

Giotta, G. J., and Wang, H. H. (1972). Reduction of Nitroxide Free Radicals by Biological Materials. *Biochem. Biophysical Res. Commun.* 46 (4), 1576–1580. doi:10.1016/0006-291x(72)90788-7

Hall, D. A., Maus, D. C., Gerfen, G. J., Inati, S., Becerra, L., Dahlquist, F. W., et al. (1997). Polarization-Enhanced NMR Spectroscopy of Biomolecules in Frozen Solution. *Science* 276 (5314), 930–932. doi:10.1126/science.276.5314.930

Hornberger, K., Yu, G., McKenna, D., and Hubel, A. (2019). Cryopreservation of Hematopoietic Stem Cells: Emerging Assays, Cryoprotectant Agents, and Technology to Improve Outcomes. *Transfus. Med. Hemother* 46, 188–196. doi:10.1159/000496068

Hu, K.-N., Havlin, R. H., Yau, W.-M., and Tycko, R. (2009). Quantitative Determination of Site-specific Conformational Distributions in an Unfolded

Protein by Solid-State Nuclear Magnetic Resonance. *J. Mol. Biol.* 392, 1055–1073. doi:10.1016/j.jmb.2009.07.073

Huebinger, J., Han, H.-M., Hofnagel, O., Vetter, I. R., Bastiaens, P. I. H., and Grabenbauer, M. (2016). Direct Measurement of Water States in Cryopreserved Cells Reveals Tolerance toward Ice Crystallization. *Biophysical J.* 110, 840–849. doi:10.1016/j.bpj.2015.09.029

Huebinger, J. (2018). Modification of Cellular Membranes Conveys Cryoprotection to Cells during Rapid, Non-equilibrium Cryopreservation. *PLOS One* 13 (10), e0205520. doi:10.1371/journal.pone.0205520

Jordan, A., Defechereux, P., and Verdin, E. (2001). The Site of HIV-1 Integration in the Human Genome Determines Basal Transcriptional Activity and Response to Tat Transactivation. *EMBO J.* 20 (7), 1726–1738. doi:10.1093/emboj/20.7.1726

Kroemer, G., Galluzzi, L., Vandenabeele, P., Abrams, J., Alnemri, E. S., Baehrecke, E. H., et al. (2009). Classification of Cell Death: Recommendations of the Nomenclature Committee on Cell Death 2009. *Cell Death Differ* 16 (1), 3–11. doi:10.1038/cdd.2008.150

Lakshmanan, I., and Batra, S. K. (2013). Protocol for Apoptosis Assay by Flow Cytometry Using Annexin V Staining Method. *Bio-protocol* 3 (6), e374. doi:10.21769/bioprotoc.374

Leavesley, A., Jain, S., Kaminker, I., Zhang, H., Rajca, S., Rajca, A., et al. (2018a). Maximizing NMR Signal Per Unit Time by Facilitating the E-E-N Cross Effect DNP Rate. *Phys. Chem. Chem. Phys.* 20 (43), 27646–27657. doi:10.1039/c8cp04909b

Leavesley, A., Wilson, C. B., Sherwin, M., and Han, S. (2018b). Effect of Water/glycerol Polymorphism on Dynamic Nuclear Polarization. *Phys. Chem. Chem. Phys.* 20, 9897–9903. doi:10.1039/c8cp00358k

Len, J. S., Koh, W. S. D., and Tan, S.-X. (2019). The Roles of Reactive Oxygen Species and Antioxidants in Cryopreservation. *Biosci. Rep.* 39 (8), BSR20191601. doi:10.1042/BSR20191601

Liao, S. Y., Lee, M., Wang, T., Sergeyev, I. V., and Hong, M. (2016). Efficient DNP NMR of Membrane Proteins: Sample Preparation Protocols, Sensitivity, and Radical Location. *J. Biomol. NMR* 64 (3), 223–237. doi:10.1007/s10858-016-0023-3

Luchinat, E., Barbieri, L., Cremonini, M., Nocentini, A., Supuran, C. T., Banci, L., et al. (2020). Intracellular Binding/Unbinding Kinetics of Approved Drugs to Carbonic Anhydrase II Observed by In-Cell NMR. *ACS Chem. Biol.* 15 (10), 2792–2800. doi:10.1021/acschembio.0c00590

Martorana, A., Bellapadrona, G., Feintuch, A., Di Gregorio, E., Aime, S., and Goldfarb, D. (2014). Probing Protein Conformation in Cells by EPR Distance Measurements Using Gd3+ Spin Labeling. *J. Am. Chem. Soc.* 136, 13458–13465. doi:10.1021/ja5079392

Mazur, P. (1984). Freezing of Living Cells: Mechanisms and Implications. *Am. J. Physiology-Cell Physiol.* 247 (3 Pt 1), C125–C142. doi:10.1152/ajpcell.1984.247.3.C125

Mazur, P. (2004). "Principles of Cryobiology," in *Life In the Frozen State*. Editors B. J Fuller, N. Lane, and E. Benson (Boca Raton, FL: CRC Press), 3–65. doi:10.1201/9780203647073.ch1

McCoy, K. M., Rogawski, R., Stovicek, O., and McDermott, A. E. (2019). Stability of Nitroxide Biradical TOTAPOL in Biological Samples. *J. Magn. Reson.* 303, 115–120. doi:10.1016/j.jmr.2019.04.013

Meneghel, J., Kilbride, P., Morris, J. G., and Fonseca, F. (2019). Physical Events Occurring during the Cryopreservation of Immortalized Human T Cells. *PLOS One* 14, e0217304. doi:10.1371/journal.pone.0217304

Mentink-Vigier, F., Vega, S., and De Paëpe, G. (2017). Fast and Accurate MAS-DNP Simulations of Large Spin Ensembles. *Phys. Chem. Chem. Phys.* 19, 3506–3522. doi:10.1039/c6cp07881h

Müntener, T., Häussinger, D., Selenko, P., and Theillet, F.-X. (2016). In-Cell Protein Structures from 2D NMR Experiments. *J. Phys. Chem. Lett.* 7, 2821–2825. doi:10.1021/acs.jpclett.6b01074

Narasimhan, S., Scherpe, S., Lucini Paioni, A., van der Zwan, J., Folkers, G. E., Ovaa, H., et al. (2019). DNP-Supported Solid-State NMR Spectroscopy of Proteins inside Mammalian Cells. *Angew. Chem. Int. Ed.* 58, 12969–12973. doi:10.1002/anie.201903246

Overall, S. A., Price, L. E., Albert, B. J., Gao, C., Alaniva, N., Judge, P. T., et al. (2020). *In Situ* Detection of Endogenous HIV Activation by Dynamic Nuclear Polarization NMR and Flow Cytometry. *Ijms* 21 (13), 4649. doi:10.3390/ijms21134649

Overall, S. A., Zhu, S., Hanssen, E., Separovic, F., and Sani, M.-A. (2019). *In Situ* Monitoring of Bacteria under Antimicrobial Stress Using 31P Solid-State NMR. *Ijms* 20 (1), 181–212. doi:10.3390/ijms20010181

Pan, B.-B., Yang, F. Q. Wu., Ye, Y., Wu, Q., Li, C., Huber, T., et al. (2016). 3D Structure Determination of a Protein in Living Cells Using Paramagnetic NMR Spectroscopy. *Chem. Commun.* 52, 10237–10240. doi:10.1039/C6CC05490K

Raju, R., Bryant, S. J., Wilkinson, B. L., and Bryant, Gary. (2021). The Need for Novel Cryoprotectants and Cryopreservation Protocols: Insights into the Importance of Biophysical Investigation and Cell Permeability. *Biochim. Biophys. Acta (Bba) - Gen. Subjects* 1865 (1), 129749. doi:10.1016/j.bbagen.2020.129749

Rezazadeh Valojerdi, M., Eftekhari-Yazdi, P., Karimian, L., Hassani, F., and Movaghar, B. (2009). Vitrification versus Slow Freezing Gives Excellent Survival, Post Warming Embryo Morphology and Pregnancy Outcomes for Human Cleaved Embryos. *J. Assist. Reprod. Genet.* 26 (6), 347–354. doi:10.1007/s10815-009-9318-6

Sakakibara, D., Sasaki, A., Ikeya, T., Hamatsu, J., Hanashima, T., Mishima, M., et al. (2009). Protein Structure Determination in Living Cells by In-Cell NMR Spectroscopy. *Nature* 458, 102–105. doi:10.1038/nature07814

Sauvée, C., Rosay, M., Casano, G., Aussenac, F., Weber, R. T., Ouari, O., et al. (2013). Highly Efficient, Water-Soluble Polarizing Agents for Dynamic Nuclear Polarization at High Frequency. *Angew. Chem. Int. Ed.* 52 (41), 10858–10861. doi:10.1002/anie.201304657

Schlagnitweit, J., Friebe Sandoz, S., Jaworski, A., Guzzetti, I., Aussenac, F., Carbajo, R. J., et al. (2019). Observing an Antisense Drug Complex in Intact Human Cells by in-Cell NMR Spectroscopy. *ChemBioChem* 20, 2474–2478. doi:10.1002/cbic.201900297

Schrader, A. M., Cheng, C. Y., Israelachvili, J. N., and Han, S. (2016). Communication: Contrasting Effects of Glycerol and DMSO on Lipid Membrane Surface Hydration Dynamics and Forces. *J. Chem. Phys.* 145 (4), 041101. doi:10.1063/1.4959904

Schrader, A. M., Donaldson, S. H., Song, J., Cheng, C.-Y., Lee, D. W., Han, S., et al. (2015). Correlating Steric Hydration Forces with Water Dynamics through Surface Force and Diffusion NMR Measurements in a Lipid-DMSO-H2o System. *Proc. Natl. Acad. Sci. USA* 112 (34), 10708–10713. doi:10.1073/pnas.1512325112

Sekhon, L., Lee, J. A., Flisser, E., Copperman, A. B., and Stein, Daniel. (2018). Blastocyst Vitrification, Cryostorage and Warming Does Not Affect Live Birth Rate, Infant Birth Weight or Timing of Delivery. *Reprod. BioMedicine Online* 37 (1), 33–42. doi:10.106/j.rbmo.2018.03.02310.1016/j.rbmo.2018.03.023

Selenko, P., Serber, Z., Gadea, B., Ruderman, J., and Wagner, G. (2006). Quantitative NMR Analysis of the Protein G B1 Domain in *Xenopus laevis* Egg Extracts and Intact Oocytes. *Proc. Natl. Acad. Sci.* 103 (32), 11904–11909. doi:10.1073/pnas.0604667103

Shi, R., Qiao, X., Emerson, N., and Malcom, A. (2001). Dimethylsulfoxide Enhances CNS Neuronal Plasma Membrane Resealing after Injury in Low Temperature of Low Calcium. *J. Neurocytol* 30 (9–10), 829–839. doi:10.1023/a:1019645505848

Siegel, B. V., Lund, R. O., Wellings, S. R., and Bostic, W. L. (1960). Growth of HeLa Cells in Heavy Water. *Exp. Cell Res.* 19 (1), 187–189. doi:10.1016/0014-4827(60)90050-1

Theillet, F.-X., Binolfi, A., Bekei, B., Martorana, A., Rose, H. M., Stuiver, M., et al. (2016). Structural Disorder of Monomeric α-synuclein Persists in Mammalian Cells. *Nature* 530, 45–50. doi:10.1038/nature16531

Thongsomboon, W., Serra, D. O., Possling, A., Hadjineophytou, C., Hengge, R., Cegelski, L., et al. (2018). Phosphoethanolamine Cellulose: A Naturally Produced Chemically Modified Cellulose. *Science* 359 (6373), 334–338. doi:10.1126/science.aao4096

Wolkers, W. F., Oldenhof, H., Tang, F., Han, J., Bigalk, J., and Sieme, H. (2019). Jiale Han, Judith Bigalk, and Harald Sieme.Factors Affecting the Membrane Permeability Barrier Function of Cells during Preservation Technologies. *Langmuir* 35 (23), 7520–7528. doi:10.1021/acs.langmuir.8b02852

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

REVIEW
published: 25 October 2021
doi: 10.3389/fmolb.2021.727553

Model-Free or Not?

*Kai Zumpfe and Albert A. Smith**

Institute for Medical Physics and Biophysics, Medical Faculty, Leipzig University, Leipzig, Germany

Relaxation in nuclear magnetic resonance is a powerful method for obtaining spatially resolved, timescale-specific dynamics information about molecular systems. However, dynamics in biomolecular systems are generally too complex to be fully characterized based on NMR data alone. This is a familiar problem, addressed by the Lipari-Szabo model-free analysis, a method that captures the full information content of NMR relaxation data in case all internal motion of a molecule in solution is sufficiently fast. We investigate model-free analysis, as well as several other approaches, and find that model-free, spectral density mapping, LeMaster's approach, and our *detector* analysis form a class of analysis methods, for which behavior of the fitted parameters has a well-defined relationship to the distribution of correlation times of motion, independent of the specific form of that distribution. In a sense, they are all "model-free." Of these methods, only detectors are generally applicable to solid-state NMR relaxation data. We further discuss how detectors may be used for comparison of experimental data to data extracted from molecular dynamics simulation, and how simulation may be used to extract details of the dynamics that are not accessible via NMR, where detector analysis can be used to connect those details to experiments. We expect that combined methodology can eventually provide enough insight into complex dynamics to provide highly accurate models of motion, thus lending deeper insight into the nature of biomolecular dynamics.

Keywords: solid-state NMR, dynamics detectors, model-free analysis, NMR relaxation, molecular dynamics simulation

OPEN ACCESS

Edited by:
Józef Romuald Lewandowski, University of Warwick, United Kingdom

Reviewed by:
Tharin Blumenschein, University of East Anglia, United Kingdom
Paul Schanda, Institute of Science and Technology Austria, Austria

***Correspondence:**
Albert A. Smith
albert.smith-penzel@medizin.uni-leipzig.de

Specialty section:
This article was submitted to Structural Biology, a section of the journal Frontiers in Molecular Biosciences

Received: *18 June 2021*
Accepted: *21 September 2021*
Published: *25 October 2021*

Citation:
Zumpfe K and Smith AA (2021) Model-Free or Not? Front. Mol. Biosci. 8:727553. doi: 10.3389/fmolb.2021.727553

INTRODUCTION

Study of biomolecular function requires understanding the dynamics of the biological system. Nuclear magnetic resonance (NMR), despite many recent technological advances in other techniques, remains a premier method for detailed dynamics characterization. In NMR, one may measure a variety of site-specific relaxation experiments, which provide timescale sensitive information about the motion. By varying the type of experiment (T_1, $T_{1\rho}$, NOE, etc.) or experimental conditions (external magnetic field, applied field strength, magic-angle spinning (MAS) frequency, etc.), the timescale sensitivity of the measurement is modified. Then, one may resolve the dynamics both in space, via site resolution, and in timescale, via multiple experiments (Palmer, 2004; Schanda and Ernst, 2016).

However, is it possible to fully characterize the motions leading to the observed relaxation behavior? Many relaxation experiments in NMR are sensitive to the reorientational motion of anisotropic NMR interaction tensors (NMR relaxation can also be sensitive to change in scalar terms, e.g., isotropic chemical shift). For a given spin, relaxation is usually dominated by only one to two interactions. For example, relaxation of ^{15}N in a protein backbone is determined almost entirely by the reorientation of the one-bond ^{1}H–^{15}N dipole coupling and the ^{15}N chemical shift anisotropy (CSA). But, multiple sources of motion lead to reorientation of the bond. For example, if we suppose

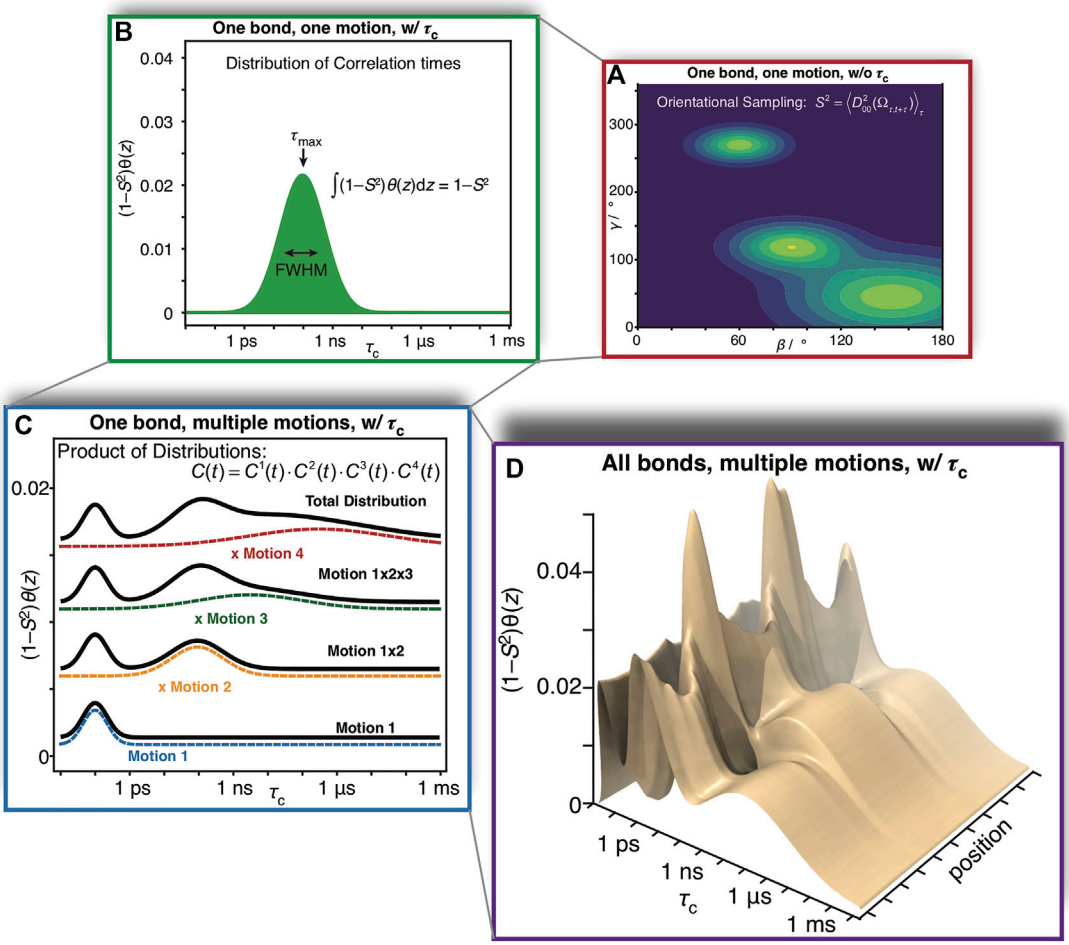


FIGURE 1 | Complexity of reorientational dynamics. For each bond in a molecule, multiple types of motion result in orientational sampling, where the distribution of angles for each motion result in a generalized order parameter, S^2. Therefore, in **(A)** we plot a possible distribution of Euler angles for a single type of motion (population is plotted as a function of angles β and γ, where α is not required for a symmetric interaction tensor). A single motion is furthermore described by a correlation time, and may be distributed over a range of correlation times. In **(B)** we plot a possible distribution of correlation times $(1-S^2)\theta(z)$, that is, amplitude of motion as a function of the log-correlation time, $z = \log_{10}(\tau_c/\mathrm{s})$. Each distribution is characterized by an amplitude, center, and width. Note that the integral of the distribution is $(1-S^2)$, S^2 being determined by the distribution of angles in **(A)**. While **(A,B)** illustrate aspects of a single motion, multiple motions influence a given bond, where the total correlation function is the product of individual correlation functions. In **(C)**, we plot four distributions of motion (color). Above each motion, we plot the distribution resulting from the product of that motion and all motions below it (black), eventually resulting in the total distribution seen at the top. Finally, we note that the total distribution varies as a function of position in the molecule, resulting in the 3D plot of the distribution as a function of correlation time and position in the molecule observed in **(D)**. While this is just an illustration, one could imagine that motion in **(D)** results from three α-helices in a protein, each having a slightly different behavior, and varying dynamics as one approaches the end of each helix.

the H–N bond to be in a protein, within a helix, then we would have local distortion of the peptide plane (one-bond libration), motion of the peptide plane within the helix, motion of the helix within the protein, and motion of the protein either in solution, in a crystal, a fibril, a membrane, etc.

This degree of complexity is illustrated in **Figure 1**. For a given bond in a molecule, and a given motion acting on that bond, a distribution of orientations is sampled as illustrated in **Figure 1A**. The orientational distribution determines the contribution of that motion to the total order parameter, S^2. However, not only are there many orientations sampled by a bond due to a motion, but those orientations are sampled at some rate, such that the motion has an associated correlation time or distribution of correlation times (denoted $(1-S^2)\theta(z)$). We illustrate this in **Figure 1B**; note that not only is the width of such a distribution variable, but also the functional form of the distribution itself. This results in a correlation function that decays from 1 to S^2, where integrating over the distribution of correlation times yields the total amplitude of the decay. Already, a single bond with just one motion acting on it yields potentially a high degree of complexity; however, we must still consider that multiple motions act on each bond, where the total correlation function is the product of the correlation functions of each individual motion (if those motions are independent from one another, **Figure 1C**). Finally, motion varies throughout a

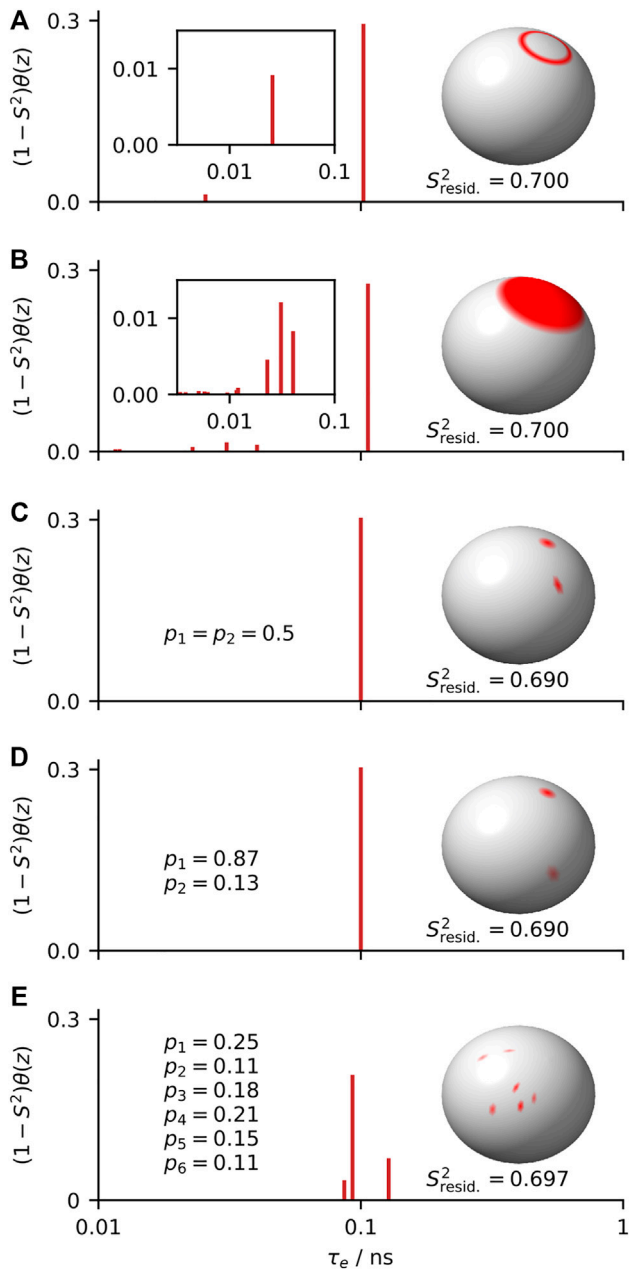


FIGURE 2 | Five distributions of orientations and correlation times that yield the same model-free parameters ($(1-S^2) = 0.3$, $\langle\tau_e\rangle = 0.1$ ns). In **(A–E)**, we plot a distribution of orientations (sphere, right); on the axes, we plot the distribution of correlation times resulting from exchange among that set of orientations. Models of motion are wobbling-on-a-cone ($\theta_{\text{cone}} = 19^\circ$), wobbling-in-a-cone ($\theta_{\text{cone}} = 28^\circ$), symmetric two-site hop ($\theta_{\text{hop}} = 39^\circ$), asymmetric two-site hop ($\theta_{\text{hop}} = 70^\circ$), and 6-site asymmetric exchange. Insets in **(A,B)** show correlation times with small amplitudes. $S^2_{\text{resid.}}$ refers to the order parameter from residual couplings (see **Supplementary Section S3**), which deviates from the generalized order parameter for asymmetric motion.

molecule, as a function of position, resulting in a complex, multi-dimensional description as illustrated in **Figure 1D**.

While NMR is powerful, obtaining a complete description of the complex dynamics stretches beyond the limit of what is possible based on experimental data alone, especially for large molecules such as proteins. This problem is a familiar one, addressed almost 40 years ago by Lipari and Szabo (Lipari and Szabo, 1982a), who developed a method known as the *model-free* approach. While we will discuss the details of this approach below, the name tells us a critical advantage of such an approach: model-free analysis allows the extraction of dynamics information from NMR relaxation data *without having knowledge of the specific model of motion.* Furthermore, the resulting parameters have a well-defined relationship to the distribution of orientations sampled and the distribution of correlation times.

Lipari and Szabo described the internal motion of a molecule with just two parameters: a generalized order parameter related to the amplitude of motion, S^2, and a mean effective correlation time, $\langle\tau_e\rangle$ (a third parameter, τ_M, gives the correlation time of the molecule tumbling in solution). While only two parameters suggests a simple analysis, it is important to note that Lipari and Szabo did not intend to only describe simple motions having just a single correlation time and amplitude: theoretical tests of their model were performed on a wobbling-on-a-cone model (Kinosita et al., 1977) that results in a weighted sum of correlation times, and experimental work was performed on methyl groups in a protein, for which the total motion is determined by the product of methyl rotation and by reorientation of the methyl group's C–C bond. Rather, the two parameters contain the aggregated information describing all motions that is available from the set of relaxation experiments alone.

The advantage of model-free analysis is that it does not require knowing the model of motion. For example, for relatively low fields (~90 MHz, as used by Lipari and Szabo), all distributions of orientations and correlation times shown in **Figure 2** should yield identical relaxation rate constants for the set of experiments. If we do not know which model is the correct model, the best we can do is to parameterize the results in a way that does not depend on the model of motion, as can be done with the model-free parameters S^2 and $\langle\tau_e\rangle$.

When analyzing data, a model provides a framework for understanding the data, and by using a model we are always adding some information to the experimental data. In some cases, we add further information depending on how we interpret a model. A model is advantageous if that information is correct, and disadvantageous if that information is wrong. Suppose, for example, we know that the correct model in **Figure 2** is a symmetric two-site hop, shown in **Figure 2C**; then we may extract the hop angle and exchange rates from S^2 and $\langle\tau_e\rangle$, resulting in $\theta_{\text{hop}} = 39^\circ$ and $k_{\text{ex}}^{1\to2} = k_{\text{ex}}^{2\to1} = 5 \times 10^9$/s. However, if the true model is an asymmetric two-site hop, shown in **Figure 2D**, the true angle and exchange rates may be significantly different (for **Figure 2D**, these are $\theta_{\text{hop}} = 70^\circ$ with $k_{\text{ex}}^{1\to2} = 1.3 \times 10^9$/s and $k_{\text{ex}}^{2\to1} = 8.7 \times 10^9$/s). Then, a model-free approach is the more reliable method when the correct model cannot be independently determined.

In this review, we will first discuss the original model-free approach, and then examine methods descended from it, including discussion of our own detector analysis, a relatively new approach that also provides a model-free analysis in the spirit

of the original Lipari-Szabo approach, but can extract the full information content of relaxation data sets in instances where the model-free approach cannot. We discuss analysis of microsecond motions using $R_{1\rho}$ relaxation, and finally consider how other methods, in particular molecular dynamics (MD) simulation, may be used to supply the information that NMR lacks, thus improving the interpretation of NMR parameters.

MODEL-FREE

While dynamics analysis methods have existed for application to solid-state NMR for some years now (Chevelkov et al., 2009b; Schanda et al., 2010; Zinkevich et al., 2013; Lamley et al., 2015a; Smith et al., 2016; Lakomek et al., 2017; Kurauskas et al., 2017), most of the approaches applied have evolved from methodology first developed for solution-state NMR. Probably the most important advance in solution-state analysis was the development of the model-free approach (Lipari and Szabo, 1982a; Lipari and Szabo, 1982b), and related two-step techniques (Wennerström et al., 1979; Halle and Wennerström, 1981; Brown, 1982). Then, we begin by reviewing some of the existing methodology, to understand advantages and disadvantages to various approaches.

Model-Free Theory

Typical solution-state NMR data sets consist of relaxation rate constants for R_1 $(1/T_1)$, R_2 $(1/T_2)$, and nuclear Overhauser effect (NOE, σ_{IS}), acquired at one or more magnetic fields. The rate constants describe the signal decay $(I(t) = I_0 e^{-R_\zeta t})$ or recovery $(I(t) = I_{\mathrm{eq}} + (I_0 - I_{\mathrm{eq}})e^{-R_\zeta t})$. In solid-state NMR, this behavior can be multi-exponential, whereas we use the rate constant that describes the powder-averaged value (Krushelnitsky et al., 2018). Relaxation is often driven by reorientation of a few anisotropic interactions, for example, for backbone ^{15}N relaxation, a one-bond H–N dipole coupling and CSA are responsible for relaxation. For these experiments, the relaxation rate constants may be calculated from the spectral density, $J(\omega)$:

$$R_1^{\mathrm{I}} = \underbrace{\left(\frac{\delta^{\mathrm{IS}}}{4}\right)^2 (J(\omega_{\mathrm{I}} - \omega_{\mathrm{S}}) + 3J(\omega_{\mathrm{I}}) + 6J(\omega_{\mathrm{I}} + \omega_{\mathrm{S}}))}_{\text{dipolar relaxation}} + \underbrace{\frac{1}{3}(\omega_{\mathrm{I}}\Delta\sigma_{\mathrm{I}})^2 J(\omega_{\mathrm{I}})}_{\text{CSA relaxation}}$$

$$R_2^{\mathrm{I}} = \frac{1}{2}R_1^{\mathrm{I}} + \underbrace{\left(\frac{\delta^{\mathrm{IS}}}{4}\right)^2 (3J(\omega_{\mathrm{S}}) + 2J(0))}_{\text{dipolar relaxation}} + \underbrace{\frac{2}{9}(\omega_{\mathrm{I}}\Delta\sigma_{\mathrm{I}})^2 J(0)}_{\text{CSA relaxation}}$$

$$\sigma_{\mathrm{IS}} = \underbrace{\left(\frac{\delta^{\mathrm{IS}}}{4}\right)^2 (-J(\omega_{\mathrm{I}} - \omega_{\mathrm{S}}) + 6J(\omega_{\mathrm{I}} - \omega_{\mathrm{S}}))}_{\text{dipolar relaxation}} \tag{1}$$

Here, ω_{I} is the Larmor frequency (in radians/s) of the nucleus being relaxed, ω_{S} the Larmor frequency of the coupled spin (usually ^{1}H), and δ^{IS} and $\Delta\sigma_{\mathrm{I}}\omega_{\mathrm{I}}$ are the anisotropies of the dipolar coupling and CSA, respectively ($\delta^{\mathrm{IS}} = -2\frac{\mu_0}{4\pi}\frac{h\gamma_{\mathrm{I}}\gamma_{\mathrm{S}}}{r_{\mathrm{IS}}^2}$, with μ_0 the vacuum permeability in T^{2}m^3/J, $\gamma_{\mathrm{I}}, \gamma_{\mathrm{S}}$ the gyromagnetic ratios of the two spins in radians/s, h is Planck's constant in J·s, and r_{IS} the distance between the spins in meters, resulting in δ^{IS}, which is the full breadth of the dipolar powder pattern in radians/s. $\Delta\sigma_{\mathrm{I}}\omega_{\mathrm{I}}$ is similarly the full breadth $(\Delta\sigma_{\mathrm{I}} = \frac{3}{2}(\sigma_{zz} - \sigma_{\mathrm{iso}}))$ of the CSA powder pattern in radians/s when the Larmor frequency of spin I is given by ω_{I}, also in radians/s (Schanda and Ernst, 2016)). The spectral density may be obtained from the Fourier transform of the correlation function of motion. The correlation function itself is the rank-2 tensor correlation function, and describes the reorientational behavior of an NMR interaction tensor in time. If we assume the correlation function is symmetric in time, we may replace $e^{i\omega t}$ with $\cos(\omega t)$ in the Fourier transform. We can also change the integration bounds from $(-\infty, \infty)$ to $(0, \infty)$, and must multiply the integral by two in order to compensate for only integrating over half the space.

$$\begin{aligned} J(\omega) &= \int_{-\infty}^{\infty} C(t)e^{i\omega t}dt \\ &= \int_{-\infty}^{\infty} \underbrace{C(t)\cos(\omega t)}_{\text{symmetric in time}} dt + i\int_{-\infty}^{\infty} \underbrace{C(t)\sin(\omega t)}_{\text{antisymmetric in time}\to 0} dt \\ J(\omega) &= 2\int_0^{\infty} C(t)\cos(\omega t)dt \end{aligned} \tag{2}$$

Then, model-free analysis makes a few assumptions about the correlation function:

1) The total motion of a given bond is the result of overall tumbling of the molecule in solution and internal motion of the bond within the molecule, and these two motions are statistically independent.
2) Decay of the correlation function due to internal motion is fast compared to all ω sampled by the set of experimental relaxation rate constants (i.e., the extreme narrowing limit).

The decay of the correlation due to internal motion does not need to be mono-exponential (or even multi-exponential, although we will later apply this assumption). Instead of the second assumption, we may assume that the correlation function due to internal motion is mono-exponential, in which case we do not require its decay to be fast (we will visit this case only briefly, as it is less likely to occur in practice). We also assume tumbling is isotropic, although this is not necessarily required. Note that separate methods exist in case overall tumbling and internal motion are coupled (Tugarinov et al., 2001), although we will not consider these here. As a set of equations, this yields

$$\begin{aligned} C(t) &= C^{\mathrm{intern.}}(t) \cdot C^{\mathrm{rot.}}(t) \\ C^{\mathrm{rot.}}(t) &= \frac{1}{5}e^{-t/\tau_{\mathrm{M}}} \\ C^{\mathrm{intern.}}(t) &= S^2 + (1 - S^2)G(t) \\ G(0) &= 1, \quad \lim_{t\to\infty} G(t) = 0 \end{aligned} \tag{3}$$

The first equation is the result of statistical independence of internal and overall motion, such that we may write the total correlation function, $C(t)$, as a product of a correlation function resulting from the internal motion ($C^{\text{intern.}}(t)$), and a correlation function resulting from the overall rotational tumbling ($C^{\text{rot.}}(t)$). The overall motion may be described by a single decaying exponential, with correlation time τ_{M} if that overall motion is isotropic (occurring if the molecule is approximately spherical). For internal motion, $C^{\text{intern.}}(t)$ has an initial value of 1, and equilibrates at S^2. S^2 is referred to as the generalized order parameter, and is related to, but not always equal to order parameters that may be extracted from measurement of residual couplings, as will be discussed in *Determining* S^2. $G(t)$ is simply the decaying part of $C^{\text{intern.}}(t)$, normalized such that its initial value is 1, and final value is 0. If the second assumption, fast decay of the correlation function due to internal motion is fulfilled, we may calculate $J(\omega)$ using the parameters τ_{M}, S^2, and $\langle\tau_e\rangle$, where

$$\langle\tau_e\rangle = \int_0^\infty e^{-t/\tau_{\text{M}}} G(t)dt \tag{4}$$

We calculate $J(\omega)$ in order to see how it is a function of the parameters τ_{M}, S^2, and $\langle\tau_e\rangle$.

$$\begin{aligned} J(\omega) &= \frac{2}{5}\int_0^\infty \left[S^2 e^{-t/\tau_{\text{M}}} + \left(1-S^2\right)e^{-t/\tau_{\text{M}}}G(t)\right]\cos(\omega t)dt \\ &= \frac{2}{5}\left[\frac{S^2\tau_{\text{M}}}{1+(\omega\tau_{\text{M}})^2} + \left(1-S^2\right)\int_0^\infty e^{-t/\tau_{\text{M}}}G(t)\underbrace{\cos(\omega t)}_{\approx 1}\,dt\right] \\ &= \frac{2}{5}\left[\frac{S^2\tau_{\text{M}}}{1+(\omega\tau_{\text{M}})^2} + \left(1-S^2\right)\langle\tau_e\rangle\right] \end{aligned} \tag{5}$$

We see that if $e^{-t/\tau_{\text{M}}}G(t)$ decays quickly compared to ω, then we may replace $\cos(\omega t)$ with 1, since the exponential approaches zero more quickly than the cosine term can evolve away from 1. Then, regardless of the precise form of $G(t)$, $J(\omega)$ may always be calculated from the parameters S^2, $\langle\tau_e\rangle$, and τ_{M}. Furthermore, if τ_{M} is known (usually from the analysis of R_1 and R_2 throughout a molecule (Kay et al., 1989)), $J(\omega)$ becomes a linear function of the parameters S^2 and $(1-S^2)\langle\tau_e\rangle$.

Instead of assuming fast decay of $G(t)$, one may alternatively assume that it is mono-exponential ($G(t) = e^{-t/\tau}$), yielding

$$\begin{aligned} J(\omega) &= \frac{2}{5}\int_0^\infty \left[S^2 e^{-t/\tau_{\text{M}}} + \left(1-S^2\right)e^{-t/\tau_{\text{M}}}e^{-t/\tau}\right]\cos(\omega t)dt \\ \langle\tau_e\rangle^{-1} &= \tau_{\text{M}}^{-1} + \tau^{-1} \\ J(\omega) &= \frac{2}{5}\left[\frac{S^2\tau_{\text{M}}}{1+(\omega\tau_{\text{M}})^2} + \left(1-S^2\right)\frac{\langle\tau_e\rangle}{1+(\omega\langle\tau_e\rangle)^2}\right] \end{aligned} \tag{6}$$

In the extreme narrowing limit, where decay of the correlation function is fast, we have $\omega\langle\tau_e\rangle \ll 1$ such that this result equals the result in **Eq. 5**. The expression in **Eq. 6** is equivalent to **Eq. 1** in (Lipari and Szabo, 1982a), and is valid either in the case of mono-exponential decay or fast decay of the internal correlation function. However, we find the case of fast, multi-exponential decay the more likely scenario, and so focus on this assumption.

The notation $\langle\tau_e\rangle$ is used to indicate the average of the effective correlation time. To understand how the integral of $e^{-t/\tau_{\text{M}}}G(t)$ is related to this average, we must assume that $G(t)$ is the sum of decaying exponentials. This may be achieved with a sum over a discrete number of correlation times, weighted with A_i, or a continuous distribution, defined by the function $\theta(z)$.

$$\begin{aligned} &G(t) = \sum_i A_i e^{-t/\tau_i} \\ &\text{where } \Sigma_i A_i = 1 \\ &\text{–or–} \\ &G(t) = \int_{-\infty}^{\infty} \theta(z) e^{-t/(10^z\cdot 1\text{ s})}dz \\ &\text{where } \int_{-\infty}^{\infty}\theta(z)dz = 1 \end{aligned} \tag{7}$$

Since $G(0) = 1$, it is clear that the sum of amplitudes (A_i) must be 1. For the former equation, we take a simple sum, and for the latter form, we use a distribution of correlation times, $\theta(z)$, given on a logarithmic scale, such that $z = \log_{10}(\tau_c/\text{s})$. The distribution must similarly integrate to 1. The two forms can be treated equivalently. We have recently re-introduced the latter form (Smith et al., 2018), which was previously used to describe a variety of continuous correlation time distributions, e.g., see Beckmann (1988). We may insert this expression for $G(t)$ (**Eq. 7**) into **Eq. 4** in order to obtain the relationship between $\theta(z)$ and $\langle\tau_e\rangle$.

$$\begin{aligned} \langle\tau_e\rangle &= \int_0^\infty e^{-t/\tau_{\text{M}}} \underbrace{\int_{-\infty}^{\infty}\theta(z)e^{-t/(10^z\cdot 1\text{ s})}dz}_{G(t)}\,dt \\ &= \int_0^\infty\int_{-\infty}^{\infty}\theta(z)e^{-t\left(\tau_{\text{M}}^{-1}+(10^z\cdot 1\text{ s})^{-1}\right)}dzdt \\ (\tau_e(z))^{-1} &= \tau_{\text{M}}^{-1} + (10^z\cdot 1\text{ s})^{-1} \\ \langle\tau_e\rangle &= \int_0^\infty\int_{-\infty}^{\infty}\theta(z)e^{-t/\tau_e(z)}dzdt = \int_{-\infty}^{\infty}\theta(z)\left(-\tau_e(z)e^{-t/\tau_e(z)}\right)\Big|_{t=0}^{\infty}dz \\ \langle\tau_e\rangle &= \int_{-\infty}^{\infty}\theta(z)\tau_e(z)dz \\ &\text{equivalently: } \langle\tau_e\rangle = \sum_i A_i\tau_e^i, \text{for } \left(\tau_e^i\right)^{-1} = \tau_{\text{M}}^{-1} + \tau_i^{-1} \end{aligned} \tag{8}$$

$(\tau_e^i)^{-1} = \tau_{\text{M}}^{-1} + (10^z\cdot 1\text{ s})^{-1}$ is the effective correlation time, resulting from decay of both the correlation function due to the internal correlation time, $z = \log_{10}(\tau_c/s)$ and correlation time of the overall motion, τ_{M}. Since $\theta(z)$ integrates to 1, $\int_{-\infty}^{\infty}\theta(z)\tau_e(z)dz$ yields the weighted average of the effective correlation time, $\langle\tau_e\rangle$. Then, one fits experimental data to a correlation function having the following model:

$$C(t) = \frac{1}{5}\left(S^2 e^{-t/\tau_M} + (1 - S^2)e^{-t/\langle\tau_e\rangle}\right) \tag{9}$$

Applying this model does not require that the true correlation function has exactly this form, but rather, the model correlation function simply must have the same values of S^2 and $\langle\tau_e\rangle$ as the true correlation function. In this sense, the analysis itself remains model-free, although equating $\langle\tau_e\rangle$ with the averaged effective correlation time requires the true correlation function to be a sum of decaying exponentials, as in **Eq. 7**.

A Few Notes on Linearity

We will later note that many of the methods used for analyzing relaxation rate constants result in parameters that are linear functions of the distribution of correlation times, $(1 - S^2)\theta(z)$. Specifically, we mean that any parameter, P_m, is linear to $(1 - S^2)\theta(z)$ if it can be written as

$$P_m = (1 - S^2)\int_{-\infty}^{\infty} \theta(z)p_m(z)dz \tag{10}$$

That is, for every correlation time, $z = \log_{10}(\tau_c/\mathrm{s})$, P increases proportionally to $(1 - S^2)\theta(z)$ at that correlation time, where the proportionality is defined by $p_m(z)$. Furthermore, any linear combination of parameters, P_m, is then also linear to $(1 - S^2)\theta(z)$, as we can see by integrating a sum of parameters, P_m, and swapping the order of the integration and the summation.

$$\begin{aligned}\sum_m a_m P_m &= \sum_m a_m (1 - S^2)\int_{-\infty}^{\infty} p_m(z)\theta(z)dz \\ &= (1 - S^2)\int_{-\infty}^{\infty} \underbrace{\left[\sum_m a_m p_m(z)\right]}_{=\Sigma(z)} \theta(z)dz \\ &= (1 - S^2)\int_{-\infty}^{\infty} \Sigma(z)\theta(z)dz \end{aligned} \tag{11}$$

We define the function $\Sigma(z)$ to be the weighted sum of the sensitivities, $p_m(z)$, which then defines the linear relationship of the sum of the P_m to $(1 - S^2)\theta(z)$. This principle is one of the basic tenants of linear algebra. What can be less obvious is that a *linear fit* of parameters, P_m, defined by a matrix, $\mathbf{M}$, to a new set of parameters, Q_n is also linear to $(1 - S^2)\theta(z)$. This is only the case if restrictions on the fit parameters, Q_n, are not applied (no *priors* are used). In this case, the parameters Q_n should minimize the following equation.

$$\begin{aligned}&\min\left[\sum_m \left|P_m - [\mathbf{M}]_{m,n}Q_n\right|^2\right] \\ &Q_n = \sum_m [\mathbf{M}^{-1}]_{n,m}P_m\end{aligned} \tag{12}$$

One may determine the Q_n by computing the pseudoinverse of $\mathbf{M}$ (denoted $\mathbf{M}^{-1}$) and multiplying by the P_m. Linearity of the Q_n to $(1 - S^2)\theta(z)$ results from the fact that linear combinations defined by $\mathbf{M}^{-1}$ remain unchanged regardless of the value of the parameters being fit, P_m. However, if the allowed values of the Q_n are restricted with priors, then it can be that some values of P_m will result in the latter formula in **Eq. 12** yielding Q_n outside of the allowed range. In this case, a linear least squares algorithm will search for a different solution than that given by **Eq. 12**, such that the Q_n are no longer defined by $\mathbf{M}^{-1}$, and no longer have a consistent linear relationship to $(1 - S^2)\theta(z)$. Note that if priors are used, but **Eq. 12** does not yield Q_n outside of the bounds defined by the priors, then **Eq. 12** still remains the best solution and linearity is maintained. In general, we will find analysis methods that rely on linear combination of data have more predictable behavior than those that do not.

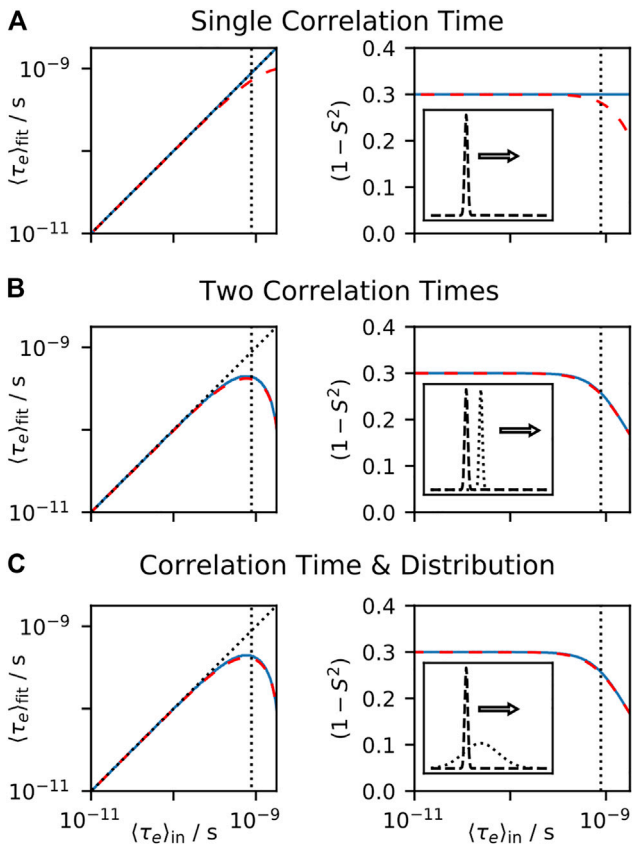


FIGURE 3 | Model-free fit parameters as a function of input parameters. For each plot, a data set is calculated, using the experiments found in from Table I of Lipari and Szabo (1982b), and the resulting rate constants are fit using the model-free approach, with the resulting $\langle\tau_e\rangle_{\mathrm{fit}}$ and $(1 - S^2)$ shown on the left and right, respectively. For all plots, the tumbling correlation time is $\tau_M = 4$ ns and $(1 - S^2) = 0.3$. One correlation time of the internal motion is varied, and we plot $\langle\tau_e\rangle_{\mathrm{in}}$ on the *x*-axis. In each plot, we fit using the full spectral density (blue, solid, see **Eq. 6**) and using a linear approximation (red, dashed, see **Eq. 5**). In **(A)**, the input correlation function only has a single correlation time. In **(B)**, one correlation time is fixed to 10 ps, and the second correlation time is swept. In **(C)**, a log-Gaussian distribution (μ = 10 ps, σ = 0.75 order of magnitude) is combined with a correlation time that is varied (with total amplitude equal). On the left plots, black dotted lines indicate where the input value, $\langle\tau_e\rangle_{\mathrm{in}}$, matches the fit, $\langle\tau_e\rangle_{\mathrm{fit}}$. In all plots, vertical black dotted lines indicate where $\omega\langle\tau_e\rangle_{\mathrm{in}} = 0.5$ for $\omega/2\pi$ = 90 MHz, where this frequency corresponds to the highest field used for the data set.

Then, the model-free parameters S^2 and $(1-S^2)\langle\tau_e\rangle$ are linear to $(1-S^2)\theta(z)$, because one can fit experimental relaxation rate constants with S^2 and $(1-S^2)\langle\tau_e\rangle$ (see **Eq. 5**), where the relaxation rate constants themselves are linear to the spectral density (**Eq. 1**), the spectral density is linear to the correlation function (via Fourier transform, **Eq. 2**), and the correlation function is linear to the distribution of correlation times, $(1-S^2)\theta(z)$ (**Eqs. 3**, **7**)). Assuming the correlation function decays quickly, this linear relationship is given by the following, where $\tau_e(z)$ is defined in **Eq. 8**.

$$S^2 = 1 - \left[(1-S^2)\int_{-\infty}^{\infty}\theta(z)dz\right]$$
$$(1-S^2)\langle\tau_e\rangle = (1-S^2)\int_{-\infty}^{\infty}\tau_e(z)\theta(z)dz \tag{13}$$

Note that $\langle\tau_e\rangle$ is not itself linear to $(1-S^2)\theta(z)$, but is easily obtained from the above parameters.

Fitting With Model-Free

In **Figure 3**, we test the performance of model-free fitting under a number of conditions. In **Figure 3A**, we calculate a number of relaxation rate constants from motion having a single internal correlation time and overall tumbling with $\tau_M = 4$ ns, and then fit the results, assuming the model-free correlation function (**Eq. 9**). We may calculate the spectral density exactly, or we may assume that the correlation function decays quickly, by using the spectral density given in **Eq. 5**, resulting in a linear fit. The former method is shown as a blue, solid line, where the input parameters always exactly match the fit parameters, whereas using a linear fit (red, dashed line) results in disagreement of input and fit parameters when the correlation function does not decay quickly compared to the frequencies sampled ($\omega\tau_e \ll 1$); in this case, **Eq. 5** is no longer a good estimate of the spectral density whereas **Eq. 6** has the correct form.

In **Figure 3B**, we include two correlation times in the input, each with equal amplitude, where one correlation time is fixed (10 ps), and a second correlation time is swept. We calculate the mean effective correlation time directly on the x-axis ($\langle\tau_e\rangle_{\text{in}}$), and compare this to the fitted parameters on the y-axis ($\langle\tau_e\rangle_{\text{fit}}$, left plot, $1-S^2$, right plot). As expected, if the assumption that $\omega\tau_e \ll 1$ holds for all frequencies sampled and all correlation times present, the fit parameters are in good agreement with their input values, but when $\omega\tau_e \not\ll 1$, $\langle\tau_e\rangle_{\text{fit}}$ and S^2 no longer reproduce the correct values. Note that performing this fit with the full spectral density (blue, solid line) and using just a linear fit (red, dashed line) produces very similar results. In **Figure 3C**, we perform the same tests, but instead of fixing a correlation time to 10 ps, we have a log-Gaussian distribution of correlation times, centered at 10 ps, with a standard deviation of 0.75 orders of magnitude. Results are similar to those found in **Figure 3B**.

Determining S^2

For model-free analysis, $\langle\tau_e\rangle$ is the average effective correlation time, and can be calculated from the distribution of correlation times. S^2, on the other hand, is determined from the distribution of orientations sampled by internal motion. By definition, it is equal to the correlation function of internal motion, taken as the limit of t goes to infinity. We may obtain S^2 by first considering the formula for the correlation function.

$$C^{\text{intern.}}(t) = \left\langle P_2\left(\vec{\mu}(\tau)\cdot\vec{\mu}(t+\tau)\right)\right\rangle_\tau \tag{14}$$

$P_2(x)$ is the second Legendre polynomial ($P_2(x) = (3x^2-1)/2$), and $\vec{\mu}(\tau)$ is a normalized vector that gives the direction of the principal component of an NMR interaction as a function of time, due to internal motion only (without tumbling). The dot product $(\vec{\mu}(\tau)\cdot\vec{\mu}(t+\tau))$ yields the cosine of the angle between the two vectors. The correlation function itself may take on a variety of complex forms, depending on the correlation times present, but S^2, its value as $t\to\infty$, depends only on the distribution of orientations sampled by the internal motion. This may be obtained by taking a weighted average over all possible starting orientations (p) and all possible final orientations (q), and calculating $P_2(\vec{\mu}_p\cdot\vec{\mu}_q)$ for each pair. Defining $p_{\text{eq}}(\vec{\mu}_p)$ to be the fraction of orientation $\vec{\mu}_p$ at thermal equilibrium, we obtain

$$S^2 = \sum_p\sum_q p_{\text{eq}}\left(\vec{\mu}_p\right)p_{\text{eq}}\left(\vec{\mu}_q\right)P_2\left(\vec{\mu}_p\cdot\vec{\mu}_q\right) \tag{15}$$

Then, if we have a precise description of the internal dynamics, we may calculate parameters $\langle\tau_e\rangle$ and S^2 using **Eqs. 8**, **15**. We may not easily go backwards, to obtain a precise description of the dynamics from only these parameters. However, this is not a limitation of the method of analysis, but rather of the information content of the data.

In solid-state NMR, we no longer have overall tumbling motion, so the term e^{-t/τ_M} vanishes from the correlation function and **Eq. 5** becomes simply

$$J(\omega) = \frac{2}{5}(1-S^2)\langle\tau\rangle \tag{16}$$

This prevents us from separating S^2 and $\langle\tau\rangle$ via relaxation data alone (we drop the subscript e from τ, since it is no longer an effective correlation time); however, one may measure the size of residual couplings in NMR (Chevelkov et al., 2009a; Schanda et al., 2011), often via DIPSHIFT (Munowitz et al., 1981) or REDOR (Gullion and Schaefer, 1989). In this case, the ratio of the anisotropies of the rigid interaction ($\delta_{\text{rigid.}}$) to the motionally averaged interaction ($\delta_{\text{resid.}}$) defines $S_{\text{resid.}}$.

$$S_{\text{resid.}} = \delta_{\text{resid.}}/\delta_{\text{rigid}} \tag{17}$$

One usually equates S^2 and $S^2_{\text{resid.}}$, although for motion that does not have at least a three-fold symmetry axis, these terms are not necessarily equal (**Supplementary Section S3**). Examples are found in Figures 2C-E, although we see the deviation is actually quite small (e.g., $S^2_{\text{resid.}} = 0.69$, vs. $S^2 = 0.7$), so that this approach may be used to obtain good separation of S^2 and $\langle\tau\rangle$.

ALTERNATIVE METHODS

In the case that all internal motion is fast, such that the correlation function decays quickly, model-free analysis is an ideal approach for extracting dynamics information from relaxation data: the full

information content of the relaxation data is captured in the parameters S^2 and $\langle\tau_e\rangle$, where these parameters have simple relationships to the distribution of correlation times, $(1-S^2)\theta(z)$ (parameters S^2 and $(1-S^2)\langle\tau_e\rangle$ are furthermore linearly related to $(1-S^2)\theta(z)$). In case the correlation function does not decay quickly compared to the sampled frequencies, our formula for the spectral density becomes significantly more complex. To obtain it, we begin from **Eq. 5** (first expression), and insert the assumed form of $G(t)$, found in **Eq. 7**, yielding the equation for the solution-state spectral density.

$$\begin{aligned}
&J(\omega) = \frac{2}{5}\int_0^\infty \left[S^2 e^{-t/\tau_M} + (1-S^2)e^{-t/\tau_M}\int_{-\infty}^{\infty}\theta(z)e^{-t/(10^z\cdot 1\,\mathrm{s})}dz\right]\cos(\omega t)dt \\
&(\tau_e(z))^{-1} = \tau_M^{-1} + (10^z\cdot 1\,\mathrm{s})^{-1} \\
&z_e(z) = \log_{10}(\tau_e(z)/\mathrm{s}) \\
&J(\omega) = \frac{2}{5}\int_0^\infty \left[S^2 e^{-t/\tau_M} + (1-S^2)\int_{-\infty}^{\infty}\theta(z)e^{-t/\left(10^{z_e(z)}\cdot 1\,\mathrm{s}\right)}dz\right]\cos(\omega t)dt \\
&\quad = \frac{2}{5}\left[\frac{S^2\tau_M}{1+(\omega\tau_M)^2} + (1-S^2)\int_{-\infty}^{\infty}\theta(z)\frac{10^{z_e(z)}\cdot 1\,\mathrm{s}}{1+\left(\omega\cdot 10^{z_e(z)}\cdot 1\,\mathrm{s}\right)^2}dz\right]
\end{aligned} \tag{18}$$

The first step is to combine the two exponential terms, where we define the log-effective correlation time, $z_e(z)$, as a function of the log-internal correlation time, z, and also the rotational correlation time, τ_M. Subsequently, each exponential term is Fourier transformed to yield the familiar Lorentzian function. The spectral density for solid-state NMR can be similarly calculated, where the overall motion is omitted.

$$\begin{aligned}
J(\omega) &= \frac{2}{5}\int_0^\infty (1-S^2)\int_{-\infty}^{\infty}\theta(z)e^{-t/(10^z\cdot 1\,\mathrm{s})}\,dz\cos(\omega t)dt \\
&= \frac{2}{5}(1-S^2)\int_{-\infty}^{\infty}\theta(z)\frac{10^z\cdot 1\,\mathrm{s}}{1+(\omega\cdot 10^z\cdot 1\,\mathrm{s})^2}dz
\end{aligned} \tag{19}$$

The integral has a complex dependence on ω, and depends on the specific form of $(1-S^2)\theta(z)$, so that by using multiple relaxation experiments, we can extract more than two parameters describing the internal motion. However, we require a different approach to extract that information. We discuss four approaches developed for treating this case: the extended model-free approach (EMF), spectral density mapping (SDM), LeMaster's approach, and IMPACT. Another approach that bears mentioning is the slowly relaxing local structure model (SRLS), which accounts for coupling of local motional modes to overall motion of a molecule in solution (Polimeno and Freed, 1992; Tugarinov et al., 2001; Mendelman and Meirovitch, 2021; Shapiro and Meirovitch, 2012). SRLS reduces to the model-free approach as coupling between local and overall motion vanishes. However, we do not include further comparison to the analytically simpler methods discussed here.

Extended Model-Free

Clore and coworkers found that when measuring relaxation data at higher fields (up to 600 MHz) that not all backbone motion could be well fit using the model-free approach for staphylococcal nuclease and interleukin-1β (Clore et al., 1990). They found that the simplest correlation function that could fit the data was obtained by adding another decaying exponential term, yielding the EMF correlation function.

$$C^{\text{intern.}}(t) = \left(1-S_f^2\right)e^{-t/\tau_f} + S_f^2\left(1-S_s^2\right)e^{-t/\tau_s} + S_f^2 S_s^2 \tag{20}$$

In this correlation function, the total internal motion is separated into fast and slow components, with order parameters S_f^2 and S_s^2, and effective correlation times, τ_f and τ_s, respectively. The product $S_f^2 S_s^2$ should yield the total order parameter, S^2. Also note that the faster motion's order parameter scales the influence of the slower motion, as seen in the term $S_f^2(1-S_s^2)e^{-t/\tau_s}$. Data analysis with EMF in solid- and solution-state NMR involves simply varying the parameters, S_f^2, S_s^2, τ_f, and

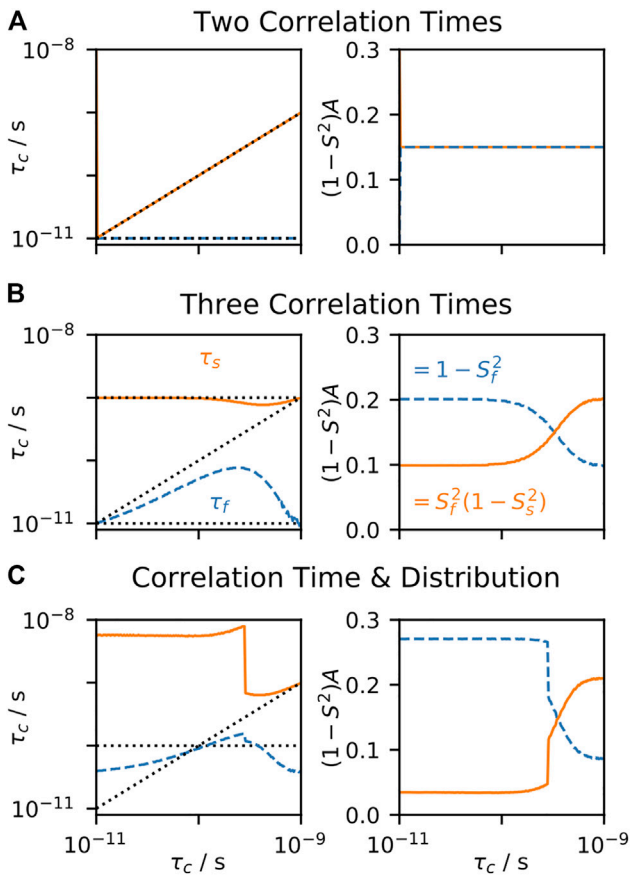


FIGURE 4 | EMF parameters as a function of input correlation time (solution-state). For each plot, a data set is calculated, using the set of experiments from Clore et al. (1990), and the resulting rate constants are fitted using the EMF approach. For all plots, τ_M = 8.3 ns, and the input $(1-S^2)$ = 0.3. In each subplot, the fitted correlation times **(left)** and amplitudes **(right)** are shown, as a function of an input correlation time (*x*-axis). In **(A)**, the input correlation function has two correlation times (with equal amplitudes), with one fixed at 10 ps, and the other swept. In **(B)**, the input correlation function has three correlation times, two fixed at 10 ps and 1 ns, and the third is swept. In **(C)**, a log-Gaussian distribution of correlation times is used (μ = 100 ps, σ = 0.75 orders of magnitude), and a single correlation time is swept. Black dotted lines show the input correlation times **(left plots)**.

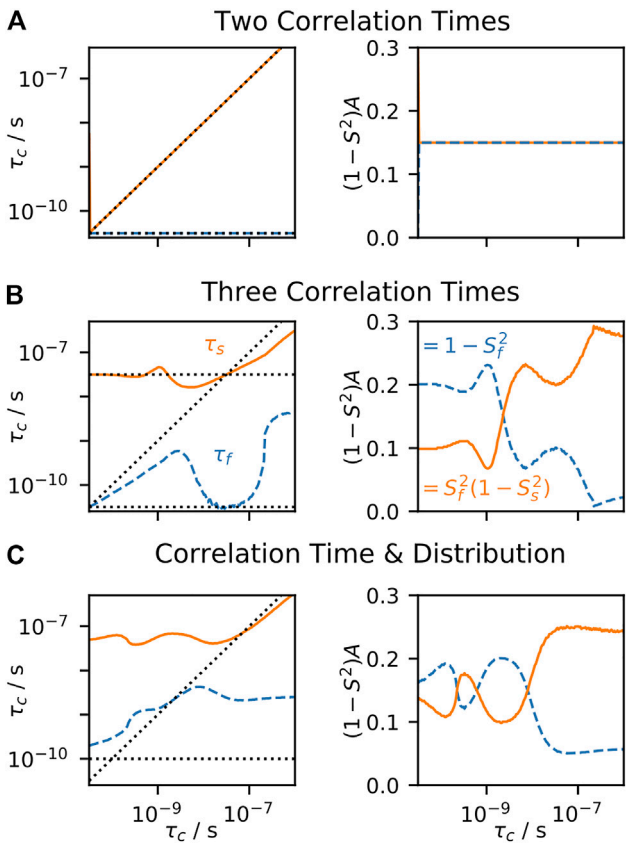


FIGURE 5 | EMF parameters as a function of input correlation time (solid-state). For each plot, a data set is calculated, including direct measurement of $S_{resid.}$ via residual couplings (**Eq. 17**), ^{15}N T_1 at 400, 500, and 850 MHz, and T_2 with MAS of 60 kHz. The resulting rate constants are fitted using the EMF approach. For all plots, $(1 - S^2) = 0.3$. In each subplot, the fitted correlation times **(left)** and amplitudes **(right)** are shown, as a function of an input correlation time (*x*-axis). In **(A)**, the input correlation function has two correlation times (with equal amplitudes), with one fixed at 3.2 ps, and the other swept. In **(B)**, the input correlation function has three correlation times, two fixed at 3.2 ps and 32 ns, and the third is swept. In **(C)**, a log-Gaussian distribution of correlation times is used (μ = 100 ps, σ = 0.75 orders of magnitude), and a single correlation time is swept. Black dotted lines show the input correlation times (left plots).

τ_s, to find an optimal fit to experimental data. Often, one also performs a model selection step, where one may determine how many parameters should be included in the fit (Mandel et al., 1995; d'Auvergne and Gooley, 2003; Zinkevich et al., 2013; Gill et al., 2016). In **Figure 4**, the behavior of EMF parameters is shown for several correlation functions. In each subplot, all terms except one correlation time are fixed, and we observe the model behavior as we sweep through the variable correlation time. In **Figure 4A**, two correlation times are used, so that the input correlation function has the same form as the correlation function used for fitting; as expected, the fitted parameters perfectly match the input parameters, since the input and fit models match. In **Figure 4B**, three correlation times are input, where the fast and slow correlation times are fixed at 10 ps and 1 ns, and the intermediate correlation time is swept. In this case, when the intermediate correlation time is fast, the fitted τ_f falls in between the fast and intermediate correlation times, and the fitted amplitude for the fast motion is the sum of the input amplitudes for the fast and intermediate motions. However, for longer correlation times, the fitted τ_f again gets shorter, eventually equaling 10 ps, so that the fitted τ_s takes over the role of fitting the intermediate correlation time. This is especially well illustrated in **Figure 4(B**, right), where the amplitude corresponding to the slow motion increases from 0.1 to 0.2, indicating that the slow motion in the model fits both the input intermediate and slow motions. Similar behavior is observed in **Figure 4C**, where a distribution of correlation times is combined with a single correlation time that is swept.

To the best of our knowledge, the behavior of the fit parameters has no well-defined relationship to the distribution of correlation times, $(1 - S^2)\theta(z)$: if we know $(1 - S^2)\theta(z)$ precisely, our only way to obtain the EMF parameters from it would be to explicitly calculate a set of relaxation rate constants, and then fit the results to **Eq. 20**. This is in sharp contrast to the original model-free parameters. Similar limitations arise for the EMF approach in solid-state NMR, as seen in **Figure 5**. Note that typical solution-state data sets are fairly continuous in their sensitivity to motion as a function of correlation time (Smith et al., 2019a), whereas solid-state NMR has a "blind-spot" in sensitivity centered around ~100 ns (Schanda, 2019), which results in some of the more unusual behavior for EMF in solids (see *Case 1: Extended Model-Free* for a detailed discussion of the behavior of typical model-free parameters in solid-state NMR).

Spectral Density Mapping

In contrast to EMF, SDM is achieved by simple linear combination of sets of relaxation data at a single magnetic field (Peng and Wagner, 1992; Ishima et al., 1999). From a set of R_1, R_2, and NOE relaxation rate constants, one calculates

$$\begin{aligned} J(0) &= \frac{R_2 - R_1/2 - 0.454\sigma_{IS}}{\delta_{IS}^2/2 + 2(\Delta\sigma_I\omega_I)^2} \\ J(\omega_I) &= \frac{R_1 - 1.249\sigma_{IS}}{3(\delta_{IS}/4)^2 + (\Delta\sigma_I\omega_I)^2/3} \\ J(0.870\omega_S) &= 16\sigma_{IS}/(5\delta_{IS}^2) \end{aligned} \tag{21}$$

The above expressions yield very close approximations of the spectral density at specific frequencies: 0, ω_I, and $0.870\omega_S$, where ω_I is the nuclear Larmor frequency of the spin being relaxed, and ω_S is a spin which is dipole coupled to that spin (usually a directly bonded ^{1}H). Differences in the representations of the anisotropies (δ_{IS}, $\Delta\sigma_I\omega_I$) result in the different appearances of the normalization factors (denominators). These terms may be interpreted as being proportional to the amount of motion near the given frequency (which corresponds to the correlation time $\tau = 1/\omega$), but otherwise they do not provide a more physical interpretation of the motion. One may subsequently fit the spectral densities to model-free parameters for better interpretation (Gill et al., 2016). If we have a precise

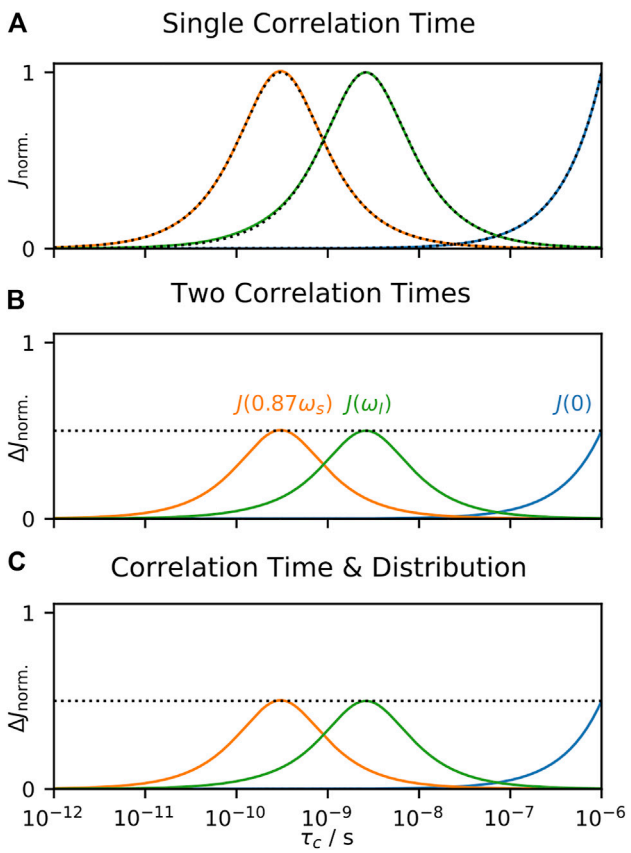


FIGURE 6 | Behavior of SDM as a function of correlation time. In each subplot, we calculate ^{15}N T_1, T_2, and σ_{NH} at 600 MHz, and analyze the results using **Eq. 21**. In **(A)**, the input total correlation function consists of a single decaying exponential term (with amplitude 1), where the terms $J(\omega)$ are plotted as the correlation time is varied (results are normalized). Black dotted lines show the spectral densities, $J(0), J(\omega_l), J(0.870\omega_S)$, calculated with **Eq. 22**, and colored lines show the results of the data analysis, yielding an almost exact correspondence. In **(B)**, the total correlation function now uses two correlation times (equal amplitudes), with one fixed at 10 ps, and the second swept (*x*-axis). On the *y*-axis, we plot contribution to the terms, $\Delta J(\omega)$, from the correlation time being varied. The resulting behavior is identical to that in **(A)**, except that the amplitude is half as large, since we have split the total amplitude between the fixed and variable correlation time (dashed line marks 0.5). In **(C)**, the same information is plotted, but the total correlation function includes a log-Gaussian distribution (μ = 630 ps, σ = 1 order of magnitude), and a single, variable correlation time.

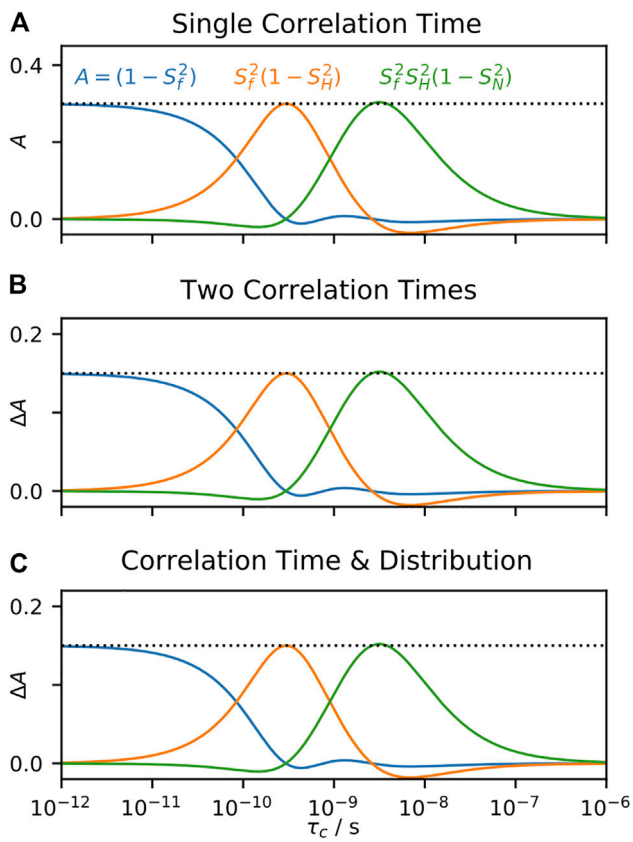


FIGURE 7 | Behavior of LeMaster's approach as a function of correlation time. In each subplot, we calculate ^{15}N T_1, T_2, and σ_{NH} at 600 MHz for motion with $(1 - S^2) = 0.3$ and tumbling correlation time of $\tau_M = 4$ ns, and analyze the results using **Eq. 24**. In **(A)** the internal correlation function consists of a single decaying exponential term (with amplitude 0.3), where the fitted amplitudes are plotted as the correlation time is varied. In **(B)** the internal correlation function uses two correlation times (both amplitudes are 0.15), with one correlation time fixed at 10 ps, and the second swept (*x*-axis). On the *y*-axis, we plot contributions to the terms from the correlation time being varied. The resulting behavior is identical to that in **(A)**, except that the amplitude is half, since we have split the total amplitude between the fixed and variable correlation time (dashed line marks 0.15). In **(C)**, the same information is plotted, but the total correlation function includes a log-Gaussian distribution (μ = 630 ps, σ = 1 order of magnitude), and a single, variable correlation time.

description of the motion (e.g., $(1 - S^2)\theta(z)$), the terms $J(\omega)$ are easily obtained:

$$J(\omega) = \frac{2}{5}(1 - S^2)\int_{-\infty}^{\infty} \theta(z)\frac{10^z \cdot 1\,\mathrm{s}}{1 + (\omega \cdot 10^z \cdot 1\,\mathrm{s})^2}dz \qquad (22)$$

The parameters resulting from SDM always behave the same way in response to a given correlation time, regardless of other correlation times present, and is the consequence of properties of linearity discussed in *A Few Notes on Linearity*. This is seen in **Figure 6A**, where we calculate relaxation rate constants resulting from a single correlation time and analyze with SDM. In **Figure 6B**, we split motion over two correlation times, and observe how the terms respond to sweeping one of them, and in **Figure 6C**, we split motion into a distribution and a single, swept correlation time and determine how the terms respond to the swept correlation time. The result is always identical (scaling by 0.5 results from dividing the total amplitude into two parts), a very useful property occurring when data is analyzed strictly by linear combination of data. Unlike EMF analysis, behavior of SDM is independent of the form of the distribution of correlation times.

Note that this approach describes the total motion, and does not separate out tumbling from internal motion in the case of solution-state NMR, which has an especially strong influence on $J(0)$. The original approach only incorporates data from one field, whereas later work has extended the method to include data

from more than one field, although one still requires specific sets of experiments (Skrynnikov et al., 2002; Hsu et al., 2018).

LeMaster's Approach

LeMaster proposed an alternative to SDM analysis of R_1, R_2, and NOE data from one field, in order to separate overall tumbling from internal motion (LeMaster, 1995). In this case, LeMaster proposed fitting data to the following correlation function:

$$C(t) = S_f^2 S_H^2 S_N^2 e^{-t/\tau_M} + S_f^2 (1 - S_H^2) e^{-t/\tau_H} + S_f^2 S_H^2 (1 - S_N^2) e^{-t/\tau_N} + \left(1 - S_f\right)^2 e^{-t/\tau_f}$$
$$\tau_H = (\omega_H + \omega_N)^{-1}, \quad \tau_N = |\omega_N|^{-1} \tag{23}$$

It is assumed that τ_f is very short so that the term $(1 - S_f^2)e^{-t/\tau_f}$ makes only negligible contributions to the spectral density, resulting in the following formula:

$$J(\omega) = \frac{2}{5} S_f^2 \left[S_H^2 S_N^2 \frac{\tau_M}{1 + (\omega\tau_M)^2} + (1 - S_H^2) \frac{\tau_H}{1 + (\omega\tau_H^2)} + S_H^2 (1 - S_N^2) \frac{\tau_N}{1 + (\omega\tau_N^2)} \right]$$
$$= \frac{2}{5} \left[\frac{\tau_M}{1 + (\omega\tau_M)^2} + (1 - S_f^2) \left(- \frac{\tau_M}{1 + (\omega\tau_M)^2} \right) + S_f^2 (1 - S_H^2) \left(\frac{\tau_H}{1 + (\omega\tau_H^2)} - \frac{\tau_M}{1 + (\omega\tau_M)^2} \right) + S_f^2 S_H^2 (1 - S_N^2) \left(\frac{\tau_N}{1 + (\omega\tau_N^2)} - \frac{\tau_M}{1 + (\omega\tau_M)^2} \right) \right] \tag{24}$$

In the latter formulation, we find that the spectral density becomes a linear combination of terms, weighted by $(1 - S_f^2)$, $S_f^2 (1 - S_H^2)$, and $S_f^2 S_H^2 (1 - S_N^2)$. Then, one must fit these terms to the experimental relaxation rate constants. We do so in **Figure 7** for calculated relaxation rate constants. Like SDM, responses as a function of correlation time are always identical (again, excepting a scaling factor of 0.5 resulting from splitting the total motion into components), although the functions themselves are different: this results from the fact that LeMaster's approach characterizes the internal motion, and not the total motion, so that we obtain one amplitude, $(1 - S_f^2)$, which captures information about the fastest correlation times (<30 ps), one amplitude, $S_f^2 (1 - S_H^2)$, which captures information for correlation times near to τ_H, and one amplitude, $S_f^2 S_H^2 (1 - S_N^2)$, which captures information for correlation times near to τ_N.

LeMaster's approach is a linear fit, without priors; as discussed in *A Few Notes on Linearity*, this means that the fitted parameters may also be obtained by a linear combination of the experimental relaxation rate constants. Therefore, the parameters $(1 - S_f^2)$, $S_f^2 (1 - S_H^2)$, and $S_f^2 S_H^S (1 - S_N^2)$ are linear to $(1 - S^2)\theta(z)$. The parameters S_H^2 and S_N^2 themselves are not linear to $(1 - S^2)\theta(z)$, but may be obtained by simple arithmetic from the linear parameters. Like SDM, LeMaster's approach is limited to data acquired at a single field.

Interpretation of Motions by a Projection onto an Array of Correlation Times Approach

Limitations of the approaches above have led Ferrage and coworkers to develop the interpretation of motions by a projection onto an array of correlation times (IMPACT) approach (Khan et al., 2015), which was applied to a protein with intrinsically disordered regions (IDR). A challenge of IDRs is that the lack of structure potentially yields a large number of distinct motions and therefore many correlation times, so that EMF approach is not appropriate for data analysis, but the limited number of parameters obtained with SDM fails to provide a complete description of the dynamics. Then, the IMPACT approach allows analysis of large, multi-field data sets, by taking the total correlation function to be a sum of several fixed correlation times, τ_k, such that

$$C(t) = \sum_k A_k e^{-t/\tau_k} \tag{25}$$

Because $C(0) = 1$ and decays to 0, the A_k must sum to 1. For the Engrailed 2 protein, ^{15}N T_1, NOE (σ_{NH}), and transverse and

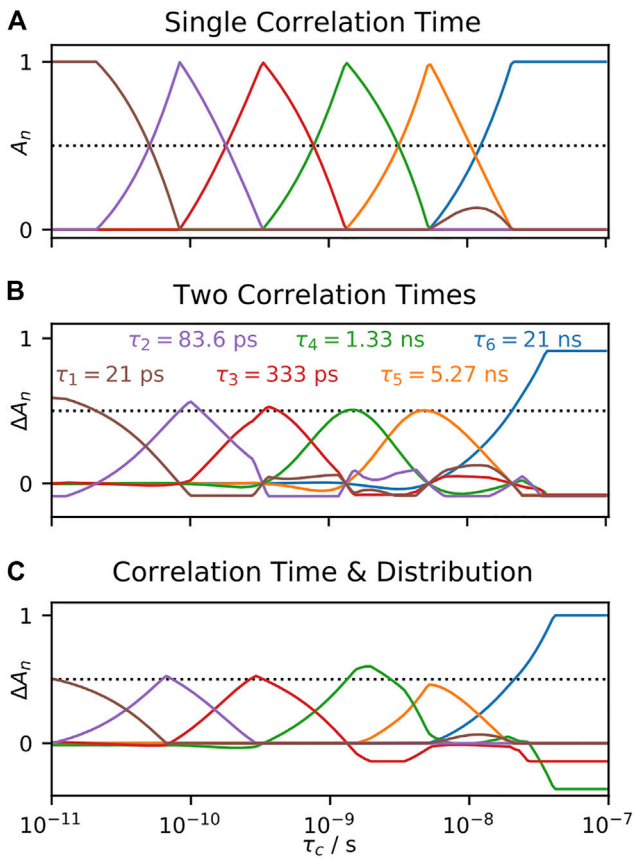


FIGURE 8 | Behavior of the IMPACT approach as a function of correlation time. In each plot, we fit calculated relaxation rate constants, and fit the amplitudes in **Eq. 25** according to the IMPACT procedure, using the set of experiments from Khan et al. (2015). In **(A)**, the input total correlation function consists of a single decaying exponential term (with amplitude 1), where the amplitudes are plotted as the correlation time is varied. In **(B)**, the total correlation function uses two correlation times (equal amplitudes), with one fixed at 1 ns, and the second swept (*x*-axis). On the *y*-axis, we plot contributions to the A_k from the correlation time being varied. In **(C)**, the same information is plotted, but the total correlation function includes a log-Gaussian distribution (μ = 630 ps, σ = 1 order of magnitude), and a single, variable correlation time.

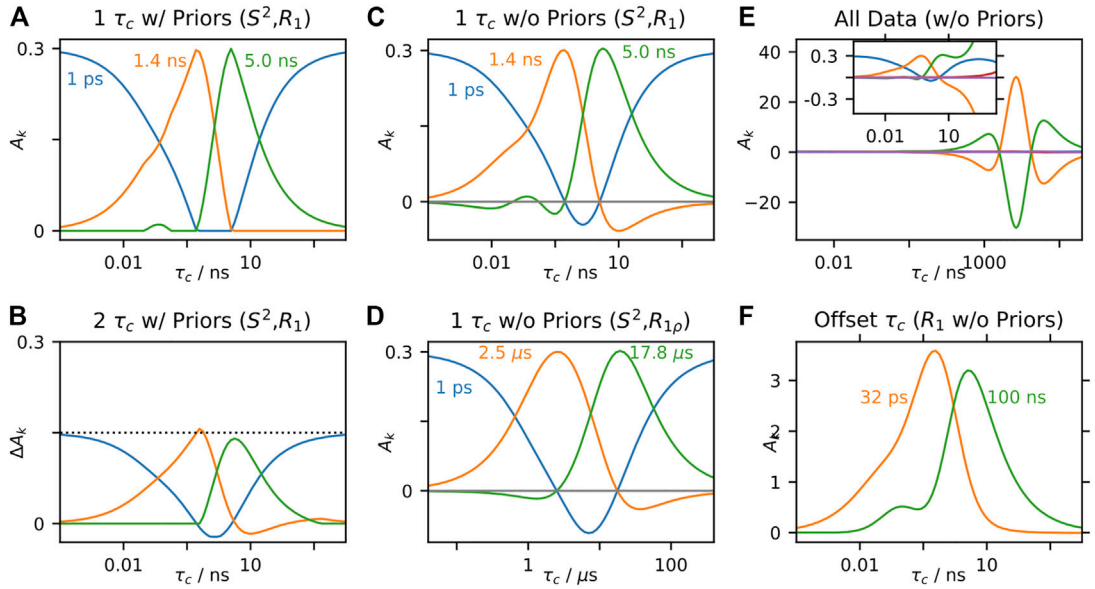


FIGURE 9 | IMPACT behavior in solids. In each plot, we test the behavior of the amplitudes, A_k, using calculated solid-state NMR data (S^2, ^{15}N R_1, ^{15}N $R_{1\rho}$, with experimental conditions taken from Smith et al. (2016)). **(A)** plots the behavior of fitting S^2 and three R_1 rate constants to three correlation times (1 ps, 1.4 ns, 5 ns), where the input correlation function has a single correlation time ($(1 - S^2)$=0.3), while restricting theA_k to fall between 0 and 1. **(B)** shows fits under the same conditions, but includes two correlation times, with one fixed at 1 ns, and the other swept (*x*-axis). The *y*-axis plots the change in the A_k due to the swept correlation time. **(C)** shows fits under the same conditions as **(A)**, without restricting the values of the A_k. **(D)** also removes restrictions on the A_k, but fits S^2 and $R_{1\rho}$ data, using correlation times of (1 ps, 2.5 μs, and 17.8 μs). **(E)** fits all data (S^2, R_1, $R_{1\rho}$) simultaneously without restrictions on the A_k, with correlation times of 1 ps, 1.4 ns, 5 ns, 2.5 μs, and 17.8 μs **(F)** fits R_1 data, but uses one very short correlation time (32 ps), and one very long correlation time (100 ns).

longitudinal cross-relaxation rate constants at five fields (400, 500, 600, 800, 1,000 MHz) could be fit to an array of six correlation times, log-spaced between 21 ps and 21 ns. When fitting to **Eq. 25**, one restricts the amplitudes to remain between zero and one, and the sum of amplitudes must be set to one.

Following our procedures for SDM and LeMaster's approach, we also examine the behavior of the IMPACT approach in **Figure 8**. When fitting a correlation function having a single correlation time in **Figure 8A**, we obtain ideal behavior from the IMPACT approach. When the input correlation time matches one of the correlation times in the IMPACT array, the corresponding amplitude is one, and all other amplitudes are zero. When the input correlation time is in between correlation times in the IMPACT array, then only the two nearest correlation times to the input value have non-zero amplitudes, and those two amplitudes sum to one (a minor deviation from this behavior occurs at 10 ns). However, if we input two correlation times in **Figure 8B**, or one correlation time and one distribution in **Figure 8C**, with motion split equally between the two correlation times or correlation time and distribution, the fit parameters' response to the swept correlation time is not an exact reproduction of the behavior in **Figure 8A**, in contrast to SDM and LeMaster's approach. While SDM and LeMaster's approach are both linear combinations of relaxation rate constants, IMPACT is a linear fit for which its behavior depends heavily on restricting the values of the fit parameters (priors), which as discussed in *A Few Notes on Linearity*, means that the fit parameters are no longer linear to $(1 - S^2)\theta(z)$. The result is that the response of the parameters A_k to a given correlation time do depend weakly on other motions present, thus not fully obtaining the ideal, linear behavior of SDM and LeMaster's approach. However, IMPACT provides a good approximation to this behavior, and is more generally applicable than SDM and LeMaster's approach.

IMPACT has not been developed for application to solid-state NMR, but it is worth investigating how such a method could work. In **Figure 9A**, we use an IMPACT-type approach to fitting R_1 at three fields and S^2, using an array of three correlation times. We restrict the fitted amplitudes (A_k) to fall between zero and one, but it does not make sense to require the A_k to sum to one, since the correlation function in solid-state NMR does not usually decay to zero. Here, we assume a motion with just one correlation time, and $(1 - S^2) = 0.3$. Then, we find that IMPACT in solids is similar to its solution-state behavior. Note that the amplitudes corresponding to 1.4 and 5 ns capture motion near those correlation times, whereas the amplitude corresponding to 1 ps captures *all* motion not in proximity to 1.4 and 5 ns, including very slow motions. As with solution-state NMR, if we split the motion over two correlation times, and determine the response to one of the two correlation times **Figure 9B**, the response changes compared to fitting just the single correlation time. However, as discussed in *A Few Notes on Linearity*, and demonstrated with SDM and LeMaster's approach, this dependence on other motions present vanishes if we eliminate restrictions on the fit parameters. Then, in **Figure 9C**, we repeat the fit from **Figure 9A**, without restrictions on the fit parameters, yielding reasonable

behavior, excepting some negative amplitudes in the A_k. **Figure 9D** shows similar results, when fitting S^2 and $R_{1\rho}$, although the fitted correlation times must be in the sensitive range of the $R_{1\rho}$ rate constants. Unfortunately, when we attempt to fit R_1 and $R_{1\rho}$ simultaneously in **Figure 9E**, using the same correlation times as in **Figures 9C,D**, we find extremely unstable behavior. Apparently, we cannot simultaneously fit data on both sides of the solid-state NMR blind spot.

In **Figures 9C,D**, we have fairly good performance, excepting that some of the amplitudes become slightly negative. Interestingly, these negative amplitudes may be eliminated by placing two correlation times further away from each other. Then, in **Figure 9F**, we fit only R_1 data, using correlation times of 32 ps and 100 ns. The fitted correlation times no longer correspond to the center of the sensitive range of the A_k (750 ps, 6.2 ns), and the amplitudes also far exceed the input value for $(1-S^2)$. Fitting while also including S^2 data allows using an additional correlation time (1 ps), but the corresponding A_k becomes large and negative (not shown). From this final result, we could simply renormalize the amplitudes to have a maximum of one, and report the center of the sensitive range instead of the correlation times to which we actually fitted. The result would still be a linear combination of the experimental data, and therefore linear to $(1-S^2)\theta(z)$, but the result would have very little to do with the correlation times chosen to obtain that linear combination.

A NEW APPROACH FOR SOLID-STATE NUCLEAR MAGNETIC RESONANCE

In the previous section, we investigated the behavior of a number of approaches to processing relaxation data. Of those approaches, model-free, SDM, and LeMaster's approach provide parameters which are linear to the distribution of correlation times, $(1-S^2)\theta(z)$ (in some cases, some additional arithmetic operations are required to obtain the reported parameters, e.g., $\langle\tau_e\rangle$ is calculated from S^2 and $(1-S^2)\langle\tau_e\rangle$). IMPACT approximates this behavior, although heavy reliance on priors prevents perfect linearity. However, each approach is limited in its application to solid-state NMR data. Therefore, we have developed the detector analysis (Smith et al., 2018), which is a general method for processing relaxation data that maintains a linear relationship between fit parameters and the distribution of correlation times.

Linear Combination of Data

As we have emphasized for the above examples, one may obtain parameters that have a well-defined (linear) relationship to the distribution of correlation times by taking linear combinations of relaxation rate constants. Thus far, we have limited ourselves to very specific linear combinations: combinations that yield the spectral density, or combinations that are related to specific correlation times. However, why shouldn't we use any linear combination that is optimized to give an ideal linear relationship to the distribution of correlation times, $(1-S^2)\theta(z)$? We first recall that the correlation function has been defined here as being a linear combination of decaying exponentials, defined by $(1-S^2)\theta(z)$, and its Fourier transform (also a series of linear combinations) must then also be linear to $(1-S^2)\theta(z)$.

$$C(t) = S^2 + \left(1-S^2\right)\int_{-\infty}^{\infty}\theta(z)e^{-t/(10^z\cdot 1\,\mathrm{s})}dz$$
$$J^{(\theta,S)}(\omega) = \frac{2}{5}\left(1-S^2\right)\int_{-\infty}^{\infty}\theta(z)\frac{10^z\cdot 1\,\mathrm{s}}{1+\left(\omega\cdot 10^z\cdot 1\,\mathrm{s}\right)^2}dz \tag{26}$$

Here, we take $J^{(\theta,S)}(\omega)$ to be the spectral density resulting from $(1-S^2)\theta(z)$. Then, any relaxation rate constant is a weighted sum of terms from the spectral density.

$$\begin{aligned}R_\zeta^{(\theta,S)} &= \sum_p a_p^\zeta J^{(\theta,S)}\left(\omega_p\right)\\ &= \sum_p a_p^\zeta\left(1-S^2\right)\int_{-\infty}^{\infty}\theta(z)\frac{10^z\cdot 1\,\mathrm{s}}{1+\left(\omega\cdot 10^z\cdot 1\,\mathrm{s}\right)^2}dz\\ &= \left(1-S^2\right)\int_{-\infty}^{\infty}\theta(z)\underbrace{\sum_p a_p\frac{10^z\cdot 1\,\mathrm{s}}{1+\left(\omega\cdot 10^z\cdot 1\,\mathrm{s}\right)^2}}_{=R_\zeta(z)}dz\\ R_\zeta^{(\theta,S)} &= \left(1-S^2\right)\int_{-\infty}^{\infty}\theta(z)R_\zeta(z)dz\end{aligned} \tag{27}$$

$R_\zeta^{(\theta,S)}$ is the relaxation rate constant for an experiment, indexed ζ, resulting from the distribution of correlation times, $(1-S^2)\theta(z)$. Coefficients a_p^ζ indicate the weightings of the spectral density for experiment ζ, sampled at frequencies ω_p. Insertion of $J^{(\theta,S)}(\omega)$ into this linear combination allows us to express $R_\zeta^{(\theta,S)}$ as a linear function of $(1-S^2)\theta(z)$, where $R_\zeta(z)$ defines the linear relationship (we refer to this as the sensitivity).

Then, as is the case for model-free, SDM, and LeMaster's approach, any sum of relaxation constants maintains linearity. Following our previous convention (Smith et al., 2018), we denote the sum as $\rho_n^{(\theta,S)}$.

$$\begin{aligned}\rho_n^{(\theta,S)} &= \sum_\zeta b_\zeta R_\zeta^{(\theta,S)}\\ &= \sum_\zeta b_\zeta\left(1-S^2\right)\int_{-\infty}^{\infty}\theta(z)R_\zeta(z)dz\\ &= \left(1-S^2\right)\int_{-\infty}^{\infty}\theta(z)\underbrace{\sum_\zeta b_\zeta R_\zeta(z)}_{=\rho_n(z)}dz\\ \rho_n^{(\theta,S)} &= \left(1-S^2\right)\int_{-\infty}^{\infty}\theta(z)\rho_n(z)dz\end{aligned} \tag{28}$$

Then, $\rho_n(z)$ defines the linear relationship between $(1-S^2)\theta(z)$ and $\rho_n^{(\theta,S)}$. The subsequent question is, how do we find the best linear combinations of the experimental relaxation rate constants for analyzing our relaxation data?

Optimizing Detectors: The Relaxation-Rate Space Approach

While any linear combination of experimental relaxation rate constants yields a linear relationship between $(1-S^2)\theta(z)$ and

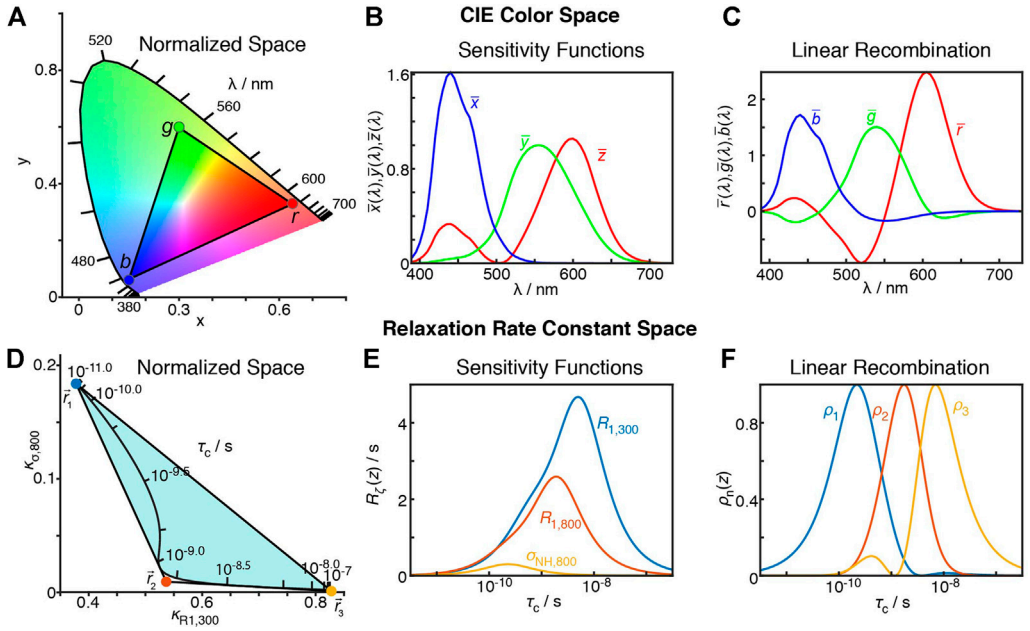


FIGURE 10 | Similarity between the CIE XYZ colorspace and the relaxation rate constant space. **(A)** plots the XYZ colorspace, black lines indicate where single wavelengths fall in the colorspace (z not shown, space is normalized such that $x + y + z = 1$). Points connected by a triangle indicate the definition of red, green, and blue colors as defined by the sRGB standard (Anderson et al., 1996). **(B)** plots the sensitivity of the $\overline{x}(\lambda)$, $\overline{y}(\lambda)$, and $\overline{z}(\lambda)$ color matching functions as a function of wavelength (λ). **(C)** plots sRGB sensitivities resulting from transformation from the XYZ to sRGB spaces. Points connected by triangles correspond to definitions of $\vec{r}_1$, $\vec{r}_2$, and $\vec{r}_3$ that define the detector space. **(D)** shows the normalized relaxation rate space for ^{13}C R_1 at 300 and 800 MHz and H–C NOE at 800 MHz. **(E)** shows the sensitivities of each of these experiments a function of correlation time. **(F)** shows detector sensitivities resulting from transformation from the relaxation rate constant space to detector space (defined by the points in **(D)**).

the resulting $\rho_n^{(\theta,S)}$, not all combinations are equally good choices. A few guidelines are, first, non-negativity of $\rho_n(z)$; we would like $\rho_n^{(\theta,S)}$ to always increase when amplitude of motion increases, whereas negative regions of $\rho_n(z)$ could cause $\rho_n^{(\theta,S)}$ to decrease with increasing amplitudes. Second, narrowness: we would like each $\rho_n^{(\theta,S)}$ to report on a specific range of correlation times. Third, when the full set of relaxation data is analyzed, one should be able to back-calculate the experimental data (within some tolerance) from the parameters $\rho_n^{(\theta,S)}$. This ensures that one captures all information in the experimental data (clearly, if the $\rho_n^{(\theta,S)}$ can reproduce the experimental data, then the $\rho_n^{(\theta,S)}$ must have retained the information in the experiments).

The question, then, is how to obtain optimized linear combinations satisfying the above requirements. Our initial answer to this question is the result of identifying a similar problem in a completely different field: When one sees the color of an object, its appearance depends on the distribution of wavelengths reflected (or emitted) by the object. The distribution of wavelengths is given by the spectral power distribution, $S(\lambda)$. Whereas $S(\lambda)$ is an infinite-dimensional description of the spectral power vs. wavelength, what is "seen" is a projection of that distribution onto a three dimensional space, corresponding to the three cones that detect color in the eye. This 3D space is often described using the CIE (Commission internationale de l'Eclairage) XYZ color space (Smith and Guild, 1931; Judd, 1951; Vos, 1978).

$$X = \int_0^\infty S(\lambda)\overline{x}(\lambda)d\lambda$$
$$Y = \int_0^\infty S(\lambda)\overline{y}(\lambda)d\lambda \qquad (29)$$
$$Z = \int_0^\infty S(\lambda)\overline{z}(\lambda)d\lambda$$

The functions $\overline{x}(\lambda)$, $\overline{y}(\lambda)$, and $\overline{z}(\lambda)$ are plotted in **Figure 10B**. Based on the color one sees, one cannot define $S(\lambda)$ precisely, but certainly we learn something about the distribution of wavelengths. In the same way, based on a set of relaxation rate constants, we cannot fully define $(1 - S^2)\theta(z)$, but certainly we can learn something about the dynamics. The matching forms of **Eqs. 27**, **29** further highlight the relationship between these problems.

The XYZ color space can be represented as a 2D space, shown in **Figure 10A**. Only x and y are shown, and z is selected so that $x + y + z = 1$ (then, a third dimension would vary this sum, corresponding to brightness). By marking points in the color space, one can indicate how the color space may be represented in another basis. Here, we have marked points corresponding to red,

green, and blue of the sRGB standard (Anderson et al., 1996). Colors within the resulting triangle may be obtained with positive linear combinations of the red, green, and blue of sRGB, so that this triangle is a good estimate of colors that may be obtained with a color monitor (which creates color by combining red, green and blue pixels–this means that in **Figure 10A**, colors outside the triangle are not correctly represented on your screen). These points also define a transformation from the XYZ color matching functions (**Figure 10B**) to the sRGB functions (**Figure 10C**). Note that any color may be represented in the sRGB space, but only those where $S(\lambda)$ results in positive R, G, and B values can actually be reproduced by a typical monitor.

Realizing that the mathematics of relaxation rate constants was essentially equivalent to color spaces, we created analogous relaxation rate constant spaces, replacing the X, Y, and Z values with normalized rate constants. However, instead of placing points within the relaxation rate space, we surrounded the space in **Figure 10D**, since we wanted to describe all points in the space with *positive* parameters. Interestingly, by surrounding the space as closely as possible, without crossing into the space, we obtained a transformation to functions with well-separated and non-negative sensitivities, see **Figure 10F**. In the example here, we use three points to transform the three experimental sensitivities into detector sensitivities, resulting in three detectors. However, redundancy in the information of larger data sets often results in the space becoming narrow in a given dimension, so that the full space may also be approximately described using fewer points, resulting in fewer detectors than experimental data points, but better signal-to-noise in the resulting parameters. Full details of this approach are described in Smith et al. (2018).

Optimizing Detector Sensitivities: Automated Approach

Investigating the relaxation rate space is a powerful way to grasp the information content of a relaxation data set, however, detector optimization using this method requires manual selection of points in the space. This quickly became excessively tedious for large data sets, as is the case for analysis of relaxometry data (Smith A. A. et al., 2021), so that we have also automated the optimization of linear combination (Smith et al., 2019a).

For automation, one still has the requirements that we capture the information in the experiments (that is, we can fit the data), while minimizing the number of parameters to describe that data, and second, that we obtain detector sensitivities that are narrow and non-negative. The first requirement may be met using singular value decomposition (Golub and Kahan, 1965). Suppose we have a matrix, **M**, for which each row is a sensitivity of one of our experiments ($R_\zeta(z)$), where we perform a normalization to prioritize fitting of higher quality data (procedure: first, we normalize all sensitivities to a maximum of one, second we multiply the sensitivity by the median of the experimental rate constants, and third we divide by the median standard deviation of those rate constants). Each column then corresponds to a correlation time. For N experiments, we obtain the best approximation of **M** that can be achieved with a linear combination of t vectors, defined by

$$\begin{aligned} \mathbf{M} &\approx \tilde{\mathbf{M}} = \mathbf{U}_t \cdot \Sigma_t \cdot \mathbf{V}'_t \\ \mathbf{V}'_t &= \Sigma_t^{-1} \cdot \mathbf{U}'_t \cdot \mathbf{M} \end{aligned} \tag{30}$$

The t rows of $\mathbf{V}'_t$ are linear combinations of the rows of **M**, with recombination defined by the product $\Sigma_t^{-1} \cdot \mathbf{U}'_t \cdot \mathbf{M}$ ($\mathbf{U}_t$, $\mathbf{V}'_t$ are unitary matrices, and Σ_t is diagonal, with the largest n singular values along the diagonal). Linear combination of the rows of $\mathbf{V}'_t$ to yield the rows of **M** is an approximate relationship, but the inverse, recombination of the rows of **M** to yield $\mathbf{V}'_t$, is exact. Then, the closer $\tilde{M}$ is to **M**, the better the data can be fit, but this requires t to be larger, and thus more noise is also present in the final analysis. In principle, this linear recombination could be directly applied to the experimental data, to obtain detectors with sensitivities given by the rows of the $\mathbf{V}'_t$. The result would capture (approximately) the maximum amount of information possible from the experiment with t parameters. However, the sensitivities found in the rows of $\mathbf{V}'_t$ are not narrow, and usually have large negative regions. On the other hand, a linear recombination of the vectors in $\mathbf{V}'_t$ would maintain the information content and fit quality, but allows one to optimize the detector sensitivities to be separated and non-negative.

$$\begin{bmatrix} \rho_1(z) \\ \rho_2(z) \\ \vdots \end{bmatrix} = \mathbf{T} \cdot \mathbf{V}'_t = \mathbf{T} \cdot \Sigma_t^{-1} \cdot \mathbf{U}'_t \cdot \mathbf{M} \tag{31}$$

Then, **T** defines the linear recombination of the $\mathbf{V}'_t$ to yield the $\rho_n(z)$, where **T** is a square matrix. The product of a row of **T** with $\mathbf{V}'_t$ defines one of the detectors sensitivities, $\rho_n(z)$. A row of **T** is determined in order to optimize a detector sensitivity, first by choosing a single correlation time, $z_{\max} = \log_{10}(\tau_c/\mathrm{s})$, for which we optimize a linear combination of the rows of $\mathbf{V}'_t$ such that $\rho_n(z_{\max}) = 1$, while simultaneously minimizing $\rho_n(z)$ for all other correlation times, and requiring that all $\rho_n(z)$ remain non-negative. This can be quickly solved using a linear programming algorithm (Kantorovich, 1960; Dantzig, 1982; Virtanen, 2020). However, if we sweep through an array of correlation times, performing this optimization at each correlation time, we find that we are only successful at t correlation times (we consider the minimization as having failed if for some z, we find that $\rho_n(z)$ exceeds 1). Currently, we find the best t detectors by sweeping over a large array of correlation times (200), although this algorithm could be improved to reduce the number of optimizations required (spaces method and automated method both implemented in MATLAB, download from https://difrate.sourceforge.io).

In the detector analysis, once we have optimized the detectors, we apply the same linear combination to the experimental relaxation rate constants as were applied to the sensitivities in order to obtain optimized detector responses. Note in practice that this is implemented as a fit, allowing one to prioritize fitting relaxation rate constants with lower measurement error. Furthermore, we place bounds on the fitted detector responses, $\rho_n^{(\theta,S)}$. In *A Few Notes on Linearity*, we noted that bounds (priors)

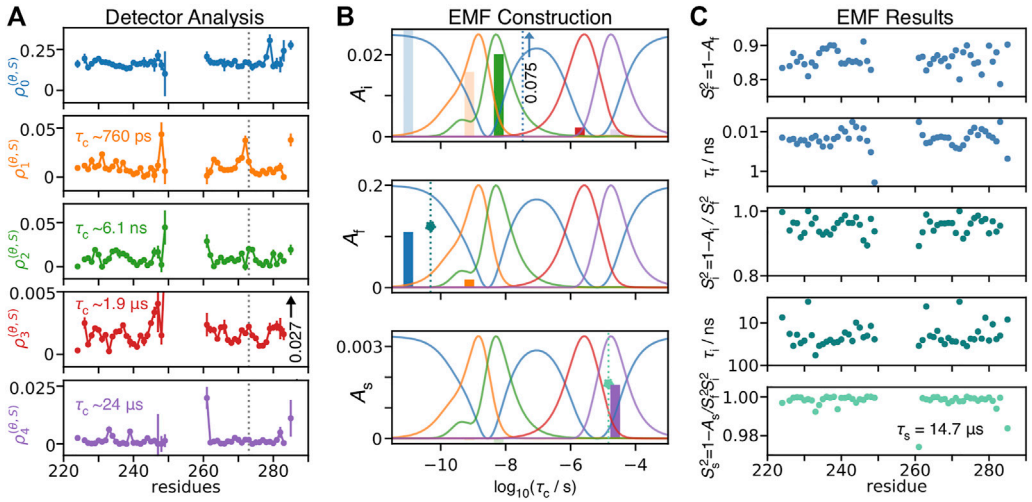


FIGURE 11 | Model-free analysis from detectors. **(A)** shows a detector analysis of HET-s (218–289) fibrils (Smith et al., 2016), with sensitivities shown in **(B)** (amplitude scale not shown; sensitivities have a maximum of 1). **(B)** illustrates the procedure to convert 273Ser detector responses into model-free parameters. Bars give the detector responses (*y*-axis), plotted at the center of the corresponding detector's sensitivity (*x*-axis, note that ρ_0, blue, does not have a well-defined center). At top, we find the ratio of $\rho_3^{(\theta,S)}/\rho_2^{(\theta,S)}$ is consistent with a correlation time of 34 ns, with corresponding amplitude of 0.075 (intermediate motion). After subtracting the contribution of this correlation time to $\rho_0^{(\theta,S)}$ (middle), we find the ratio $\rho_1^{(\theta,S)}/\rho_0^{(\theta,S)}$ is consistent with a correlation time of 49 ps, and amplitude of 0.12 (fast motion). Using a fixed correlation time of 14.7 μs, we find an amplitude for the slow motion of 1.8×10^{-3} (bottom). **(C)** shows the results of EMF analysis for all residues using the procedure in **(B)**.

on the fit parameters can cause the fit parameters to not be linear to $(1-S^2)\theta(z)$. This is only the case if the priors exclude the best fit. Detectors are constructed such that any *allowed* set of relaxation rate constants will not result in parameters that violate the priors. Allowed rate constants are any set that may occur for an arbitrary form of $(1-S^2)\theta(z)$. If, due to noise or measurement error, a dis-allowed set of relaxation rate constants is measured, then the priors will force the fitted relaxation rate constants to fall in the allowed space.

MODEL-FREE, OR NOT?

We see that the original model-free approach, SDM, LeMaster's approach, and detector analysis all belong to a family of methods that yield parameters with well-defined relationships to the distribution of correlation times, here defined by $(1-S^2)\theta(z)$. For SDM and detectors, the final parameters ($J(\omega)$, $\rho_n^{(\theta,S)}$) are linearly related to $(1-S^2)\theta(z)$; for model-free, S^2 and $(1-S^2)\langle\tau_e\rangle$ are linear, and for LeMaster's approach, $(1-S_f^2)$, $S_f^2(1-S_H^2)$, and $S_f^2 S_H^2(1-S_N^2)$ are linear, whereas the final parameters (S^2, $\langle\tau_e\rangle$, S_f^2, S_H^2, and S_N^2) must be obtained via additional arithmetic operations. Response of EMF parameters, on the other hand, may react to changes in one motion differently, depending on other motions in the system. Still, its simplicity in analysis and interpretation—one to three pairs of correlation times and amplitudes—makes it an attractive choice for relaxation data analysis. Should we then compromise in some cases, and sacrifice well-defined parameters for more easily interpreted parameters?

Case 1: Extended Model-Free

Using detectors, we may better understand how EMF parameters in solid-state NMR depend on amplitudes of motion for particular windows of correlation times. We re-analyze relaxation data of HET-s (218–289) fibrils (Smith et al., 2016), by first performing a detector analysis on the data, shown in **Figure 11A** and then iteratively fitting detector responses to correlation times and amplitudes in **Figure 11B**, resulting in the EMF analysis in **Figure 11C**.

Using the following procedure, we are able to reproduce our previous model-free results, illustrated in **Figure 11B** for residue 273Ser. The procedure is given below as a set of simple equations, where results are a good reproduction of our previous direct fit using the model-free approach.

Step 1: z_i

$$\frac{\rho_2^{(\theta,S)}}{\rho_3^{(\theta,S)}} = \frac{\rho_2(z_\mathrm{i})}{\rho_3(z_\mathrm{i})}$$

Step 2: A_i

$$A_\mathrm{i} = \frac{\rho_2^{(\theta,S)}}{\rho_2(z_\mathrm{i})} = \frac{\rho_3^{(\theta,S)}}{\rho_3(z_\mathrm{i})}$$

Step 3: z_i

$$\frac{\rho_0^{(\theta,S)} - A_\mathrm{i}\rho_0(z_\mathrm{i})}{\rho_1^{(\theta,S)} - A_\mathrm{i}\rho_1(z_\mathrm{i})} = \frac{\rho_0(z_\mathrm{f})}{\rho_1(z_\mathrm{f})}$$

Step 4: A_f

$$A_\mathrm{f} = \frac{\rho_0^{(\theta,S)} - A_\mathrm{i}\rho_0(z_\mathrm{i})}{\rho_0(z_\mathrm{f})} = \frac{\rho_1^{(\theta,S)} - A_\mathrm{i}\rho_1(z_\mathrm{i})}{\rho_1(z_\mathrm{f})}$$

Step 5: A_s

$$A_\mathrm{f} = \frac{\rho_4^{(\theta,S)} - A_\mathrm{i}\rho_4(z_\mathrm{i}) - A_\mathrm{f}\rho_4(z_\mathrm{f})}{\rho_4(z_\mathrm{s})}$$

$$S_\mathrm{f}^2 = 1 - A_\mathrm{f} \qquad \tau_\mathrm{f} = 10^{z_\mathrm{f}} \cdot 1\,\mathrm{s}$$

$$S_\mathrm{i}^2 = 1 - A_\mathrm{i}/S_\mathrm{f}^2 \qquad \tau_\mathrm{i} = 10^{z_\mathrm{i}} \cdot 1\,\mathrm{s}$$

$$S_\mathrm{s}^2 = 1 - A_\mathrm{s}/\left(S_\mathrm{f}^2 S_\mathrm{i}^2\right) \qquad \tau_\mathrm{s} = 10^{z_\mathrm{s}} \cdot 1\,\mathrm{s} \tag{32}$$

In the first and second steps, we find a correlation time for which the ratio of sensitivities of ρ_2 and ρ_3 matches the ratio of the detector responses, and then subsequently find the correct amplitude to reproduce these correlation times. With $\rho_2^{(\theta,S)} = 2.0 \times 10^{-2}$, and $\rho_3^{(\theta,S)} = 2.2 \times 10^{-3}$, we find $\tau_i = 34$ ns. Our first concern with this fit is that the intermediate correlation time, $z_i = \log_{10}(\tau_i/\mathrm{s})$, is a compromise between a detector sensitive to motions around 6 ns and a second sensitive around 2 µs. It seems unlikely that the same motion can really explain these two detector responses, which have sensitivities separated by three orders of magnitude. The second problem is because we use a compromise correlation time, both detector sensitivities are very low at this correlation time, which must be counterbalanced by using a large amplitude (A_i) in the model-free fit. Then, in our example, $A_i = 0.075$ is significantly larger than the detector responses, $\rho_2^{(\theta,S)}$ and $\rho_3^{(\theta,S)}$, from which it results, so that we are very likely overestimating the amplitude of this motion.

In the third and fourth steps, we subtract the contributions from z_i and A_i from $\rho_0^{(\theta,S)}$ and $\rho_1^{(\theta,S)}$, and similarly use the ratios of the remainder of the detector responses to obtain z_f, and their amplitudes to obtain A_f. Again, it is not clear if these detectors should be treated as if they describe a single motion. In particular, the relatively uniform behavior of $\rho_0^{(\theta,S)}$ likely is a result of primarily local librational motion, which will not be described by the same amplitudes and correlation times of motions leading to greater variation in $\rho_1^{(\theta,S)}$. Interestingly, because the amplitudes do not vary in the same way, the variation in amplitude of $\rho_1^{(\theta,S)}$ cannot be reproduced in the trends for A_f, but instead has to be fitted by variation in correlation time (τ_f). The result is that amplitude trends in $S_i^2 = 1 - A_i/S_f^2$, shown in **Figure 11(C**, middle) correlate well with trends in τ_f, especially near breaks between the β-sheets of HET-s (near 235Glu, 271Gly). However, this correlation is actually coming from similar amplitude trends observed for $\rho_1^{(\theta,S)}$ and $\rho_2^{(\theta,S)}$. The corresponding detector sensitivities are centered at 760 ps and 6.1 ns, and in fact overlap, suggesting that they may describe the same or at least related motions. EMF attributes these detector responses to different motions, having median correlation times of 22 ps and 42 ns (taken over all residues), thus being separated by three orders of magnitude.

In the final step, one fixes the slow correlation time to 14.7 µs (based on a fit optimization over the whole data set). In this case, the amplitude of $\rho_4^{(\theta,S)}$ determines A_s alone; the proximity of 14.7 µs to the center of ρ_4 (24 µs) results in fairly reasonable amplitudes (for 273Ser, $\rho_4^{(\theta,S)}$ and A_s fall within rounding error, yielding 1.8×10^{-3}).

Then, the major problems with this EMF analysis are intermediate correlation times falling within the NMR blind spot (~20–600 ns), along with correspondingly inflated amplitudes, as well as similar problems due to fitting fast correlation times to $\rho_0^{(\theta,S)}$ and $\rho_1^{(\theta,S)}$, which requires a compromise correlation time between librational motions (~ps) with nanosecond motions. Furthermore, this behavior prevented comparison of EMF parameters for HET-s to MD results, whereas detectors yielded reasonable agreement (Smith et al., 2019b). As we have previously pointed out (Smith et al., 2017), the model-free parameters in HET-s fibrils are far from being atypical, in fact they are fairly consistent across multiple protein systems, likely due most studies utilizing similar data sets and analysis methodology (Chevelkov et al., 2009b; Schanda et al., 2010; Haller and Schanda, 2013; Zinkevich et al., 2013; Lamley et al., 2015a).

Case 2: Model-Free Analysis of µs-Motion

Microsecond motion is the result of processes having higher free-energy cost than nanosecond and picosecond dynamics. We suggest dividing these motions into local and collective motions, where the free energy cost of local motions comes from higher amplitude motions (~10°) that require traversing a large energy barrier. In contrast, collective motions tend to be very low amplitude motion, where the high free-energy cost of the motion is not due to large amplitude dynamics or a significant energy barrier, but rather diffusive dynamics involving large numbers of atoms. Such dynamics are characterized by modes of motion, where a continuum of possible correlation lengths leads to a distribution of correlation times. In contrast, some local microsecond dynamics can be reasonably well approximated as a hopping motion between two orientations, and therefore described with a single correlation time (although effort should be made to determine whether relaxation might be due to multi-site exchange, and understand how this changes the interpretation of data analysis).

Local Dynamics

The availability of $R_{1\rho}$ data, including formulas for its analysis (Trott and Palmer, 2002; Abergel and Palmer, 2003; Miloushev and Palmer, 2005; Kurbanov et al., 2011; Rovo and Linser, 2017) and improving methods for its collection (Kurauskas et al., 2017; Lakomek et al., 2017; Keeler et al., 2018; Krushelnitsky et al., 2018) has recently resulted in considerable improvement in the ability to characterize local micro- to millisecond motions (Rovó, 2020). We consider two categories of $R_{1\rho}$ experiments: the first is Bloch-McConnell relaxation dispersion experiments (BMRD), for which $R_{1\rho}$ relaxation is the result of motion modulating the isotropic chemical shift, and the NEar Rotary-resonance Relaxation Dispersion (NERRD, (Kurauskas et al., 2017)), for which orientational changes in anisotropic tensors leads to $R_{1\rho}$ relaxation. For two-site exchange, BMRD $R_{1\rho}$ relaxation rate constants depend on exchange rate ($k_{ex} = 1/\tau_c$), the change in isotropic chemical shift due to exchange ($\Delta\omega_{12}$), and the population ($p_1, p_2 = 1 - p_1$). Rate constants further depend on the effective field strengths corresponding to each of the two chemical shifts, ω_{e1} and ω_{e2}, as well as the effective field for the average chemical shift. Palmer and coworkers provide us with the following expression (Trott and Palmer, 2002; Trott et al., 2003; Miloushev and Palmer, 2005), which is valid in the fast or intermediate exchange regimes:

$$R_{1\rho} = \frac{R_1}{2}\left(1+\cos^2\beta_e\right) + R_{1\rho}^{\mathrm{DD,CSA}} + R_{1\rho}^{\mathrm{ex}}$$
$$R_{1\rho}^{\mathrm{ex}} = \frac{\sin^2\beta_e p_1 p_2 \Delta\omega_{12}^2 k_{\mathrm{ex}}}{\frac{\omega_{e1}^2\omega_{e2}^2}{\omega_e^2} + k_{\mathrm{ex}}^2 - \sin^2\beta_e p_1 p_2 \Delta\omega_{12}^2\left(1+\frac{2k_{\mathrm{ex}}^2\left(p_1\omega_{e1}^2+p_2\omega_{e2}^2\right)}{\omega_{e1}^2\omega_{e2}^2+\omega_e^2 k_{\mathrm{ex}}^2}\right)}$$
$$\Omega = p_1\Omega_1 + p_2\Omega_2$$
$$\omega_e^2 = \omega_1^2 + \Omega^2$$
$$\omega_{e1}^2 = \omega_1^2 + \Omega_1^2, \omega_{e2}^2 = \omega_2^2 + \Omega_2^2 \tag{33}$$

The total $R_{1\rho}$ relaxation has contributions from longitudinal relaxation (R_1), transverse relaxation from dipole and CSA tensors ($R_{1\rho}^{\mathrm{DD,CSA}}$), and from chemical exchange ($R_{1\rho}^{\mathrm{DD,CSA}}$). (Kurbanov et al., 2011) give the formula for $R_{1\rho}^{\mathrm{DD,CSA}}$.

$$R_{1\rho}^{\mathrm{DD,CSA}} = \sin^2\beta_e \times \left[\left(\frac{\delta}{4}\right)^2\left(J(\omega_S) + \frac{1}{3}J(2\omega_r-\omega_e) + \frac{2}{3}J(\omega_r-\omega_e) + \frac{2}{3}J(\omega_r+\omega_e) + \frac{1}{3}J(2\omega_r+\omega_e)\right) + \frac{2}{27}(\omega_I\Delta\sigma_I)^2\left(\frac{1}{2}J(2\omega_r-\omega_e) + J(\omega_r-\omega_e) + J(\omega_r+\omega_e) + \frac{1}{2}J(2\omega_r+\omega_e)\right)\right] \tag{34}$$

ω_r is the magic angle spinning frequency, and ω_e is the effective field as defined above. If one assumes the microsecond dynamics are dominated by two-site hoping, the spectral density is given simply by

$$J(\omega) = \frac{2}{5}3p_1p_2\left(1-\cos^2\theta\right)\frac{k_{\mathrm{ex}}}{k_{\mathrm{ex}}^2+\omega^2} = \frac{2}{5}\left(1-S^2\right)\frac{\tau_c}{1+(\omega\tau_c)^2} \tag{35}$$

Then, the question is, how may we most efficiently extract the exchange rate ($k_{\mathrm{ex}} = 1/\tau_c$), populations ($p_1, p_2$), chemical shift changes ($\Delta\omega_{12}$), and angle changes (θ). Fitting of k_{ex} has been fairly well established using both NERDD or BMRD (Trott et al., 2003; Ma et al., 2014; Rovo and Linser, 2017; Marion et al., 2019), and combining both methods should improve the accuracy of the resulting k_{ex}. However, separation of populations from either θ (NERDD) or $\Delta\omega_{12}$ (BMRD) is non-trivial. Supposing we already know k_{ex}, a given experiment's relaxation rate constant then depends on the populations and either θ or $\Delta\omega_{12}$ (at sufficiently fast MAS, a given effective field usually results in either $R_{1\rho}^{\mathrm{DD,CSA}}$ or $R_{1\rho}^{\mathrm{ex}}$ being dominant, although in principle both terms are active in the same experiments). Inspecting **Eqs. 34**, **35** we note that terms p_1, p_2, and θ, only appear once as a product of terms, $3p_1p_1\left(1-\cos^2\theta\right)$. Then, based on NERRD data alone, these parameters are inseparable. This is seen in **Figure 12**, where we plot $R_{1\rho}$ as a function of p_1 and θ. We also calculate $R_{1\rho}^{\mathrm{DD,CSA}}$ specifically for $p_1 = 0.25$ and $\theta = 16^\circ$, and then indicate all other positions resulting in the same value of $R_{1\rho}^{\mathrm{DD,CSA}}$ **Figures 12A,B** as a black contour. In **Figures 12E,F**, we only show contours where $R_{1\rho}^{\mathrm{DD,CSA}}$ matches the value obtained for $p_1 = 0.25$ and $\theta = 16^\circ$, but show several different experimental conditions (varying the field strength, $\nu_1 = \omega_1/2\pi$). Because this results in identical contours, we are unable to disentangle these parameters based on NERRD experiments under different conditions.

In contrast, $R_{1\rho}$ relaxation resulting from chemical exchange has a more complex dependence on the various parameters. In particular, effective fields for each of the two states in exchange, ω_{e1} and ω_{e2} depend on the different offsets, Ω_1, Ω_2, but do not depend on the populations, in principle making the terms separable. Indeed, several plots in **Figure 12** show that different experimental conditions lead to different contours for p_1 vs. $\Delta\omega_{12}$ (contours correspond to $R_{1\rho}$ that is equal to $R_{1\rho}$ obtained for $p_1 = 0.25$ and $\Delta\omega_{12} = 500$ Hz, where contour intersections yield the input values). We are then able to identify the critical conditions required for separating population from chemical shift change. First, we see that if $k_{\mathrm{ex}} \gg \Delta\omega_{12}$, contours are fully overlapped so that we are not able to separate the terms, shown in **Figure 12G**. This is because, in **Eq. 33**, k_{ex}^2 must be much larger than the last term in the denominator. If it is also larger than $\omega_{e1}^2\omega_{e2}^2/\omega_e^2$, then the critical dependence of the $R_{1\rho}$ on ω_{e1} or ω_{e2} is lost. In case k_{ex}^2 is not larger than $\omega_{e1}^2\omega_{e2}^2/\omega_e^2$, then the effective field must be much larger than $\Delta\omega_{12}$, so that this term converges on ω_e^2, again losing dependence on ω_{e1} and ω_{e2} (i.e. the denominator simplifies to $\omega_e^2 + k_{\mathrm{ex}}^2$ (Trott and Palmer, 2002)). In any case, if the effective fields become large, $\omega_{e1} \to \omega_e$, $\omega_{e2} \to \omega_e$, similarly preventing separation in terms. For example, see **Figures 12J,K**, where a large offset or large field strength on the spin-locking field results in overlapping contours. Finally, note that we require a frequency offset to be applied in order to obtain the sign of $\Delta\omega_{12}$. If no frequency offset is applied, then all contours are symmetric as in **Figure 12H**.

Separability occurs only when k_{ex}, $\Delta\omega_{12}$, and ω_e are of similar size. Restricting ω_e is particularly challenging in solid-state NMR, where coherent effects may contribute to relaxation when the spin-locking field becomes too small (Öster et al., 2019). One approach would be to use increasing spinning frequencies (Penzel et al., 2015; Lakomek et al., 2017), although we note that some of the most clear improvements in **Figure 12** occur in **Figure 12L**, where the field strength is only a few times bigger than the H–N *J*-couplings, which cannot be averaged by spinning.

In case we are in the fast exchange limit for BMRD experiments, we are left only with the terms $p_1p_2\left(1-\cos^2\theta\right)$ from NERRD experiments and $p_1p_2\Delta\omega_{12}^2$ from BMRD experiments. In this case, there is little to be done to fully separate population from the other parameters. If values of $\Delta\omega_{12}$ may be bounded, it is then possible to also bound p_1p_2, and therefore one finds a restricted range for possible values of θ (the reverse approach also works). However, if we are in the range of intermediate exchange, then we may separate populations from $\Delta\omega_{12}$, and use the result to also obtain θ (note that inclusion of NERDD data should additionally improve the accuracy of k_{ex}, which in turn improves separation of $\Delta\omega_{12}$ from populations based on the BMRD data). Note that Marion et al. have recently presented similar arguments (Marion et al., 2019), although separation of terms was apparently achieved by combining NERRD and BMRD data for fairly fast exchange (k_{ex}=18,000 s^{-1}, $\Delta\omega_{12}/2\pi$ = 240 Hz). While we agree that using both data sets together is beneficial, the information to separate

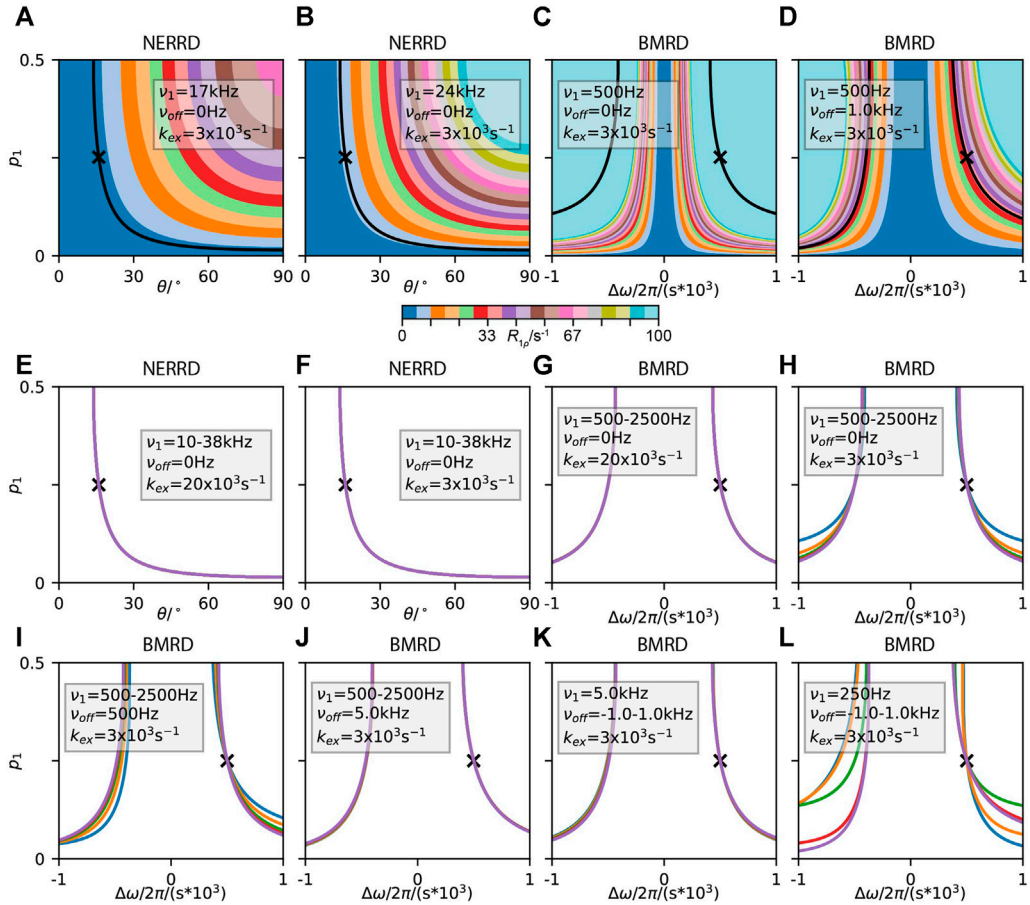


FIGURE 12 | Separating population from hop angle and change in chemical shift in NERRD and BMRD experiments. Relevant parameters are shown as insets ($\omega_r/2\pi$=40 kHz for NERDD plots). In **(A–D)**, contour plots are shown for NERDD and BMRD relaxation rate constants under various conditions, and in each plot, a contour shows all values of p_1 and θ or $\Delta\omega_1$ that yield $R_{1\rho}$ equal to the value obtained for p_1 = 0.25 and θ = 16° or $\Delta\omega_1$ = 500 Hz (marked as a cross on each plot). In **(E–L)**, we only show the contour, but for a range of experimental conditions (five experiments, linearly spaced, with range indicated in the plot). In some cases, this yields nearly identical contours, such that we only see one of the five contours.

population must come from the BMRD data and this is only possible in the intermediate exchange regime (Marion et al. calculated $R_{1\rho}$ for a set of conditions, and via a coarse grid search, were able to find the initial conditions, however, other solutions along contours as in our **Figure 12** likely were overlooked in the grid search).

A final consideration when analyzing BMRD and NERRD data is whether a two-site exchange model is reasonable. In a true two-site exchange, all moving residues should have identical exchange rates and populations, but differing $\Delta\omega_{12}$ and θ values. Then, validation of the two-site model could be achieved by independently analyzing all residues and establishing that all fits have approximately the same p_1, p_2, and k_{ex} (or just the same k_{ex} if populations cannot be determined). In case the true behavior is, for example, three-site exchange, fitting to the two-site exchange model will yield exchange rates that are a weighted average of the two non-zero eigenvalues of the 3 × 3 exchange matrix, where weighting will depend on the chemical shifts of the three sites and/or the angles sampled. In this case, it may be appropriate to apply a three-site exchange model, while jointly fitting all residues using a common set of rate constants (four to six independent parameters, depending on the model chosen). Such an approach has been demonstrated using CPMG relaxation in solution-state NMR (Korzhnev et al., 2005; Neudecker et al., 2006), with the general equations solved for CPMG (Koss et al., 2018).

Collective Dynamics

NERDD relaxation also appears throughout the whole protein in the absence of BMRD relaxation, depending on sample conditions, and is attributed to low amplitude rocking of the whole protein. This is observed very weakly in GB1 crystals (Krushelnitsky et al., 2018), and strongly in GB1 complexed with IgG (Lamley et al., 2015b), HET-s (218–289) (Smith et al., 2016), ubiquitin crystals with amplitude depending heavily on crystal form (Ma et al., 2015; Kurauskas et al., 2017; Lakomek et al., 2017), and SH3 (Krushelnitsky et al., 2018). The apparent global nature of this motion led all of these studies, with the exception of Lakomek et al., to fit $R_{1\rho}$ relaxation using a slow motion with a single correlation time for

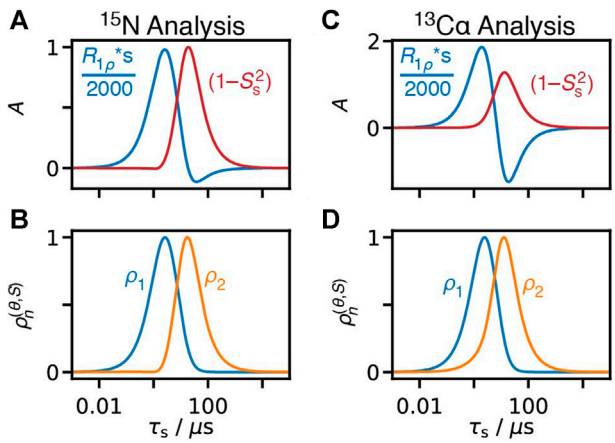


FIGURE 13 | Behavior of fitting $R_{1\rho}$ data to an offset and a fixed correlation time. **(A)** shows the offset term, $R^0_{1\rho}$, divided by 2000, and the order parameter for the slow correlation time, $(1-S^2_s)$ resulting from fitting calculated relaxation rate constants as a function of correlation time to **Eq. 36**. Experiments are ^{15}N $R_{1\rho}$ acquired with MAS frequency of 60 kHz and spin-lock strengths of 11, 16, 25, 38, and 51 kHz, and τ_s is fixed at 18.5 μs. **(B)** shows detector sensitivities optimized using the same data set. **(C)** shows $R^0_{1\rho}$ divided by 2000 and $(1-S^2_s)$ for ^{13}C $R_{1\rho}$ acquired with MAS frequency of 60 kHz and spin-lock strengths of 9, 18, 35, and 48 kHz, as well as an additional experiment with MAS frequency of 40 kHz and spin-lock strength of 25 kHz τ_s is fixed at 7.0 μs. **(D)** shows detector sensitivities optimized using the same data set.

all residues. In our HET-s analysis, we proposed fitting $R_{1\rho}$ data using a global, slow correlation time, where the corresponding order parameter could vary, and additionally an offset term that would account for faster motion that could not be fully parameterized from $R_{1\rho}$ data alone. Kurauskas et al. also followed this procedure, whereas Krushelnitsky and coworkers included explicit fitting of an additional fast motion with a distribution of correlation times. By including an offset term, and using a single correlation time globally, we again have a linear fit.

$$R_{1\rho} = \underbrace{R^0_{1\rho}}_{\text{ns motions}} + \left(1-S^2_s\right) \underbrace{R_{1\rho}(\tau_s)}_{\mu\text{s motion,fixed}} \tag{36}$$

Then, for each residue, $R^0_{1\rho}$ and $(1-S^2_s)$ are varied, where $R^0_{1\rho}$ in principle compensates for relaxation due to fast, nanosecond motion, and $(1-S^2_s)$ should determine the effective amplitude of the global motion, with correlation time τ_s, on the given residue. Practically, what happens is that the $R_{1\rho}$ rate constants measured for a given residue have certain ratios. If those ratios match the ratios calculated for τ_s, then $R^0_{1\rho} = 0$ and the relaxation rate constants are fitted only with $(1-S^2_s)$. In contrast, if all rate constants are approximately equal, then $(1-S^2_s) = 0$ and $R^0_{1\rho}$ accounts for the full relaxation. However, in most cases, the ratios are closer to one than predicted by τ_s, but not exactly one and so by including contributions from $R^0_{1\rho}$ and $(1-S^2_s)R_{1\rho}(\tau_s)$, the data may be fit. One may investigate in more detail how the two terms vary as a function of correlation time (as in **Figures 6–8**). We show the behavior for the ^{15}N and ^{13}Cα $R_{1\rho}$ data sets found in Smith et al. (2016) in **Figure 13**.

In **Figure 13A**, we calculate $R_{1\rho}$ relaxation rate constants for ^{15}N relaxation, and fit to **Eq. 36**. $(1-S^2_s)$ reaches a maximum of approximately one at 19 μs, so that this parameter describes motion at and around the fixed correlation time of τ_s = 18.5 μs. On the other hand, the offset term, $R^0_{1\rho}$, actually is most sensitive at 2.5 μs, far from fitting primarily fast, nanosecond motion. We see that the functional forms are similar to detector sensitivities optimized from the same data set, **Figure 13B**. In **Figure 13C**, the behavior is less ideal: $(1-S^2_s)$ reaches a maximum of 1.28 at 13 μs, somewhat removed from the fixed correlation time of 7.0 μs, and the offset term becomes negative for correlation times around 18 μs.

The sensitivity of the offset term in **Figure 13A** to motion near 2.5 μs as opposed to faster motions may be surprising, although perhaps it should not be. NERRD experiments are most sensitive in the μs-range of correlation times, and rate constants under different experimental conditions have nearly converged to the same value at 1.9 μs (all rate constants within 5% of each other)–only slightly faster than the 2.5 μs where we find the maximum. Then, we would expect the offset term to be sensitive both near where $R_{1\rho}$ is most sensitive, but also near where it converges, which is roughly what we find.

It is then important to note that fitting $R_{1\rho}$ to contributions from an offset term and a fixed correlation time results in an offset term that is most sensitive not to fast (nanosecond) motions, but rather to slower (microsecond) motions. In some cases, $(1-S^2_s)$ may be overly sensitive to some correlation times, with sensitivity exceeding one at positions that are removed from τ_s. Detectors are also a better choice for characterizing broad distributions of correlation times, if one does not know the form of the distribution. In fact, we suspect that global rocking motion is the result of collective dynamics over varying length scales, where increasing the correlation length also increases the correlation time, and therefore yields a broad distribution of correlation times. We demonstrated the relationship between correlation length and correlation time window for HET-s fibrils on the nanosecond timescale using a combination of NMR and MD simulation (Smith et al., 2019b), however, the question remains whether similar behavior can fully explain rocking motion of crystalline proteins; for example, Schanda and coworkers argue that a coupling between overall rocking motion and local loop motion may exist in crystalline ubiquitin (Kurauskas et al., 2017).

OUTLOOK: COMBINING METHODS

We have seen that relaxation data in NMR may be processed by a variety of different methods, however, only some of these methods can really be thought of as "model-free," such that we can establish a well-defined (linear) behavior for each parameter as a function of correlation time, independent of the actual model of the correlation function. These methods are the original model-free analysis, under the assumption that $\omega\tau_i \ll 1$, spectral density mapping, LeMaster's approach, and detector analysis. Of these, only detector analysis is generally applicable to solid-state NMR.

So, are detectors the last word in NMR dynamics analysis? We certainly hope not. Each detector response provides a "window"

into the total reorientational motion of some NMR tensor, with the window width and center defined by $\rho_n(z)$. Still, such a description is not very precise: a moderate detector response could result from a low amplitude motion near where $\rho_n(z)$ reaches its maximum, it could result from a high amplitude motion where $\rho_n(z)$ is small, or (and we suspect this is often the case), it characterizes a distribution of correlation times that overlaps the range of sensitivity of that detector. A collection of detectors, and their behavior as a function of position in a molecule gives further hints at the dynamics of a molecule, but leaves much to be desired in terms of details of motion. What we would rather have is better models of motion. If we use a *good* model, based on knowledge of the dynamics obtained from other methods, the information added to our experimental data should improve our interpretation of the experiment.

Molecular dynamics simulation is particularly powerful as a complimentary method to NMR. One obtains positions of all atoms as a function of time, allowing first, the direct calculation of the NMR-relevant correlation functions, and second, in principle allowing one to connect those correlation functions to specific motion in the molecule. $C(t)$ is explicitly calculated as

$$\begin{aligned} C(t_n) &= \frac{1}{N}\sum_{i=0}^{N-n-1} P_2\Big(\vec{\mu}(\tau_i)\cdot\vec{\mu}(\tau_{i+n})\Big) \\ &\approx S^2 + \left(1-S^2\right)\int_{-\infty}^{\infty}\theta(z)\underbrace{e^{-t_n/(10^z\cdot 1\,\mathrm{s})}}_{R_{C(t_n)}(z)}\,dz \end{aligned} \tag{37}$$

This is the discrete form of **Eq. 14**, as would be applied to an MD trajectory. To obtain the *n*th time point in the correlation function, $C(t_n)$, we simply average over all pairs of frames separated by *n* frames. The latter equation is our assumed form for the correlation function, where we note that a given time point of the correlation function, $C(t_n)$, is related to the distribution of correlation times with the same functional form as the relaxation rate constants (excepting the offset, S^2, see **Eq. 27**). This allows one to calculate detectors from the collection of time points in MD-derived correlation functions using a procedure nearly identical to that described in *Optimizing Detector Sensitivities: Automated Approach*, where the sensitivity, $R_\zeta(z)$ (**Eq. 27**), is replaced by the term $R_{C(t_n)}(z) = \exp\left(-t_n/(10^z\cdot 1\,\mathrm{s})\right)$. In fact, not only may detector analysis be easily modified to analyze MD-derived data, but it is a general approach to numerically solving the inverse Laplace transform, which avoids some of the pitfalls of more common regularization approaches (Tikhonov and Arsenin, 1977).

When analyzing MD with detectors, one has two options: find the optimal set of detectors for describing correlation time distributions found with MD (that is, as many as possible with good signal-to-noise, and as narrow/non-overlapping as possible), or optimize the detectors to match some or all of the NMR-derived detectors. The latter approach is shown in **Figure 14**, where sensitivities of seven NMR experiments in **Figure 14A** are optimized to yield five detectors in **Figure 14C**, and the linear combination used to yield $\rho_2(z)$ is explicitly illustrated in **Figure 14B**. From MD, time points in the correlation function may also be linearly combined (sensitivities for 11 time points shown in **Figure 14D**), to match the NMR-derived detectors **Figures 14E,F**. Note that in **Figure 14F**, the linear combination is a very good match for ρ_1 and ρ_2, with moderate success for ρ_0, but detector sensitivities in the microsecond range are badly reproduced. The detector optimization indicates (correctly) that a 1 μs trajectory cannot reasonably predict dynamics in the range of several microseconds (ρ_3, ρ_4 in red, violet), thus providing a means for determining what information can and cannot be compared across methods. Where sensitivities agree, *quantitative* comparison of dynamics in MD and NMR is possible. Note that in **Figures 14D–F**, we only show 11 time points for illustrative purposes, but this procedure is equally valid for ~10^6 time points (for such a long correlation function, calculating **Eq. 37** takes much longer than evaluating its result with detectors).

The detector analysis then provides a very reliable means of comparing NMR results to MD simulation. The ability to easily compare results across multiple methods is one of the primary advantages of detector analysis. We should note that carefully executed fitting of MD-derived correlation functions, followed by calculation of relaxation rate constants should yield similarly reliable rate constants, if the trajectory is sufficiently long (Mollica et al., 2012). However, the rate constants themselves are sensitive to a broader range of correlation times than detectors, so that the comparison has lower timescale resolution than detectors.

With MD and NMR data sets, one may then use NMR data via detectors (or relaxation rate constants) as a means of validating the MD, and potentially refining it; methods include selecting sections of trajectories that best reproduce experiment (Salvi et al., 2016), selecting the best force fields for a system (Antila et al., 2021), or validating the refinement of a force-field itself (Hoffmann et al., 2018a; Hoffmann et al., 2018b). One may also use NMR data (specifically order parameters) as a means of directing the simulation, so that the simulation returns parameters matching the experiment (Hansen et al., 2014). One should note that a major challenge of combining NMR and MD data is that, while NMR is highly sensitive to microsecond motions, for example, via $R_{1\rho}$ measurements, it is challenging to obtain accurate dynamics on the microsecond timescale from MD simulations. Although MD simulations now regularly extend for multiple microseconds, or longer via enhanced sampling (Bernardi et al., 2015), one still lacks sufficient statistics to obtain reliable dynamics behavior. Consider, if we investigate a 1 μs motion, using a 10 μs trajectory, we should observe 10 events, but the variance in number of events is also 10 (assuming Poisson statistics), so that large errors easily occur. Additionally, correct replication of slower motions requires all the faster motion leading up to the slow motion to occur at approximately the correct rates, so that the slower motions are more susceptible to influences like force field inaccuracies, starting structure of the system, etc. This remains a significant challenge for combining experiment in simulation, requiring creative solutions to take advantage of simulation where reproduction of experimental observables is poor.

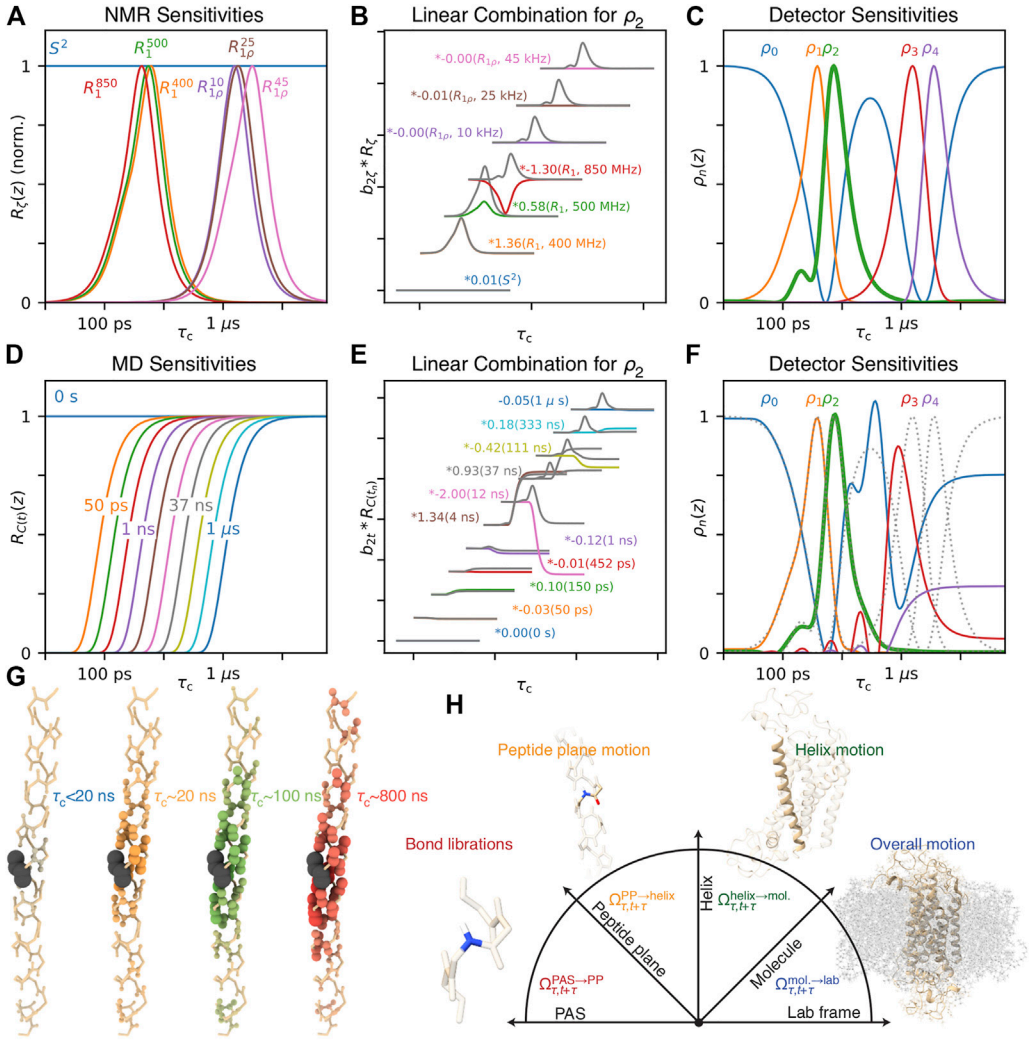


FIGURE 14 | Combining NMR and MD. **(A)** plots normalized NMR sensitivities for a selection of experiments (S^2, ^{15}N R_1 at 400, 500, 850 MHz, ^{15}N $R_{1\rho}$ at 850 MHz, 60 kHz MAS, ν_1 at 10, 25, and 45 kHz). **(B)** shows a linear combination of the normalized sensitivities (*x*, *y* positions shifted to reduce plot overlap), which yields the sensitivity of ρ_2, shown in **(C)** (green, bold). In color are the weighted contributions from each rate constant, and grey shows the cumulative sum (summing all sensitivities at and below the grey line). **(C)** shows the five sensitivities optimized from NMR data. **(D)** plots sensitivities of time points from MD-derived correlation functions (0 s, 10 points log-spaced from 50 ps to 1 μs). **(E)** shows a linear combination of those sensitivities, optimized to match the sensitivity of ρ_2 (*x, y* positions shifted to reduce plot overlap). **(F)** shows detectors optimized to match the NMR-derived detectors in **(C)**. **(G)** shows spatial correlation of motion in a helix as a function of correlation time (windows for <20, ~20, ~100, ~800 ns). Color intensity and bond radii indicate the correlation coefficient between that residue's H–N motion and the motion of the black residue. **(H)** illustrates frames used to separate transformation from the PAS to the lab frame into four steps: a peptide plane frame, a helix frame, and a molecule frame (illustration inspired by Brown (1996), molecule plots created with ChimeraX (Pettersen et al., 2021)).

With experimental validation or refinement of an MD simulation, one may analyze the simulation further, with improved confidence of the accuracy of the simulation. However, we want to use the simulation specifically to improve our interpretation of the experimental parameters. For example, we recently showed that it was possible to calculate the spatial correlation of motions within a given detector window between different residues in HET-s (218–289) fibrils (Smith et al., 2019b), using a modified iRED analysis (Prompers and Brüschweiler, 2001; Prompers and Brüschweiler, 2002). The result is that we could see that detector windows corresponding to longer correlation times tended to result in correlation over longer distances, providing at least some explanation for the presence of slow, low amplitude motion in fibrils. A similar correlation analysis is shown in **Figure 14G**, in this case for residues in an α-helix, where similarly, detector windows corresponding to longer correlation times yield longer correlation lengths. We suspect this behavior to be nearly universal: even in well-defined structures, there is always some residual flexibility. Then, both short- and long-range modes of motion should be thermally populated (in terms of modes, these are more

accurately described as having short and long wavelengths). However, the longer-range modes usually have longer correlation times, resulting in the trends in **Figure 14G**. Note that this implies that there should almost always be distributions of correlation times due to varying correlation length, further complicating the interpretation of the two to three correlation times provided by the EMF approach.

For fairly rigid regions of a molecule, we expect detector-specific correlation analysis to help explain dynamic trends. However, what should we do for regions that are more mobile, with multiple types of motion contributing? Having all of the atom positions in an MD simulation should provide the detail that would allow us to separate different motions. Then, we could define the total motion of a bond as resulting from the product of these motions. For example, for an H–N dipole coupling in an α-helix, the total rotation of the dipole is the result of the reorientation of the principal axis system (PAS) of the dipole within the peptide plane (PP), the peptide plane reorienting within the helix, the helix reorienting with the molecule, and the molecule reorienting within the lab frame.

$$\begin{aligned}\vec{v}(t+\tau) &= \mathbf{R}\left(\Omega_{\tau,t+\tau}\right)\cdot\vec{v}(\tau)\\ &= \mathbf{R}\left(\Omega_{\tau,t+\tau}^{\text{mol.}\to\text{lab}}\right)\cdot\mathbf{R}\left(\Omega_{\tau,t+\tau}^{\text{helix}\to\text{mol.}}\right)\cdot\mathbf{R}\left(\Omega_{\tau,t+\tau}^{\text{PP}\to\text{helix}}\right)\cdot\mathbf{R}\left(\Omega_{\tau,t+\tau}^{\text{PAS}\to\text{PP}}\right)\cdot\vec{v}(\tau)\end{aligned} \tag{38}$$

This concept is illustrated in **Figure 14H**. In the case that it is possible to derive a correlation function from each rotation, one then may effectively achieve an *in silico* model-free type separation of the correlation functions motion. A similar approach for the specific separation of librations, φ/ψ reorientation, and peptide plane tumbling in intrinsically disorded proteins has been demonstrated by Salvi et al. (2017), however we find that it is possible to fully generalize this concept for separation of arbitrary definitions of independent motions (manuscript under revision, (Smith et al., 2021b)). Then, separated motions may also be analyzed with detectors, to determine how both experimental and simulated detector responses depend on both timescale and position in the molecule. Separation of motions could also be coupled with mode analyses such as iRED (Prompers and Brüschweiler, 2001, 2002) or principal component analysis (Amadei et al., 1993; Altis et al., 2007), providing a method to better characterize distributions of correlation times arising from different motions and complex mode-like dynamics. In each proposed case, comparison of the different MD analyses is possible via the detector analysis. Our eventual goal is that one may extract enough detail from the MD to build explicit models of motion for direct application to the NMR experimental results, so that the final characterizations are no longer model-free at all, but rather yield highly detailed models based on the combined information from experiment and simulation.

CONCLUSION

We show that the original model-free approach, SDM, LeMaster's approach, and detectors all belong to a class of methods where fit parameters are resulting from a linear combination of experimental relaxation rate constants (potentially requiring an additional arithmetic step to yield the final parameters). IMPACT is a close approximation to this behavior, whereas EMF parameters exhibit significantly different behavior. Analysis methods belonging to this class are particularly useful because it is straightforward to estimate the resulting parameters if the distribution of correlation times, $(1-S^2)\theta(z)$, is known. This is particularly advantageous when determining if a model is consistent with experimentally determined parameters, and also allows easy comparison of multiple methods.

The detector analysis is the most general of these approaches, being applicable to any collection of NMR relaxation experiments probing reorientational motion, and can be generalized for other methods such as MD simulation, requiring very little modification of the analysis. Then, the resulting detector responses from NMR and MD are easily compared. With experimental validation of MD, one may then use the wealth of detail in MD simulation to better understand how experimentally derived parameters are related to specific motion, via correlation of motion, separation of motion, and other existing and yet-to-be developed techniques. This has the potential to lead to improved models of motion for NMR analysis, which in turn can help obtain a more fundamental understand of dynamics in biomolecular systems.

AUTHOR CONTRIBUTIONS

AS has prepared the manuscript text and KZ has prepared the figures.

FUNDING

The authors acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG) grant SM 576/1-1 and by European Social Funds (ESF) and the Free State of Saxony (Junior Research Group UniDyn, Project No. SAB 100382164). The authors also acknowledge support from the German Research Foundation (DFG) and Universität Leipzig within the program of Open Access Publishing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.727553/full#supplementary-material

REFERENCES

Abergel, D., and Palmer, A. G. (2003). On the Use of the Stochastic Liouville Equation in Nuclear Magnetic Resonance: Application to R1ρ Relaxation in the Presence of Exchange. *Concepts Magn. Reson.* 19A, 134–148. doi:10.1002/cmr.a.10091

Altis, A., Nguyen, P. H., Hegger, R., and Stock, G. (2007). Dihedral Angle Principal Component Analysis of Molecular Dynamics Simulations. *J. Chem. Phys.* 126, 244111. doi:10.1063/1.2746330

Amadei, A., Linssen, A. B. M., and Berendsen, H. J. C. (1993). Essential Dynamics of Proteins. *Proteins* 17, 412–425. doi:10.1002/prot.340170408

Anderson, M., Motta, R., Chandrasekar, S., and Stokes, M. (1996). "Proposal for a Standard Default Color Space for the Internet—sRGB," in Color and Imaging Conference (Society for Imaging Science and Technology), 238–245.

Antila, H. S., Ferreira, T. M., Ollila, O. H. S., and Miettinen, M. S. (2021). Using Open Data to Rapidly Benchmark Biomolecular Simulations: Phospholipid Conformational Dynamics. *J. Chem. Inf. Model.* 61, 938–949. doi:10.1021/acs.jcim.0c01299

Beckmann, P. A. (1988). Spectral Densities and Nuclear Spin Relaxation in Solids. *Phys. Rep.* 171, 85–128. doi:10.1016/0370-1573(88)90073-7

Bernardi, R. C., Melo, M. C. R., and Schulten, K. (2015). Enhanced Sampling Techniques in Molecular Dynamics Simulations of Biological Systems. *Biochim. Biophys. Acta (Bba) - Gen. Subjects* 1850, 872–877. doi:10.1016/j.bbagen.2014.10.019

Brown, M. (1982). Theory of Spin-Lattice Relaxation in Lipid Bilayers and Biological Membranes. 2H and 14N Quadrupolar Relaxation. *J. Chem. Phys.* 77 (3), 1576–1599. doi:10.1063/1.443940

Brown, M. F. (1996). "Biological Membranes," in *Biological Membranes: A Molecular Perspective From Computation And Experiment*. Editors K. M. Merz and B. Roux (Birkhäuser Basel). doi:10.1007/978-1-4684-8580-6

Chevelkov, V., Fink, U., and Reif, B. (2009a). Accurate Determination of Order Parameters from 1H,15N Dipolar Couplings in MAS Solid-State NMR Experiments. *J. Am. Chem. Soc.* 131, 14018–14022. doi:10.1021/ja902649u

Chevelkov, V., Fink, U., and Reif, B. (2009b). Quantitative Analysis of Backbone Motion in Proteins Using MAS Solid-State NMR Spectroscopy. *J. Biomol. NMR* 45, 197–206. doi:10.1007/S10858-009-9348-5

Clore, G. M., Szabo, A., Bax, A., Kay, L. E., Driscoll, P. C., and Gronenborn, A. M. (1990). Deviations from the Simple Two-Parameter Model-free Approach to the Interpretation of Nitrogen-15 Nuclear Magnetic Relaxation of Proteins. *J. Am. Chem. Soc.* 112, 4989–4991. doi:10.1021/ja00168a070

d'Auvergne, E. J., and Gooley, P. R. (2003). The Use of Model Selection in the Model-free Analysis of Protein Dynamics. *J. Biomol. NMR* 25, 25–39. doi:10.1023/A:1021902006114

Dantzig, G. B. (1982). Reminiscences about the Origins of Linear Programming. *Operations Res. Lett.* 1, 43–48. doi:10.1016/0167-6377(82)90043-8

Gill, M. L., Byrd, R. A., and Palmer, A. G. (2016). Dynamics of GCN4 Facilitate DNA Interaction: a Model-free Analysis of an Intrinsically Disordered Region. *Phys. Chem. Chem. Phys.* 18, 5839–5849. doi:10.1039/c5cp06197k

Golub, G., and Kahan, W. (1965). Calculating the Singular Values and Pseudo-inverse of a Matrix. *J. Soc. Ind. Appl. Maths. Ser. B Numer. Anal.* 2, 205–224. doi:10.1137/0702016

Gullion, T., and Schaefer, J. (1989). Rotational-Echo Double-Resonance NMR. *J. Magn. Reson. (1969)* 81, 196–200. doi:10.1016/0022-2364(89)90280-1

Halle, B., and Wennerström, H. (1981). Interpretation of Magnetic Resonance Data from Water Nuclei in Heterogeneous Systems. *J. Chem. Phys.* 75, 1928–1943. doi:10.1063/1.442218

Haller, J. D., and Schanda, P. (2013). Amplitudes and Time Scales of Picosecond-To-Microsecond Motion in Proteins Studied by Solid-State NMR: a Critical Evaluation of Experimental Approaches and Application to Crystalline Ubiquitin. *J. Biomol. NMR* 57, 263–280. doi:10.1007/s10858-013-9787-x

Hansen, N., Heller, F., Schmid, N., and van Gunsteren, W. F. (2014). Time-averaged Order Parameter Restraints in Molecular Dynamics Simulations. *J. Biomol. NMR* 60, 169–187. doi:10.1007/s10858-014-9866-7

Hoffmann, F., Mulder, F. A. A., and Schäfer, L. V. (2018a). Accurate Methyl Group Dynamics in Protein Simulations with AMBER Force Fields. *J. Phys. Chem. B.* 122, 5038–5048. doi:10.1021/acs.jpcb.8b02769

Hoffmann, F., Xue, M., Schäfer, L. V., and Mulder, F. A. A. (2018b). Narrowing the gap between Experimental and Computational Determination of Methyl Group Dynamics in Proteins. *Phys. Chem. Chem. Phys.* 20, 24577–24590. doi:10.1039/c8cp03915a

Hsu, A., O'Brien, P. A., Bhattacharya, S., Rance, M., and Palmer, A. G. (2018). Enhanced Spectral Density Mapping through Combined Multiple-Field Deuterium 13CH2D Methyl Spin Relaxation NMR Spectroscopy. *Methods* 138-139, 76–84. doi:10.1016/j.ymeth.2017.12.020

Ishima, R., Louis, J. M., and Torchia, D. A. (1999). Transverse 13C Relaxation of CHD2 Methyl Isotopmers to Detect Slow Conformational Changes of Protein Side Chains. *J. Am. Chem. Soc.* 121, 11589–11590. doi:10.1021/ja992836b

Judd, D. B. (1951). *Report of U.S. Secretariat Committee on Colorimetry and Artificial Daylight, Proceedings of the Twelfth Session of the CIE*. Stockholm: Paris: Bureau Central de la CIE, 11.

Kantorovich, L. (1960). On the Calculation of Production Inputs. *Probl. Econ.* 3, 3–10. doi:10.2753/PET1061-199103013

Kay, L. E., Torchia, D. A., and Bax, A. (1989). Backbone Dynamics of Proteins as Studied by Nitrogen-15 Inverse Detected Heteronuclear NMR Spectroscopy: Application to Staphylococcal Nuclease. *Biochemistry* 28, 8972–8979. doi:10.1021/bi00449a003

Keeler, E. G., Fritzsching, K. J., and McDermott, A. E. (2018). Refocusing CSA during Magic Angle Spinning Rotating-Frame Relaxation Experiments. *J. Magn. Reson.* 296, 130–137. doi:10.1016/j.jmr.2018.09.004

Khan, S. N., Charlier, C., Augustyniak, R., Salvi, N., Déjean, V., Bodenhausen, G., et al. (2015). Distribution of Pico- and Nanosecond Motions in Disordered Proteins from Nuclear Spin Relaxation. *Biophysical J.* 109, 988–999. doi:10.1016/j.bpj.2015.06.069

Kinosita, K., Kawato, S., and Ikegami, A. (1977). A Theory of Fluorescence Polarization Decay in Membranes. *Biophysical J.* 20, 289–305. doi:10.1016/S0006-3495(77)85550-1

Korzhnev, D. M., Neudecker, P., Mittermaier, A., Orekhov, V. Y., and Kay, L. E. (2005). Multiple-Site Exchange in Proteins Studied with a Suite of Six NMR Relaxation Dispersion Experiments: An Application to the Folding of a Fyn SH3 Domain Mutant. *J. Am. Chem. Soc.* 127, 15602–15611. doi:10.1021/ja054550e

Koss, H., Rance, M., and Palmer, A. G. (2018). General Expressions for Carr-Purcell-Meiboom-Gill Relaxation Dispersion for N-Site Chemical Exchange. *Biochemistry* 57, 4753–4763. doi:10.1021/acs.biochem.8b00370

Krushelnitsky, A., Gauto, D., Rodriguez Camargo, D. C., Schanda, P., and Saalwächter, K. (2018). Microsecond Motions Probed by Near-Rotary-Resonance R1ρ 15N MAS NMR Experiments: the Model Case of Protein Overall-Rocking in Crystals. *J. Biomol. NMR* 71, 53–67. doi:10.1007/s10858-018-0191-4

Kurauskas, V., Izmailov, S. A., Rogacheva, O. N., Hessel, A., Ayala, I., Woodhouse, J., et al. (2017). Slow Conformational Exchange and Overall Rocking Motion in Ubiquitin Protein Crystals. *Nat. Commun.* 8, 145. doi:10.1038/s41467-017-00165-8

Kurbanov, R., Zinkevich, T., and Krushelnitsky, A. (2011). The Nuclear Magnetic Resonance Relaxation Data Analysis in Solids: General R1/R1ρ Equations and the Model-free Approach. *J. Chem. Phys.* 135 (1–9), 184104. doi:10.1063/1.3658383

Lakomek, N.-A., Penzel, S., Lends, A., Cadalbert, R., Ernst, M., and Meier, B. H. (2017). Microsecond Dynamics in Ubiquitin Probed by Solid-State 15N NMR Spectroscopy R1ρ Relaxation Experiments under Fast Mas (60-110 Khz). *Chem. Eur. J.* 23, 9425–9433. doi:10.1002/chem.201701738

Lamley, J. M., Lougher, M. J., Sass, H. J., Rogowski, M., Grzesiek, S., and Lewandowski, J. R. (2015a). Unraveling the Complexity of Protein Backbone Dynamics with Combined 13C and 15N Solid-State NMR Relaxation Measurements. *Phys. Chem. Chem. Phys.* 17, 21997–22008. doi:10.1039/c5cp03484a

Lamley, J. M., Öster, C., Stevens, R. A., and Lewandowski, J. R. (2015b). Intermolecular Interactions and Protein Dynamics by Solid-State NMR Spectroscopy. *Angew. Chem. Int. Ed.* 54, 15374–15378. doi:10.1002/anie.201509168

LeMaster, D. (1995). Larmor Frequency Selective Model Free Analysis of Protein NMR Relaxation. *J. Biomol. NMR* 6, 366–374. doi:10.1007/BF00197636

Lipari, G., and Szabo, A. (1982a). Model-free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. *J. Am. Chem. Soc.* 104, 4546–4559. doi:10.1021/ja00381a009

Lipari, G., and Szabo, A. (1982b). Model-free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 2. Analysis of Experimental Results. *J. Am. Chem. Soc.* 104, 4559–4570. doi:10.1021/ja00381a010

Ma, P., Haller, J. D., Zajakala, J., Macek, P., Sivertsen, A. C., Willbold, D., et al. (2014). Probing Transient Conformational States of Proteins by Solid-State R1ρ Relaxation-Dispersion NMR Spectroscopy. *Angew. Chem. Int. Ed.* 53, 4312–4317. doi:10.1002/anie.201311275

Ma, P., Xue, Y., Coquelle, N., Haller, J. D., Yuwen, T., Ayala, I., et al. (2015). Observing the Overall Rocking Motion of a Protein in a Crystal. *Nat. Commun.* 6, 8361. doi:10.1038/ncomms9361

Mandel, A. M., Akke, M., and Palmer, A. G. (1995). Backbone Dynamics of Escherichia coli Ribonuclease HI: Correlations with Structure and Function in an Active Enzyme. *J. Mol. Biol.* 246, 144–163. doi:10.1006/jmbi.1994.0073

Marion, D., Gauto, D. F., Ayala, I., Giandoreggio-Barranco, K., and Schanda, P. (2019). Microsecond Protein Dynamics from Combined Bloch-McConnell and Near-Rotary-Resonance R1p Relaxation-Dispersion MAS NMR. *ChemPhysChem* 20, 276–284. doi:10.1002/cphc.201800935

Mendelman, N., and Meirovitch, E. (2021). Structural Dynamics from NMR Relaxation by SRLS Analysis: Local Geometry, Potential Energy Landscapes, and Spectral Densities. *J. Phys. Chem. B.* 125, 6130–6143. doi:10.1021/acs.jpcb.1c02502

Miloushev, V. Z., and Palmer, A. G. (2005). R1ρ Relaxation for Two-Site Chemical Exchange: General Approximations and Some Exact Solutions. *J. Magn. Reson.* 177, 221–227. doi:10.1016/j.jmr.2005.07.023

Mollica, L., Baias, M., Lewandowski, J. R., Wylie, B. J., Sperling, L. J., Rienstra, C. M., et al. (2012). Atomic-Resolution Structural Dynamics in Crystalline Proteins from NMR and Molecular Simulation. *J. Phys. Chem. Lett.* 3, 3657–3662. doi:10.1021/Jz3016233

Munowitz, M. G., Griffin, R. G., Bodenhausen, G., and Huang, T. H. (1981). Two-dimensional Rotational Spin-echo Nuclear Magnetic Resonance in Solids: Correlation of Chemical Shift and Dipolar Interactions. *J. Am. Chem. Soc.* 103, 2529–2533. doi:10.1021/ja00400a007

Neudecker, P., Korzhnev, D. M., and Kay, L. E. (2006). Assessment of the Effects of Increased Relaxation Dispersion Data on the Extraction of 3-site Exchange Parameters Characterizing the Unfolding of an SH3 Domain. *J. Biomol. NMR* 34, 129–135. doi:10.1007/s10858-006-0001-2

Öster, C., Kosol, S., and Lewandowski, J. R. (2019). Quantifying Microsecond Exchange in Large Protein Complexes with Accelerated Relaxation Dispersion Experiments in the Solid State. *Sci. Rep.* 9, 11082. doi:10.1038/s41598-019-47507-8

Palmer, A. G. (2004). NMR Characterization of the Dynamics of Biomacromolecules. *Chem. Rev.* 104, 3623–3640. doi:10.1021/cr030413t

Peng, J. W., and Wagner, G. (1992). Mapping of Spectral Density Functions Using Heteronuclear NMR Relaxation Measurements. *J. Magn. Reson. (1969)* 98, 308–332. doi:10.1016/0022-2364(92)90135-T

Penzel, S., Smith, A. A., Agarwal, V., Hunkeler, A., Org, M.-L., Samoson, A., et al. (2015). Protein Resonance Assignment at MAS Frequencies Approaching 100 kHz: a Quantitative Comparison of J-Coupling and Dipolar-Coupling-Based Transfer Methods. *J. Biomol. NMR* 63, 165–186. doi:10.1007/s10858-015-9975-y

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., et al. (2021). UCSF ChimeraX : Structure Visualization for Researchers, Educators, and Developers. *Protein Sci.* 30, 70–82. doi:10.1002/pro.3943

Polimeno, A., and Freed, J. H. (1992). *Advances In Chemical Physics*. John Wiley & Sons, 89–206. doi:10.1002/9780470141410.ch3A Many-Body Stochastic Approach to Rotational Motions in Liquids.

Prompers, J. J., and Brüschweiler, R. (2002). General Framework for Studying the Dynamics of Folded and Nonfolded Proteins by NMR Relaxation Spectroscopy and MD Simulation. *J. Am. Chem. Soc.* 124, 4522–4534. doi:10.1021/ja012750u

Prompers, J. J., and Brüschweiler, R. (2001). Reorientational Eigenmode Dynamics: A Combined MD/NMR Relaxation Analysis Method for Flexible Parts in Globular Proteins. *J. Am. Chem. Soc.* 123, 7305–7313. doi:10.1021/ja0107226

Rovó, P., and Linser, R. (2017). Microsecond Time Scale Proton Rotating-Frame Relaxation under Magic Angle Spinning. *J. Phys. Chem. B.* 121, 6117–6130. doi:10.1021/acs.jpcb.7b03333

Rovó, P. (2020). Recent Advances in Solid-State Relaxation Dispersion Techniques. *Solid State. Nucl. Magn. Reson.* 108, 101665. doi:10.1016/j.ssnmr.2020.101665

Salvi, N., Abyzov, A., and Blackledge, M. (2017). Analytical Description of NMR Relaxation Highlights Correlated Dynamics in Intrinsically Disordered Proteins. *Angew. Chem. Int. Ed.* 56, 14020–14024. doi:10.1002/anie.201706740

Salvi, N., Abyzov, A., and Blackledge, M. (2016). Multi-Timescale Dynamics in Intrinsically Disordered Proteins from NMR Relaxation and Molecular Simulation. *J. Phys. Chem. Lett.* 7, 2483–2489. doi:10.1021/acs.jpclett.6b00885

Schanda, P., and Ernst, M. (2016). Studying Dynamics by Magic-Angle Spinning Solid-State NMR Spectroscopy: Principles and Applications to Biomolecules. *Prog. Nucl. Magn. Reson. Spectrosc.* 96, 1–46. doi:10.1016/j.pnmrs.2016.02.001

Schanda, P., Meier, B. H., and Ernst, M. (2011). Accurate Measurement of One-Bond H-X Heteronuclear Dipolar Couplings in MAS Solid-State NMR. *J. Magn. Reson.* 210, 246–259. doi:10.1016/j.jmr.2011.03.015

Schanda, P., Meier, B. H., and Ernst, M. (2010). Quantitative Analysis of Protein Backbone Dynamics in Microcrystalline Ubiquitin by Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 132, 15957–15967. doi:10.1021/Ja100726a

Schanda, P. (2019). Relaxing with Liquids and Solids - A Perspective on Biomolecular Dynamics. *J. Magn. Reson.* 306, 180–186. doi:10.1016/j.jmr.2019.07.025

Shapiro, Y. E., and Meirovitch, E. (2012). Slowly Relaxing Local Structure (SRLS) Analysis of 15N-H Relaxation from the Prototypical Small Proteins GB1 and GB3. *J. Phys. Chem. B.* 116, 4056–4068. doi:10.1021/jp300245k

Skrynnikov, N. R., Millet, O., and Kay, L. E. (2002). Deuterium Spin Probes of Side-Chain Dynamics in Proteins. 2. Spectral Density Mapping and Identification of Nanosecond Time-Scale Side-Chain Motions. *J. Am. Chem. Soc.* 124, 6449–6460. doi:10.1021/ja012498q

Smith, A. A., Bolik-Coulon, N., Ernst, M., Meier, B. H., and Ferrage, F. (2021a). How Wide Is the Window Opened by High-Resolution Relaxometry on the Internal Dynamics of Proteins in Solution? *J. Biomol. NMR* 75, 119–131. doi:10.1007/s10858-021-00361-1

Smith, A. A., Ernst, M., and Meier, B. H. (2017). Because the Light Is Better Here: Correlation-Time Analysis by NMR Spectroscopy. *Angew. Chem.* 129, 13778–13783. doi:10.1002/ange.201707316

Smith, A. A., Ernst, M., Meier, B. H., and Ferrage, F. (2019a). Reducing Bias in the Analysis of Solution-State NMR Data with Dynamics Detectors. *J. Chem. Phys.* 151, 034102. doi:10.1063/1.5111081

Smith, A. A., Ernst, M., and Meier, B. H. (2018). Optimized "detectors" for Dynamics Analysis in Solid-State NMR. *J. Chem. Phys.* 148, 045104. doi:10.1063/1.5013316

Smith, A. A., Ernst, M., Riniker, S., and Meier, B. H. (2019b). Localized and Collective Motions in HET-s(218-289) Fibrils from Combined NMR Relaxation and MD Simulation. *Angew. Chem.* 131, 9483–9488. doi:10.1002/ange.201901929

Smith, A. A., Testori, E., Cadalbert, R., Meier, B. H., and Ernst, M. (2016). Characterization of Fibril Dynamics on Three Timescales by Solid-State NMR. *J. Biomol. NMR* 65, 171–191. doi:10.1007/s10858-016-0047-8

Smith, A. A., Vogel, A., Engberg, O., Hildebrand, P. W., and Huster, D. (2021b). A Method to Construct the Dynamic Landscape of a Bio-Membrane with experiment and Simulation. *ResearchSquare*, 1–28. doi:10.21203/rs.3.rs-645823/v1

Smith, T., and Guild, J. (1931). The C.I.E. Colorimetric Standards and Their Use. *Trans. Opt. Soc.* 33, 73–134. doi:10.1088/1475-4878/33/3/301

Tikhonov, A. N., and Arsenin, V. I. A. (1977). *Solutions of Ill-Posed Problems*. Washington, DC: Winston and Sons.

Trott, O., Abergel, D., and Palmer, A. G. (2003). An Average-Magnetization Analysis of R1ρ Relaxation outside of the Fast Exchange Limit. *Mol. Phys.* 101, 753–763. doi:10.1080/0026897021000054826

Trott, O., and Palmer, A. G. (2002). R1ρ Relaxation outside of the Fast-Exchange Limit. *J. Magn. Reson.* 154, 157–160. doi:10.1006/jmre.2001.2466

Tugarinov, V., Liang, Z., Shapiro, Y. E., Freed, J. H., and Meirovitch, E. (2001). A Structural Mode-Coupling Approach to 15N NMR Relaxation in Proteins. *J. Am. Chem. Soc.* 123, 3055–3063. doi:10.1021/ja003803v

Virtanen, P. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* 17, 261–272. doi:10.1038/s41592-019-0686-2

Vos, J. J. (1978). Colorimetric and Photometric Properties of a 2° Fundamental Observer. *Color Res. Appl.* 3, 125–128. doi:10.1002/col.5080030309

Wennerström, H., Lindman, B., Soederman, O., Drakenberg, T., and Rosenholm, J. B. (1979). Carbon-13 Magnetic Relaxation in Micellar Solutions. Influence of Aggregate Motion on T1. *J. Am. Chem. Soc.* 101, 6860–6864. doi:10.1021/ja00517a012

Zinkevich, T., Chevelkov, V., Reif, B., Saalwächter, K., and Krushelnitsky, A. (2013). Internal Protein Dynamics on ps to μs Timescales as Studied by Multi-Frequency 15N Solid-State NMR Relaxation. *J. Biomol. NMR* 57, 219–235. doi:10.1007/s10858-013-9782-2

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

REVIEW
published: 11 November 2021
doi: 10.3389/fmolb.2021.761227

NMR Studies of Tau Protein in Tauopathies

*Kristine Kitoka[1], Rostislav Skrabana[2,3], Norbert Gasparik[4,5], Jozef Hritz[4,6] and Kristaps Jaudzems[1,7]**

[1]Laboratory of Physical Organic Chemistry, Latvian Institute of Organic Synthesis, Riga, Latvia, [2]Institute of Neuroimmunology, Slovak Academy of Sciences, Bratislava, Slovakia, [3]AXON Neuroscience R&D Services SE, Bratislava, Slovakia, [4]Central European Institute of Technology, Masaryk University, Brno, Czech Republic, [5]Faculty of Science, National Centre for Biomolecular Research, Masaryk University, Brno, Czech Republic, [6]Department of Chemistry, Faculty of Science, Masaryk University, Brno, Czech Republic, [7]Faculty of Chemistry, University of Latvia, Riga, Latvia

Tauopathies, including Alzheimer's disease (AD), are the most troublesome of all age-related chronic conditions, as there are no well-established disease-modifying therapies for their prevention and treatment. Spatio-temporal distribution of tau protein pathology correlates with cognitive decline and severity of the disease, therefore, tau protein has become an appealing target for therapy. Current knowledge of the pathological effects and significance of specific species in the tau aggregation pathway is incomplete although more and more structural and mechanistic insights are being gained using biophysical techniques. Here, we review the application of NMR to structural studies of various tau forms that appear in its aggregation process, focusing on results obtained from solid-state NMR. Furthermore, we discuss implications from these studies and their prospective contribution to the development of new tauopathy therapies.

Keywords: tau, nuclear magnetic resonance, protein structure, Alzheimer's disease, filaments

OPEN ACCESS

Edited by:
Amir Goldbourt,
Tel Aviv University, Israel

Reviewed by:
Mei Hong,
Massachusetts Institute of Technology, United States
Łukasz Jaremko,
King Abdullah University of Science and Technology, Saudi Arabia

****Correspondence:***
Kristaps Jaudzems
kristaps.jaudzems@osi.lv

Specialty section:
This article was submitted to Structural Biology, a section of the journal Frontiers in Molecular Biosciences

Received: *19 August 2021*
Accepted: *25 October 2021*
Published: *11 November 2021*

Citation:
Kitoka K, Skrabana R, Gasparik N, Hritz J and Jaudzems K (2021) NMR Studies of Tau Protein in Tauopathies. Front. Mol. Biosci. 8:761227. doi: 10.3389/fmolb.2021.761227

INTRODUCTION

Neurodegenerative tauopathies form a large group of heterogeneous incurable diseases characterised by deposits of abnormal forms of tau protein in specific parts of the brain (Kovacs., 2015). The most frequent tauopathy is Alzheimer's disease (AD). The number of patients with AD is estimated to be more than 30 million worldwide and is expected to increase dramatically (World Health Organisation, 2020). AD and other tauopathies lead to severe personality changes, decline of thinking skills and loss of patients' ability to carry out everyday tasks, leaving them fully dependent on medical care until death, which usually occurs 5–10 years after the clinical diagnosis (Reitz et al., 2011). The socioeconomic burden of the disease is enormous due to the protracted disease course and dependence on care of AD patients (Wimo et al., 2010).

Hitherto efforts to develop a therapy for AD or tauopathies in general have had an extremely low success rate in comparison with other chronic conditions (Cummings et al., 2014). Only one new AD treatment (Aducanumab) has been approved since 2003 despite hundreds of clinical trials in the last two decades (Cavazzoni, 2021). Aducanumab is a passive immunotherapy which can remove aggregated forms of amyloid-beta peptide, one of the two hallmarks of AD (Sevigny et al., 2016); however, its ability to slow cognitive decline of patients was not ascertained yet. Other potential therapies in clinical research focus on the protein tau, which forms intracellular neurofibrillary tangles in neurons and is ubiquitous in tauopathies. The tau protein neurofibrillary pathology may appear early in the pre-symptomatic phase of disease and its spatio-temporal distribution correlates with cognitive decline and disease severity (Braak and Braak, 1991; Braak et al., 2011). For these

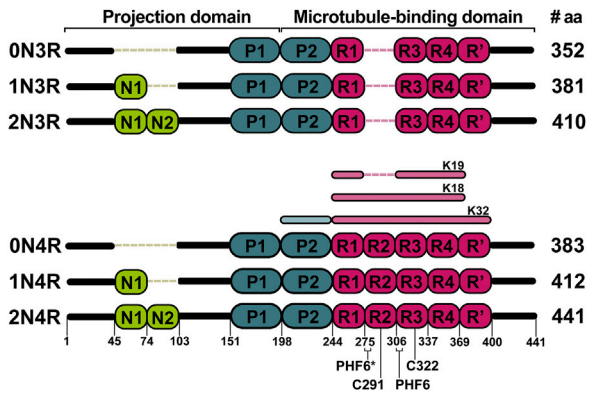


FIGURE 1 | Six major tau isoforms generated by alternative splicing of the MAPT gene. N1 and N2 indicate N-terminal inserts (green). The presence of inserts is encoded by exon 2 and 3. P1 and P2 indicate proline-rich domains (blue). R1 to R4 indicate repeat domains (pink). The R2 is encoded by exon 10. The K18, K19, and K32 represent the most studied truncated constructs of tau both by solution and solid-state NMR. Locations of PHF6*, PHF6, C291, and C322 are shown on the 2N4R isoform.

reasons, tau has become an appealing target for various therapeutic strategies including small-molecule inhibition of tau aggregation and phosphorylation, anti-sense oligonucleotide therapy, passive and active immunotherapy (Li and Götz, 2017; Novak et al., 2018b; Congdon and Sigurdsson, 2018; Jadhav et al., 2019). Particularly, the recently completed Phase 2 clinical trial of AADvac1, an active immunotherapy targeting tau mid-region indicated that therapy may potentially slow cognitive decline in a subgroup of patients with ascertained tau pathology (Novak et al., 2021).

In the brain, tau is expressed as six different isoforms comprising 352–441 residues. The isoforms are generated via alternative splicing of the MAPT gene and contain either zero, one, or two 29-residue inserts at the N-terminal part (0, 1 or 2N isoforms, respectively) and three to four repeat sequences at the C-terminal part (3R or 4R isoforms) (**Figure 1**). Compared to some other fibril-forming proteins, tau is a very soluble protein due to its high content of charged and hydrophilic amino acids (Sibille et al., 2006). Structurally, tau belongs to the class of intrinsically disordered proteins (IDPs), which do not form a stable tertiary fold or secondary structure elements and exist as an ensemble of interconverting conformations (conformational ensemble). Nevertheless, tau has many binding sites that are specific for different partners (Melkova et al., 2019). The protein can additionally undergo a large variety of posttranslational modifications, mostly phosphorylations, glycations and truncations, which modulate its physiological and pathophysiological functions (Zilka et al., 2012). The primary function of tau is to promote assembly and maintain stability of axonal microtubules, however, it has also been implicated in cellular signaling and regulation of other cellular processes (Habchi et al., 2014; Sotiropoulos et al., 2017). Self-assembly of tau is associated with tauopathies. Under normal conditions, tau has a low propensity to form aggregates. However, upon hyperphosphorylation (Alonso et al., 2001), metal ion binding (Jiji et al., 2017; Ahmadi et al., 2019) or truncation (Al-Hilaly et al., 2017; Novak et al., 2018a) tau self-assembles into insoluble paired helical (PHF) or straight filaments (SF), which can contribute to the pathogenesis. Repeat sequences present in the microtubule binding domain of all six isoforms are involved in the filament assembly (Goedert et al., 1988). Thus, the formation of filaments is linked to a reduced ability to bind microtubules.

A major bottleneck for development of new anti-tau therapies is the identification of the most relevant biological target (isoform, posttranslational modification, and aggregation state) to tackle. Understanding the structural and dynamic basis of tau assembly leading to disease is therefore crucial for developing new strategies for the treatment of AD and other tauopathies. The structure of tau filaments from AD patient-isolated material has been recently elucidated by cryo-EM (Fitzpatrick et al., 2017; Falcon et al., 2018). However, the filament structures fail to explain what is the exact basis for tauopathy-specific and structurally different filaments (Zhang et al., 2019). This points to a necessity for integrative structural and mechanistic studies addressing the interplay between truncation, phosphorylation, and aggregation of tau with respect to progression of pathology. NMR spectroscopy is unique in this sense as it can provide structural and functional information on tau's disordered conformational ensembles, aggregated and filamentous states as well as directly probe phosphorylation and its effects.

In this paper, we review solution and solid-state NMR structural and interaction studies of tau in monomeric, oligomeric and filamentous forms. We start by reviewing the wide efforts to characterize secondary structure propensities of monomeric tau by solution NMR, and how it is influenced by phosphorylation or interaction with different partners. The following section describes the few studies of tau aggregation intermediates including oligomers. The last section is devoted to tau filament studies using solid-state NMR, which includes most recent results on *in vitro* tau fibrils generated without any inducer. Finally, we discuss the implications and perspectives of NMR studies to decipher the complex mechanisms of tau aggregation.

MONOMERIC STATE AS THE STARTING POINT OF TAU SELF-ASSEMBLY

Until now, it has been unclear whether aggregation or phosphorylation is the leading event in the process of tau self-assembly into filaments (Lippens et al., 2003; Wegmann et al., 2021). However, for both of these events, the initial object of interest is a protein in the monomeric state. Protein aggregation is driven by a transition from α-helix or random coil to β-sheet structures (von Bergen et al., 2005; Ding et al., 2003; Bibow et al., 2011). Therefore, it is crucial to identify sites in proteins with a propensity to form β-strands, polyproline II helices or other extended structures amenable to self-association (Sillen et al., 2005a).

On the other hand, it is also necessary to explore whether any phosphorylation site may trigger the aggregation of tau. Hence, solution NMR studies of tau in the monomeric state have been

performed with the aim to determine secondary structure propensities of the different tau protein regions and to measure the effects of phosphorylation at specific sites (Mukrasch et al., 2009; Smet et al., 2004; Bibow et al., 2011; Sibille et al., 2012).

NMR Assignment

Assignment of backbone resonances is the first step towards any site-specific studies by NMR. However, a complete assignment of the backbone resonances of tau monomer in solution has been a major bottleneck due to its large molecular size and IDP character. Therefore, initial studies were performed on short peptides comprising fragments of the repeat sequences. Lippens and colleagues were among the first who tried to assign full-length tau. Their motivation was to use the assignment to study the impact of phosphorylation on tau. Unfortunately, using conventional assignment schemes, they were able to assign only a limited set of resonances (Lippens et al., 2003; Smet et al., 2004). Protein size, particular amino acid composition, repetitive regions, all together cause an immense overlap of signals, which makes assignment complicated (Lippens et al., 2006). The following attempts using higher magnetic fields and a 3D heteronuclear experiment setup were more successful, and Mukrasch and co-workers, whose primary motivation was to study structural propensities of tau and its interactions with microtubules and polyanions, presented the backbone assignment (except prolines) of full-length tau (Mukrasch et al., 2009). Later, other groups also succeeded in providing full-length tau assignments by implementing various non-uniform sampling (NUS)-NMR strategies allowing for higher dimensionality experiments (typically 4-7D) (Narayanan et al., 2010; Harbison et al., 2012) and analysis based on comparison with shorter constructs (Harbison et al., 2012). Some of these NUS-NMR strategies also succeeded to assign proline residues, in contrast to the conventional NMR assignment approaches. This is of particular importance for determination of the preferred trans/cis conformational states of individual prolines within tau in non-phosphorylated and a variety of phosphorylated states (Ahuja et al., 2016).

Secondary Structure Propensities

In contrast to other methods such as CD spectroscopy which describe the overall populations of secondary structure elements within the studied protein, solution NMR allows to determine secondary structure propensities (SSP) at the individual residue level by mostly using the chemical shifts of Hα, Cα and Cβ atoms and their differences with respect to random coil values (Marsch et al., 2006).

More than two decades ago, von Bergen and colleagues reported that V306-K311 sequence (PHF6) in the R3 repeat (**Figure 1**) is a minimal tau self-interaction motif supporting the formation of filaments (von Bergen et al., 2000). This observation served as a motivation to study tau R3 by ^{1}H NMR. Analysis of NOESY cross-peak pattern and $^{3}J_{HNH\alpha}$ coupling data in 2,2,2-trifluoroethanol (TFE) suggested that the hexapeptide V306-K311 exhibits an extended structure, L315-L325 exhibits α-helix character, while the remaining part of the repeat remains unstructured. In addition, a possible model for the self-assembly via the helical structure was proposed, where the dimer formation and aggregation are promoted by hydrophilic and hydrophobic interactions, respectively. Contrary to previous findings, the sequence L315-L325 in H_2O did not exhibit α-helix, but L315-S320 adopted an extended structure as V306-K311 (Minoura et al., 2002; Minoura et al., 2003). The possible reason for such a difference could be the tendency of TFE to stabilize α-helices (Shiraki et al., 1995). Unlike the R3, other repeats (R1, R2, and R4) in TFE do not exhibit any β-sheet propensity (Minoura et al., 2004; Tomoo et al., 2005). Deviations of Cα chemical shifts, which serve as sensitive probes for identifying local secondary structures, also identified a 6-residue region Q307-P312 with a β-sheet propensity in the cysteine-free K19 construct, which lacks R2 (**Figure 2**). (Eliezer et al., 2005). Mukrasch and colleagues expanded their interest and studied the K18 construct in addition to K19. Their data suggested that the beginning of each repeat except R1 has a stretch of 10–11 residues with a high β-structure propensity (**Figure 2**). The highest propensity was observed for K274-L284 and S305-L315, enclosing PHF6* and PHF6 hexapeptides (Mukrasch et al., 2005), which have been identified as seeds of filament formation (von Bergen et al., 2000). Moreover, a weak propensity for β-structures was observed in the ends of repeats R2 and R4 (Mukrasch et al., 2005). A later study on full-length tau supported the findings mentioned above (**Figure 2**). Additionally, several stretches with polyproline II helix propensity in the proline-rich regions P1 and P2 were observed, and a random coil character for 343 of 441 residues was confirmed (Mukrasch et al., 2009). Detailed comparison of SSP along the chain of tau and its homologue Map2c protein was also reviewed by Melkova et al. (Melkova et al., 2019).

Phosphorylation Effects

Hyperphosphorylation has been suggested to be sufficient for the induction of tau filament formation (Alonso et al., 2001; Despres et al., 2019). Phosphate groups have the potential to impact the local and global structure of tau and alter its interaction preferences from microtubules to itself. Serines, threonines, and tyrosines, which are potential sites for phosphorylation, comprise nearly one-fifth of tau residues making it an ideal substrate for kinases (Lippens et al., 2003). Over the last two decades NMR has been used to decipher the pattern of tau phosphorylation in both qualitative and quantitative manner (Lippens et al., 2003; Landrieu et al., 2006; Amniai et al., 2009; Amniai et al., 2011; Qi et al., 2016; Despres et al., 2017). The NMR studies were focused on the phosphorylation of residues in the proline-rich domains and repeat regions of the protein (**Figure 3**). Solution NMR has also allowed monitoring the phosphorylation kinetics of multiple residues by multiple kinases. Recent advances in the use of non-uniformly sampled NMR have offered superior time resolution of such processes (Mayzel et al., 2014; Louša et al., 2017). This resulted not only in the efficient comparison of different phosphorylation rates at several sites but also allowed to determine whether any phosphorylation is pre-conditioned by a previous phosphorylation at a different site.

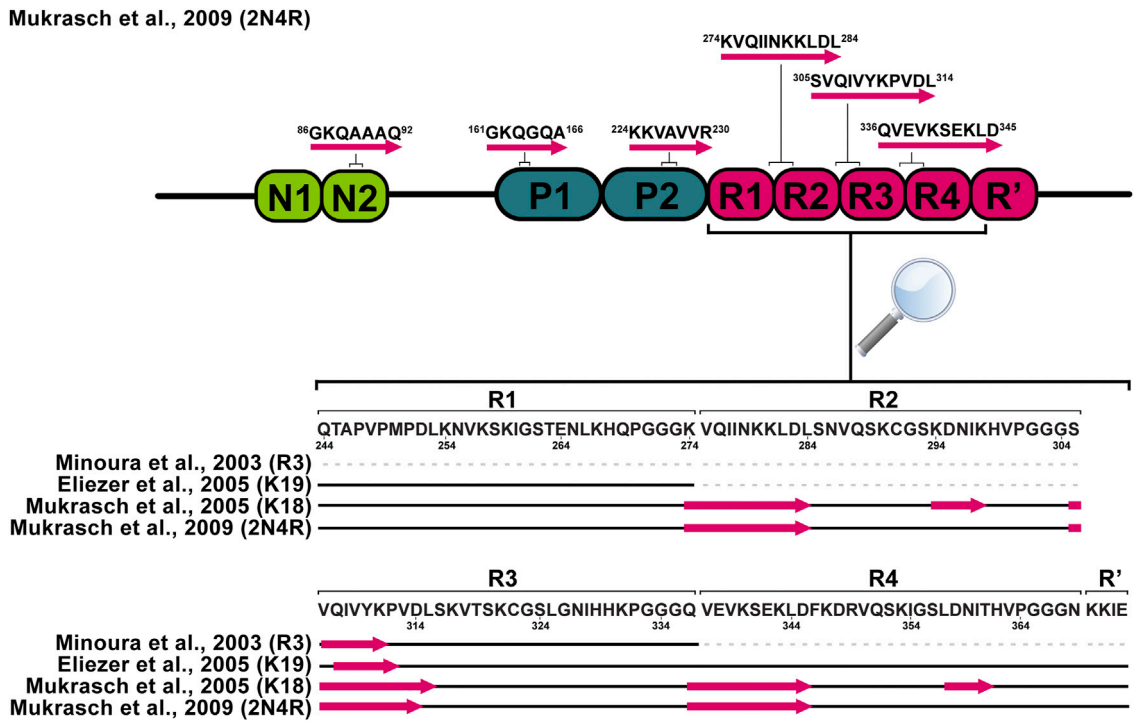


FIGURE 2 | Schematic representation of tau monomer regions with preference for the β-sheet structure derived from solution NMR data. Pink arrows represent regions with increased β-sheet propensity.

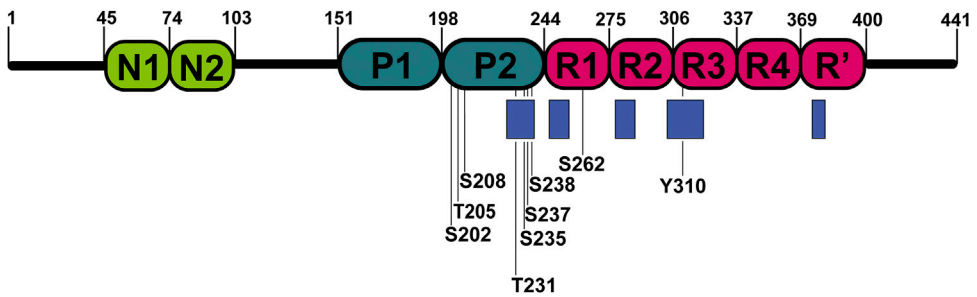


FIGURE 3 | Locations of phosphorylated residues and tau microtubule-binding sites discussed in this review. Blue rectangles represent tau microtubule-binding sites with the strongest affinity.

According to Hα chemical shift deviations, in a short peptide (residues K224-K240) from the proline-rich domain double phosphorylation at T231 and S235 induced β-turn propensity for residues V229–T231 (Daly et al., 2000), while multiple-site phosphorylation induced a structural change to polyproline II helix in engineered proline-rich domain peptides. Moreover, the data derived from $^{3}J_{H\alpha HN}$ suggested that these changes may be involved in the global structural transition of phosphorylated tau to the aggregated form (Bielska and Zondlo, 2006). Likewise, an enhanced propensity for β-turn formation of V229–T231 was also observed in a case where phosphorylation was mimicked by mutations T231E and S235E. In addition, a transient helix between S238 and R242 was found, which does not depend on phosphorylation of T231 but is stabilized by phosphorylation of S235, S237, and S238 (Sibille et al., 2012; Schwalbe et al., 2015). Based on NOE and molecular dynamics (MD) data, a turn conformation was hypothesised as a response to phosphorylation at S202 and T205. The turn is stabilized by a hydrogen bond between the phosphorylated T205 and the amide proton of G207 (Gandhi et al., 2015). The phosphorylation of only S202 and T205 has been suggested to be protective against aggregation. If G209, which forms the stabilizing hydrogen bond, is mutated to valine and combined with phosphorylation at S202, T205, and S208, tau forms filaments without any other aggregation inducer (Despres et al., 2017).

The impact of phosphorylation in the repeat region of tau has been also studied. In the R1 peptide (residues V256-G273), phosphorylation of S262 enhanced the assembly rate of the peptide in comparison to the non-phosphorylated version. The different assembly rates can be explained by differences in conformation between R1 and phospho-R1 (pR1), derived from chemical shift perturbations of HN and Hα between R1 and pR1 (Zhou et al., 2006). In contrast, other studies suggest that phosphorylation of S262 inhibits aggregation of tau (Schneider et al., 1999; Haj-Yahya et al., 2020). The same site was studied in the K18 construct (**Figure 1**); instead of introducing genuine phosphate groups, the phosphorylation effects were mimicked by mutating serine to glutamic acid in the tau repeats. The mutation did not significantly influence the secondary structure propensities of repeats. Although, mutations induced selective conformational changes in the R1 and R2 (Fischer et al., 2009). Phosphorylation of another site Y310 in the R3, associated with the formation of PHFs, was sufficient to delay tau aggregation (Ait-Bouziad et al., 2020). A possible explanation could be that the interaction of I308 and non-phosphorylated Y310 is crucial for forming tau filaments (Naruto et al., 2010; Sogawa et al., 2014).

Interaction Studies

The primary function of tau is to regulate essential functions of microtubules (MTs), including polymerization, stabilization, and modulation of dynamics. Therefore, details about tau binding to microtubules are highly important for understanding these functions (Mukrasch et al., 2005). The reduced affinity of tau to MTs is mediated by genetic mutations or hyperphosphorylation and leads to axonal transport perturbation. When it occurs, tau detaches from the MT complex and tends to aggregate (Kadavath et al., 2015a). On the other hand, specific tau truncation may result in formation of abnormal forms of tau-MT complexes (Novak et al., 2018a). Tau is also able to interact with many other partners, including several molecules that have been implicated in pathogenesis, such as polyanions, metal ions and other amyloid proteins. Hub proteins, for example 14-3-3 proteins, which connect multiple cellular pathways should also be considered. These interactions have been extensively studied using solution NMR.

The overlay of 2D ^{1}H-^{15}N HSQC spectra of free and MT-bound tauF4 (residues S208-S324) revealed the disappearance of a large fraction of the resonances, which was caused by slow tumbling in solution upon binding to the microtubules (Sillen et al., 2007; Kadavath et al., 2015a). The most affected residues were K224–S237, T245–L253, V275–L284, and V300–K317, comprising P2, R1, R2, and R3 parts of the protein. Results were nearly identical for both 3R and 4R tau isoforms, indicating that they share the same attachment mechanism to the MTs (Kadavath et al., 2015a). Chemical shift perturbations measured from 2D ^{1}H-^{15}N HSQC spectra highlighted positively charged lysine and histidine residues in the repeat sequences preceding the PGGG motifs as mediators of the microtubule binding (Mukrasch et al., 2005). The most significant chemical shift changes, pointing to a strong involvement in the MT-binding process, were observed for residues K225-T231 in P2 domain, K240-V248 in R1 domain, V275-S285 and I297-V300 in R2, and K370-K375 in R' (Mukrasch et al., 2007). Other studies have suggested that FTDP-17-associated mutations (ΔK280 and P301L), phosphorylation of S214, and pseudophosphorylation of KXGS motifs significantly attenuate the binding to microtubules (Fischer et al., 2007; Sillen et al., 2007; Fischer et al., 2009). These results are in agreement with the "jaws" model of tau binding whereby the regions flanking the repeats are considered as targeting domains, responsible for positioning and high affinity tau binding on the MTs surface, and the repeats act as catalytic domains for microtubule assembly (Fischer et al., 2007; Kadavath et al., 2018). The results are also in agreement with atomic models of MT-bound tau derived from a combination of cryo-EM data and Rosetta modeling (Kellogg et al., 2018), which indicate that tau attaches to MTs through repeat sequences. However, slightly different residues in R1 were found to be involved in the binding in comparison to NMR results (Kadavath et al., 2015b).

In-cell NMR has over the years evolved into a well-established, valuable tool to study proteins in close-to-native conditions (Serber and Dötsch, 2001; Montheith and Pielak., 2014; Theillet et al., 2016; Luchinat and Banci., 2018). Until now, only a single study has been published on in-cell NMR experiments of tau protein, focusing on a shorter fragment (Zhang et al., 2018). In this work, the authors acquired 2D HSQC spectra of the ^{15}N-labelled K19 fragment in several buffer conditions and subsequently, upon electroporating the protein into HEK-293 cells. The spectrum of K19 under *in situ* conditions was most similar to the one where K19 was incubated *in vitro* with polymerized MTs, indicating that tau may primarily interact with MTs *in situ*. The authors also highlighted that the most significant change in signal intensity was observed in the PHF6 region (V306-K311). Additionally, a MARK2-phosphorylated version of K19 in the HEK-293 cells was observed to undergo rapid dephosphorylation of residues S262, S324, S352, S356 shortly after introduction into the cells. The spectrum of full-length tau showed widespread chemical shift perturbations, with most significant changes around the PHF6 region, similarly to K19, and with signals of V309, Y310, and K311 broadened beyond detection. Finally, immunofluorescence-based co-localization with tubulin in SH-SY5Y cells pointed to tau interacting with MTs, and NMR suggested the involvement of the PHF region in this interaction *in vivo*. These experiments highlight the potential of in-cell NMR to examine structural changes of tau in near-native conditions at atomistic resolution (Zhang et al., 2018).

In most cases, tau aggregation *in vitro* is initiated by polyanions, such as heparin, polyglutamic acid, and RNA (Mukrasch et al., 2005). Several polyanions were also found in brain-derived tissues of AD patients (Goedert et al., 1996; Paudel and Li., 1999). Upon polyanion binding tau protein's highly positive net charge is partially shielded, which facilitates formation of β-sheet structures. The polyanions directly stabilize the regions essential for aggregation (Mukrasch et al., 2005; Akoury et al., 2016). NMR experiments suggested that tau binds polyanions via the same interaction sites as MTs. Thus, the most noticeable changes of backbone amide chemical shifts were

similarly observed in proximity of lysine and histidine residues (Mukrasch et al., 2005; Sibille et al., 2006; Fischer et al., 2007; Mukrasch et al., 2007). Moreover, NMR revealed that the polyanion binding increases residual β-sheet propensity within R2 and R3 hexapeptides, which were identified as the seeds of tau filament formation (Sibille et al., 2006; Akoury et al., 2016). Taken together, phosphorylation and polyanions diminish tau interaction with microtubules by blocking the interaction sites and adapt them for the formation of filaments (Mukrasch et al., 2005).

Metal ions are essential for normal brain function. However, during AD pathogenesis these ions accumulate in the brain of patients (Ahmadi et al., 2019). *In vitro* studies revealed that metals like Zn^{2+} and Cu^{2+} bind to tau and increase its aggregation rate (Jiji et al., 2017). The mode of metal-binding to tau was investigated by NMR titration. The binding site of Cu^{2+} was located in repeats R2 and R3, in particular, V287-S293 and Y310-S324. Moreover, H299 in R2, H329 and H330 in R3 contributed to the interaction (Soragni et al., 2008). Another study proposed an alternative interaction mode of Cu^{2+} coordinated to H268 in R1 and H363 in R4, which promoted the dimerization of R2 and R3 via C291-C322 disulfide bond formation (Ahmadi et al., 2019). In the case of Zn^{2+}, its binding was mapped to the R3 repeat, particularly C322, and histidines H268, H329/H330, were found to complement the cysteine in Zn^{2+} binding (Jiji et al., 2017). These studies show a slightly different interaction mechanism between Zn^{2+} and Cu^{2+}, although in both cases the histidines and cysteines present in the repeat sequences participate in the coordination of these ions. Such interactions may result in stabilization of a particular conformation, particularly β-turn structures.

Another important connection can be found between tau, phosphorylation and the 14-3-3 protein family. The 14-3-3 protein isoforms are highly expressed in the human brain and interact with thousands of protein partners (Sluchanko and Bustos, 2019; Gogl et al., 2021). It has been proposed that 14-3-3 proteins might be in a competitive relationship with tubulin for binding to tau (Hashiguchi et al., 2000; Qureshi et al., 2013; Chen et al., 2019). 14-3-3 proteins have been found colocalised with tau in NFTs extracted from AD brains (Layfield et al., 1996; Umahara et al., 2004), and they were suggested to promote tau protein aggregation and fibrillization in a phosphorylation-dependent manner (Hernández et al., 2004; Sadik et al., 2009). As described above, tau's interaction with microtubules is heavily influenced by its phosphorylation status (Lindwall and Cole, 1984; Biernat et al., 1993). For example, phosphorylation by PKA is known to decrease tau-tubulin binding, by modifying S214, T231 or S356 (Scott et al., 1993; Scheider et al., 1999). At the same time, it generates binding epitopes for the 14-3-3 proteins (Sluchanko et al., 2009; Joo et al., 2015). This may seriously affect microtubule stability and cell viability. Indeed, *in vivo* experiments have shown significantly retarded axonal development in neuronal cultures overexpressing 14-3-3, via microtubule destabilization (Joo et al., 2015; Li and Paudel, 2016). The precise role of 14-3-3 in the process of tau aggregation remains to be elucidated, however, NMR studies have highlighted their interaction sites.

Using chemical shift perturbation mapping, the binding region(s) of 14-3-3 on phosphorylated full-length tau have been described in detail (Joo et al., 2015; Andrei et al., 2018). High quality ^{1}H-^{15}N HSQC spectra demonstrated a significant signal reduction of residues throughout the MTBR and proline-rich regions of tau (mainly in the vicinities of S214 and S324), while resonances in the projection and C-terminal domains were nearly unaffected. Notably, this description is strikingly similar to the regions responsible for tubulin binding (Sillen et al., 2007).

Tau has been shown to form soluble complexes with amyloid beta that may promote their aggregation into the insoluble forms observed in AD (Guo et al., 2006; Jin et al., 2011). Similarly, the presence of monomeric α-synuclein was found to promote formation of tau co-aggregated fibrils (Lu et al., 2020; Hojjatian et al., 2021). Therefore, cross-interactions of tau with other proteins involved in neurodegenerative diseases (α-synuclein and Aβ40) have been studied by NMR. NMR chemical shift perturbations revealed that α-synuclein interacts mainly with the PHF6 motif of tau through its negatively charged C-terminal region (Lu et al., 2020). α-Synuclein fibrils formed in the presence of tau were recently characterized by ssNMR experiments revealing that they share similar conformation with one particular type of fibrils obtained in the absence of tau (Hojjatian et al., 2021). In contrast, addition of full-length tau did not induce chemical shift perturbations in the ^{1}H-^{15}N HSQC spectrum of ^{15}N labeled Aβ40, although small losses in signal intensity were observed immediately after addition of tau pointing to weak interaction. As a result of this, the co-incubation of Aβ and tau induced amorphous aggregates and inhibited Aβ40 from fibrillization (Wallin et al., 2018). These examples show that tau cross-interactions also need to be considered in its aggregation mechanism.

OLIGOMERS AND OTHER INTERMEDIATES AS A BLACK BOX OF TAU AGGREGATION

In spite of all gained knowledge of the potential triggers of tau aggregation, the mechanism of how soluble tau undergoes assembly into insoluble filaments is not well explored. Intermediates in this process are oligomers. Over the last decades, several studies have shown that toxic soluble oligomers could initiate the neurodegeneration cascade. Moreover, there is rising evidence that the onset of Alzheimer's disease and other tauopathies occurs earlier than tau filaments are found in the brain. Therefore, it is crucial to characterize intermediates along the aggregation pathway (Berger et al., 2007; de Calignon et al., 2010; Lasagna-Revees et al., 2012; Ghag et al., 2018), which may be the most relevant form to target with anti-tau therapies.

The contribution of NMR to the characterization of oligomers remains scarce. Soluble oligomers of tau187 (residues N255-L441), which comprises all four repeats and the C-terminal domain of tau, were detected by solution NMR. Backbone resonances of monomeric tau187 were assigned using conventional 3D triple-resonance experiments. Further, paramagnetic relaxation enhancements (PREs) were measured in response to heparin-induced aggregation of ^{15}N labeled tau, which highlighted two MTSL-broadened regions, V275-K280 and V306-K311. In line

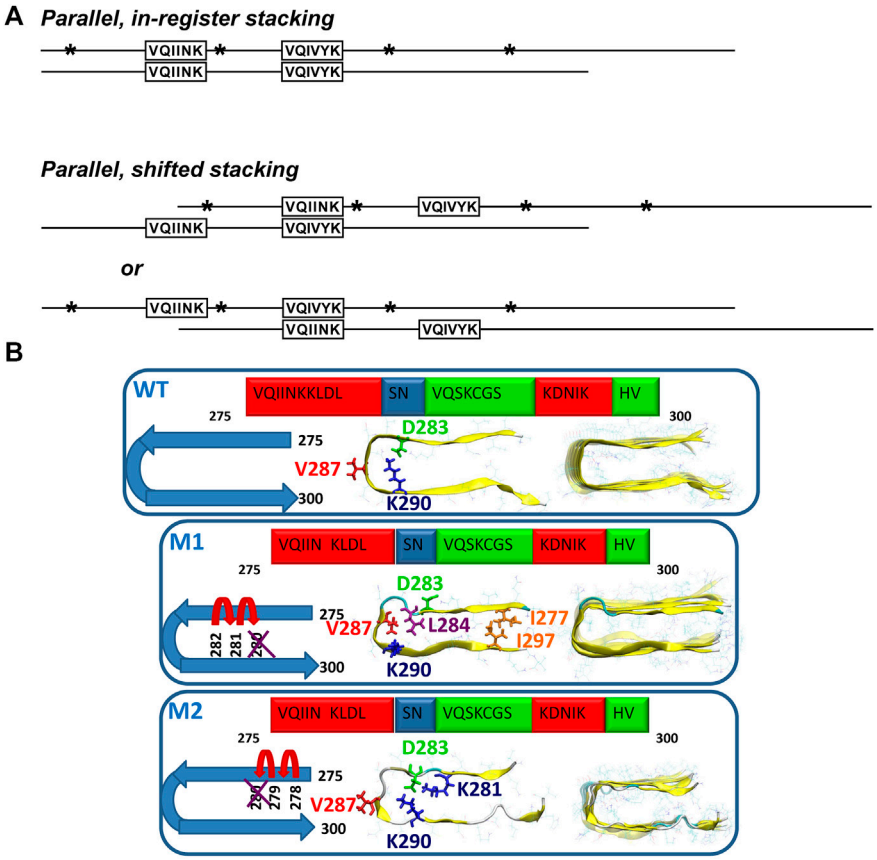


FIGURE 4 | (A) Models of possible tau oligomer complexes generated during heparin-induced aggregation (parallel, in-register stacking, and parallel, shifted stacking). Asterisks show the positions of single MTSL derivatizations in relation to the regions observed to be broadened: ^{275}V-K^{280} and ^{306}V-K^{311}. Adapted with permission from Peterson et al. (2008). Copyright 2008 American Chemical Society. **(B)** Schematic representation of two oligomer structural models of the mutated ΔK280 tau repeat R2: models M1 and M2. In M1, the deletion mutation was obtained by 'shifting' the C-terminal sequence towards K280. In M2, the deletion mutation was obtained by 'shifting' the N-terminal sequence towards K280. Reproduced from Raz et al. (2014) with permission from the PCCP Owner Societies.

with this observation, it was proposed that soluble oligomers are generated via parallel, in-register and parallel, shifted-register intermolecular interactions at these regions (**Figure 4A**). However, it remained unclear whether these oligomer species belong to the on-pathway oligomers (Peterson et al., 2008). PRE NMR was also used to capture the soluble oligomers of tau in response to interaction with the organic compound pthalocyanine tetrasulfonate (PcTS), which inhibits tau aggregation. Moreover, the study results indicated that the formed off-pathway oligomers are structurally distinct from toxic oligomers of tau (Akoury et al., 2013).

The ΔK280 deletion mutant, which is known to accelerate tau aggregation and is associated with the development of frontotemporal dementia (Wegmann et al., 2011) was studied by a combination of solid-state NMR, atomic force microscopy, transmission electron microscopy (TEM) and all-atom explicit molecular dynamics simulations in the R2 peptide (Raz et al., 2014). The authors found that this deletion mutation induces the formation of oligomers and reduces the generation of fibrils. Two structural models of the oligomers were proposed by 'shifting' the sequence from the C- or the N-terminal end towards the ΔK280 mutation site (**Figure 4B**). Model M1 is characterized by mostly hydrophobic contacts, whereas in the model M2, the dominant interactions are salt bridges. The ssNMR chemical shift assignment of R2 in ΔK280 revealed that labeled residues D283, V287, and K290 are in β-sheet conformation. In the simulated model M1, V286 and K290 show the β-sheet conformation, while, in the M2 model, D283 and V287 show the β-sheet conformation. The molecular dynamics data illustrated that M1 adopts a relatively well-packed structure compared with the M2 model. Comparison of the relative conformational energies and the populations of models shows that model M1 is more stable and strongly preferred over model M2. Therefore, the authors proposed that larger populations of the self-assembled ΔK280 tau R2 repeat oligomers and fibrils are organized as in the model M1 (Raz et al., 2014).

FILAMENTS AS AGGREGATION END-PRODUCTS

Until the mid-1990s, studies of tau filaments were limited to patient-derived material due to the unavailability of well-

established methods to spontaneously aggregate tau protein *in vitro* (Despres et al., 2019). Since the breakthrough discovery by Goedert et al. that heparin, a polyanionic cofactor, can trigger the formation of filaments of non-phosphorylated tau protein, aggregation of tau *in vitro* has been done by its addition. Heparin screens electrostatic interactions, which result in conformational rearrangement of tau protein, leading to its self-assembly (Goedert et al., 1996; Fichou et al., 2018). Besides heparin, other polyanionic compounds such as heparan sulfate (Zhao et al., 2020), RNA, arachidonic acid (Sibille et al., 2006), polyglutamic acid (Akoury et al., 2016) can be used for spontaneous aggregation of tau. It has also been reported that metal (Cu^{2+}, Zn^{2+}) ions can trigger the aggregation process (Soragni et al., 2008; Jiji et al., 2017; Ahmadi et al., 2019). This variety of inducers demonstrates that tau aggregation is rather influenced by electrostatics than by the specific interactions with the inducer (Sibille et al., 2006).

Recent cryo-EM progress showed that heparin-induced tau filaments are structurally heterogeneous and distinct from those in Alzheimer's and Pick's disease, questioning the relevance of such aggregation protocols (Fichou et al., 2018; Zhang et al., 2019). For this reason, more recently, methods to obtain filaments without an aggregation inducer are used in the tau scientific community and latest NMR studies. Usually, truncated constructs of tau comprising only the repeat sequences are used in such an approach. For example, tauF4 (residues S208-S324) (Huvent et al., 2014), dGAE (residues I297-E391) (Al-Hilaly et al., 2019), R3R4 (residues V306-F378) (Carlomagno et al., 2021; Jayan et al., 2021) are able to form filaments without addition of an inducer reagent. Moreover, Carlomagno et al. showed that R3R4 can serve as a seed and promote the aggregation of full-length tau (Carlomagno et al., 2021).

Heparin-Induced Aggregates

In electron micrographs, a rigid core of tau filaments appears to be surrounded by a fuzzy outer coat (Wischik et al., 1988; Sillen et al., 2005b). Protein regions, which are incorporated in the core of the filaments of tau, are broadened beyond the detection limits of solution NMR due to immobilization, while residues in the outer coat outside the core retain a significant degree of mobility and should be observable by solution NMR (Sillen et al., 2005a; Sillen et al., 2005b). The first attempt to study heparin-induced tau filaments by ssNMR was made by Sillen and coworkers (Sillen et al., 2005b). The authors analysed the intensity of backbone amide peaks in the ^{1}H-^{15}N HSQC spectrum as a function of primary sequence location, which allowed identification of protein regions with distinct mobility after assembly into PHFs. At the flexible N-terminal part of the protein, full intensity was recovered for the amide peaks up to A77. The following residues up to the proline-rich region showed a linear reduction of peak intensity. Most of the proline-rich region (residues T205-R230) displayed a residual intensity ratio below 30% and was defined as semi-rigid with complex dynamics. The rigid PHF core was mapped to the region G261-T386 with the lowest signal intensities. The C-terminal residue peak intensities were partially recovered indicating that the rigid core does not extend to the C terminus of tau (Sillen et al., 2005a). Later PRE measurements were used to probe long-range interactions of the rigid core and the fuzzy coat of tau filaments. Nitroxide spin label attached to C322 caused signal broadening in the first 30 residues of the N-terminus as well as in residue stretches close to Q124, A152, N167–T212, and S409-A426 at the C-terminal part (Bibow et al., 2011).

Tau variants intensively studied by ssNMR were the truncated constructs K18, K19 and K32 (**Figure 1**) encompassing the core of native PHFs. In the case of heparin-fibrilized filaments, the K19 is the most studied construct of tau. Despite high sample heterogeneity causing considerable line broadening in ssNMR spectra, a complete resonance assignment was obtained for 43 residues in the rigid parts of the construct and 29 residues in the mobile N- and C-terminal part (Andronesi et al., 2008). Secondary chemical shifts were derived based on NCA and NCOCA data sets, which indicated strong β-sheet character for several residues in R1, R4, and the entire R3 as manifested by largely negative secondary chemical shift values, and random-coil or α-helical conformation near the N-and C- termini. The exact locations of β-strands were determined by analysing the chemical shifts together with correlations observed in an hNhhC experiment. These data indicated a short β-strand at the end of R1 (βR1, residues S262–K267), two β-strands in R3 (residues Q307–I328), and two β-strands in R4 (residues Q336–I354) (**Figure 5**). Moreover, the H_2O-edited NCA experiment suggested that βR1 and βR4 strands are more solvent-exposed in comparison to βR3 strand, which is more likely to be buried in the filament core. Lack of long-range and intermolecular contacts hampered the generation of a structural model, however, the authors proposed the relative arrangement of molecules in filaments. The burial of the βR3 strand suggested that it may form the interface within the minimal structural unit of K19 filaments comprising two molecules connected via a disulfide bridge (Andronesi et al., 2008).

In the following study Daebel and coworkers compared the dynamics of the repetitive regions (Daebel et al., 2012). Unique residues in each repeat, A246, Y310 and F346 were used as sentinel residues for R1, R3 and R4, respectively. Chemical shifts assigned using an INEPT-CC-TOBSY spectrum probed a random coil character for A246 and F346, while Y310 resonance was absent. This indicated that R1 and R4 are much more flexible compared to R3. Additionally, the role of the R2 repeat in the rigid core was investigated using filaments obtained from the K18 construct, which showed an overall behavior similar to K19. The R2 repeat was found to be protected from the exchange with solvent, similarly to R3. In contrast to Andronesi et al., about 80% of K19 resonances were not observed, possibly, due to increased heterogeneity. Nevertheless, PDSD experiments at various temperatures revealed that most residues become rigid at sub-zero temperatures. Using a selectively labeled sample, the oxidation state of the single cysteine residue (C322) in the K19 sequence was investigated. Two cysteine resonances were detected, which both exhibited oxidized cysteine chemical shifts, suggesting the presence of at least two conformations (**Figure 5**). Upon cysteine mutation to alanine, a simplified resonance pattern was observed, thereby confirming the role of the cysteine in producing structural heterogeneity (Daebel et al., 2012).

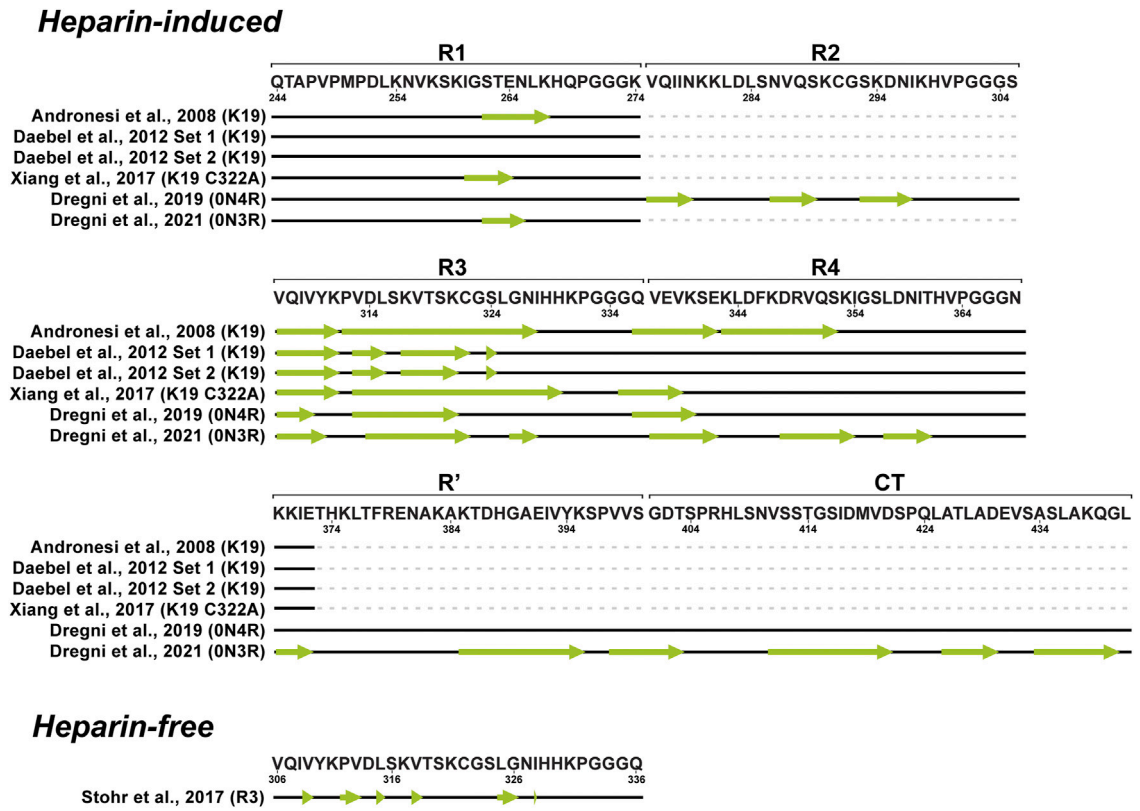


FIGURE 5 | Schematic representation of β-strand locations (indicated with green arrows) in fibrillar tau derived from solid-state NMR data.

The K19 C322A mutant was studied by Xiang and colleagues. ssNMR spectra and electron micrographs showed noticeable heterogeneity with several types of filaments within the sample although the same aggregation protocol as before was used implying batch-to-batch variability (Xiang et al., 2017; Daebel et al., 2012; Andronesi et al., 2008). Despite this heterogeneity, the authors employed proton-detected experiments with up to four chemical shift dimensions resulting in extensive chemical shift assignments for residues G260-E264 and V306-K340 (**Figure 5**). The authors concluded that tau filaments inherently exist as an ensemble of structures with a consistent and well defined structure only in the hexapeptide motif (Xiang et al., 2017).

On the other hand, Savastano and coworkers studied the involvement of the P2 region in filaments using the K32 construct that spans residues S198-S400 (Savastano et al., 2020). Regrettably, the resulting PDSD spectrum displayed signal overlap, giving just a few isolated peaks. However, the comparison of K32 and K19 PDSD spectra suggested that the R1 and the R3 repeats are part of the rigid cross-β structure in the K32 filaments. Evidence of P2 involvement in the rigid core structures was provided by ssNMR analysis of the model peptides P2R2 and P2R3. Upon aggregation, the resonances of P2 domain peptide K225-T231, which resembles the hexapeptides in repeats R2 and R3, lost their intensities compared to the monomeric state of the peptides.

An ssNMR structural model of 0N4R tau filaments has been published by Dregni et al. (**Figure 6**) (Dregni et al., 2019). For the first time, homogeneous filaments yielding high-quality ssNMR data were obtained. Assignment was performed for residues G270-K340 and the chemical shift–derived torsion angles indicated six β-strands (β1–β6) starting from the R2 hexapeptide motif ^{275}VQIINK^{280} and ending with the ^{336}QVEVK^{340} segment at the beginning of R4 (**Figure 5**). Remarkably, the assigned ^{13}C and ^{15}N chemical shifts of the 0N4R construct differed significantly from the truncated K18 and K19 tau constructs as well as showed different locations of β-strands (Andronesi et al., 2008; Daebel et al., 2012; Xiang et al., 2017). Therefore, the fibril core of 0N4R tau is distinct from tau filaments studied previously. 2D CC and 3D NCACX correlation spectra with a long mixing by ^{13}C–^{13}C CORD (Hou et al., 2013) spin diffusion were used to determine the overall fold of the 0N4R fibril core. Cross-peaks indicating close proximities were observed between the β3 strand (in R2) and the β4 strand (hexapeptide motif in R3), β1 strand and β5 strand as well as the β3 and β4 strands. In the model, the rigid β-sheet core spans residues V275-Q336 and is shaped like a hairpin (**Figure 6**), with β1 and β5 marking the approximate beginning and end. The intermolecular packing corresponds to parallel-in-register. The authors also noted the presence of a semi-rigid β-sheet domain flanking the filament core (Dregni et al., 2019).

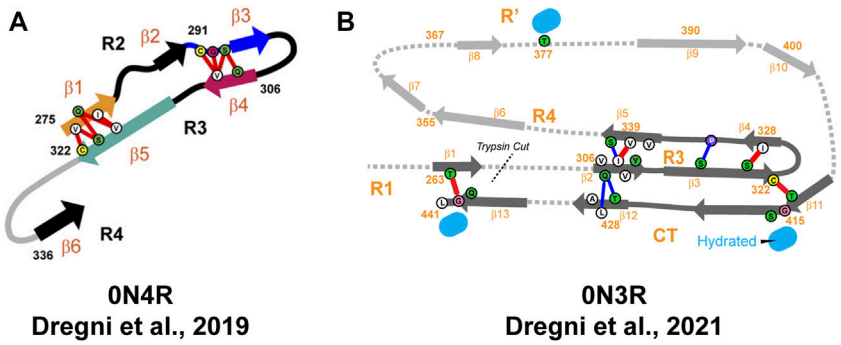


FIGURE 6 | (A) Structural model of the heparin-fibrillized 0N4R tau core. Schematic arrangement of the β-strands (thick arrows) and long-range correlations (red lines) measured in the ssNMR spectra. Hypothetical locations of segments outside the R2–R3 core are shown as gray lines. Orange, blue, magenta, and green arrows highlight the crucial R2 hexapeptide motif, the C291-containing segment, the R3 hexapeptide motif, and the C322-containing segment. Adapted with permission from Dregni et al. (2019). Copyright 2019 National Academy of Sciences. **(B)** Structural model of the heparin-fibrillized 0N3R tau. Unambiguous long-range contacts are shown as red lines while ambiguous contacts are shown as blue lines. The 3D fold of residues 342–411 is indicated by gray arrows and dotted lines. Adapted with permission from Dregni et al. (2021). Copyright 2021 American Chemical Society.

The previous structural model of 0N4R fibrils included only R2 and R3 parts of the protein sequence. In order to characterize the other parts in 0N4R, NMR relaxation and hydration studies were performed (Dregni et al., 2020). The authors concluded that the exterior of a β-sheet hairpin formed by R2-R3 is well-protected from water by other residues. Interestingly, the less rigid R1 and R4 domains did not exhibit significant differences in water accessibility compared to R2 and R3 domains, indicating their limited exposure to water molecules. Water-edited 2D^{13}C–^{13}C and ^{15}N-^{13}C correlation spectra showed that S285 and S316 are the most hydrated residues in the β-sheet core, while other serine, threonine and cysteine residues are poorly hydrated. S285 and S316 face each other in a central pocket, which led to the conclusion that the interior of the R2-R3 hairpin contains a small water pore. In addition, this water pore is local because the water-edited S285 signal is significantly more intense than that of the neighboring S289. The poor hydration of the 0N4R tau fibril core and small size of the water pore suggest that semi-rigid R1 and R4 domains or flanking regions outside repeat sequences are better targets of small-molecule drugs and imaging agents than the R2-R3 region. Although the core of the 0N4R fibril model does not include the R4 and R′ domains, the authors spotted similarities with the corticobasal degeneration CBD tau structure comprising the R2-R′ domains (Zhang et al., 2020). Both models exhibit the R2-R3 hairpin and a significant kink between the R2 hexapeptide and S285. Additionally, the hydration and dynamics data suggested that R4 is rigid and participates in hydrogen bonding in some units, while R3 is protected from water. The authors concluded that R3 and R4 are packed together like in the CBD tau fold and increased dynamics of R4 and R' is in agreement with lower resolution in the cryo-EM structure (Dregni et al., 2020).

Recently, the same group published another ssNMR derived structural model of homogeneous 0N3R filaments (**Figure 6**) (Dregni et al., 2021). A comparison of 2D ^{1}H–^{15}N correlation INEPT spectra between 0N3R and 0N4R highlighted missing peaks in the 0N3R spectrum corresponding to residues at the C-terminus, which indicated that the C-terminal domain is not isotropically mobile in the 0N3R isoform in contrast to 0N4R. Furthermore, a larger number of alanine peaks in the NCA spectrum of 0N3R were observed, which is consistent with the inclusion of the C-terminus into the β-sheet core since part of R′ and the C-terminus include seven alanines. The site-specific backbone assignments of the 0N3R tau fibril core were obtained for 104 out of 149 residues covering the region S262-L441 confirming the rigidity of the C-terminal residues. The peak intensities of R3, R4, R′, and C-terminal residues were comparable in dipolar correlation spectra, whereas the R1 signals were less intense, consistent with the start of the rigid core from the middle of R1. The proposed model showed that the 0N3R isoform has six β-strands in similar locations to 0N4R (G261-S262 in R1, V306–Y310 and V313–H330 in R3, V337–S341, V350–K353, and N359-T361 in R4) and four additional β-strands K385–S400, V411–S422, A426-A429, and A434-G440 in R′ and C-terminal part (**Figure 5**). The model was created based on 90 medium-range contacts obtained from spectra with long spin diffusion times, which included four unambiguous and four ambiguous long-range contacts that could not be explained by any short-range contacts. These eight long-range contacts provided explicit constraints on the tertiary fold of the 0N3R tau core indicating that the R1-R3 stretch is packed against the C-terminal region in an antiparallel fashion and that the R3 packs against the R4 repeat. In this filament model, the core has an elongated C-shape resembling an alligator head (**Figure 6**), which differs qualitatively from all *in vivo* and *in vitro* tau fibril core structures known to date (Dregni et al., 2021).

Inducer-free Spontaneously Formed Aggregates

The observed polymorphism and doubts about the biological relevance of heparin-induced filaments have recently inspired a surge in the studies of spontaneously aggregated tau. Stohr and

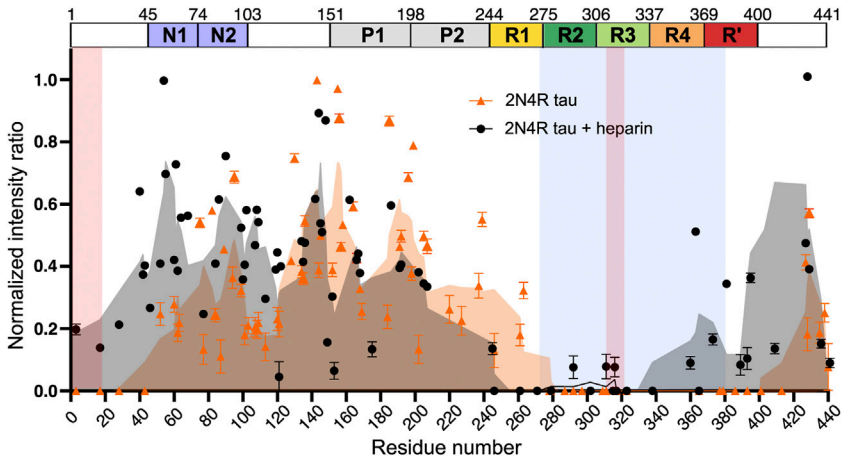


FIGURE 7 | Intensity ratio plot of INEPT signal of 2N4R tau fibrils aggregated in the absence (orange) and presence (black) of heparin. The intensity ratio was calculated by dividing the signal intensity of each residue in the fibril state by the monomeric state. The rigid cross-ß-sheet core of the tau fibril extracted from a CBD patient brain (PDB code: 6TJO) is marked in light blue. Adapted with permission from Chakraborty et al. (2021). Copyright 2021 Nature Publishing Group.

colleagues studied a set of different peptides derived from the R3 domain of tau (Stöhr et al., 2017). For ssNMR analysis, filaments of full-length R3 peptide formed under reducing conditions ($R3_{SH}$) were used, chosen for its high biological activity and macroscopic homogeneity confirmed by TEM and fibre diffraction. Two peptides with specific labeling schemes were synthesized for the magic angle spinning (MAS) ssNMR studies, and 2D^{13}C–^{13}C DARR correlation spectra acquired at 12 kHz MAS were used for assignment. Single peaks with narrow ^{13}C line widths were observed for all labelled residues except L315. Secondary chemical shift analysis indicated that all labelled residues, except L315 and G326, exhibit β-sheet conformation (**Figure 5**). For L315, the major conformer was predicted to be in the β-sheet conformation, while a minor conformer was in a non-beta conformation. The authors hypothesized that the proximity of P312 could break the β-sheet conformation. The ssNMR data were in excellent agreement with HDX experiments, which showed increased mobility in proximity to P312. The authors concluded that the R3 peptides are a small but biologically relevant system for addressing detailed biophysical questions regarding tau filament formation. Moreover, the full R3 repeat could serve as a template for full-length tau aggregation into filaments (Stöhr et al., 2017).

Previously, it has been shown that short peptides from tau repeat regions can self-assemble and form fibrils, and that spontaneous aggregation is less efficient as the length of tau fragment increases (Schweers et al., 1995; Stohr et al., 2017). Jayan and coworkers investigated the aggregation of R3R4 tau. In order to confirm the presence of R4 repeat in the fibril core, fibrils formed *in vitro* without heparin were studied using ssNMR (Jayan et al., 2021). Lack of NMR peaks in the INEPT-CC-TOBSY experiments suggested the absence of flexible residues in the R3R4 fibrils. The ^{13}C–^{13}C PDSD experiments showed slightly increased line widths of the peaks suggesting that there was still structural polymorphism in the fibrils. Secondary chemical shift analysis showed that all assigned residues in the R3 have β-strand conformation. The PDSD spectra also allowed to assign 13 amino acids of the R4 repeat, which was concluded to be a part of the rigid core, in line with the various fibril structure models from cryo-EM (Fitzpatrick et al., 2017; Falcon et al., 2018). Thus, this study suggests that the R3R4 is a suitable model system for *in vivo* tau filaments (Jayan et al., 2021).

Recently, 2N4R isoform filaments prepared in the absence of heparin were studied by ssNMR (Chakraborty et al., 2021). The analysis of ^{1}H-^{15}N INEPT spectra, which detects only highly dynamic residues in solid samples, revealed a complete loss of signals from residues P270-S400 (**Figure 7**). This is in contrast to the case of heparin-fibrilized 2N4R, for which signals from residues I260-H330 were not detected in the INEPT spectrum (**Figure 7**). These data suggest that filaments generated in the absence of heparin have a similar fibril core length as CBD tau fibrils (Zhang et al., 2020). Interestingly, 30 residues at the N-terminus could also not be observed suggesting a transient interaction between fibril core and the N-terminus. The peak widths in the 2D RFDR and 2D NCA spectra of the heparin-free 2N4R tau fibrils indicated structural homogeneity of the rigid core. However, only a few residues in the fibril core could be assigned due to strong signal overlap. To gain further insight into the structural properties of heparin-free 2N4R tau fibrils, dynamic nuclear polarisation (DNP)-enhanced 2D hChhC and hNhhC spectra of selectively labeled (^{13}Cγ valine, ^{13}C-ring phenylalanine, ^{15}N histidine) fibrils were measured. Two cross-peaks observed between the Cγ of valine and the ring carbons of phenylalanine suggested that the side chains of one or two valine residues are in proximity to the side chain of a phenylalanine residue. Another cross-peak between the Nε1/Nδ2 atoms of histidine side chain and Cγ of valine indicated that these groups are in ~4 Å distance in the structure of 2N4R fibrils. These contacts are in agreement with the cryo-EM structure of CBD fibrils, which shows that F346 is in proximity of V350, F378 contacts V275 and aromatic ring of H330 is close to Cγ of V318 (Zhang et al., 2020).

CONCLUSION

Solution and solid-state NMR have allowed bridging the gaps in structural knowledge of tau in the monomeric, oligomeric, and filamentous states. In comparison to other structural methods, solution NMR has provided residue specific insights into tau monomer regions with transient secondary structures, phosphorylation patterns and interactions with binding partners. Thus, the hexapeptide sequence motifs present in R2 and R3 repeats have been identified by several studies as regions with highest (up to 25%) β-sheet propensity (or extended conformation). Phosphorylation using site-specific kinases has been found to increase the β-sheet propensity in proline-rich regions, which could result in reduced tau binding to microtubules. The effect of polyanions is similar to phosphorylation as they were also shown to increase residual β-sheet propensity within R2 and R3 hexapeptides, thereby facilitating filament assembly. However, solution NMR studies on tau were primarily devoted to examining the role of repeats while regions outside the filament-forming core have not been characterized to the same extent. Phosphorylation studies have remained scarce and have mostly addressed isolated sites, therefore effects of hyperphosphorylation should be studied in the future, potentially using other kinases and employing also ^{31}P NMR. Additionally, it would be important to apply NMR for characterization of tau interactions with synthetic compounds or peptides that stabilize the monomeric state, thereby contributing to development of new AD therapies.

The weakest link in the chain of tau research has been oligomer characterization despite their high biological relevance and potential toxicity. NMR studies of tau oligomers have been hampered mainly because of limited capabilities to prepare stabilized aggregation intermediates. Soluble oligomers obtained by heparin-induced aggregation were characterized using PRE-NMR in solution, which indicated that the hexapeptides V275-K280 and V306-K311 participate in parallel, in-register and parallel, shifted-register intermolecular interactions. However, it remains unclear whether the oligomers detected in this way are on the aggregation pathway. The ΔK280 deletion mutant of the R2 repeat, which favors formation of oligomers over fibrils, was studied by solid-state NMR. The authors proposed a structural model of the ΔK280 oligomers, which comprised two β-strands (V275-L284 and K290-V300) assembled in a hairpin structure. Further studies of tau aggregation intermediates are necessary to elucidate the aggregation pathway and understand the disease-specificity of filament folds. The recent discovery of shorter tau constructs, which can be aggregated spontaneously, may lead to new approaches for studying oligomers by solution and solid-state NMR. In particular, sample freezing in solid-state DNP-NMR experiments could be used to stabilize and detect low-populated aggregation intermediates with sufficient sensitivity. Additionally, solution NMR could be used to monitor monomer depletion and oligomer formation during tau aggregation in a site-specific manner using selectively labeled samples.

NMR has played a major role in the initial structural characterization of tau filaments, obtained by heparin-induced aggregation. The rigid PHF core was mapped to the repeat region (residues G261-T386), whereas the proline-rich region (residues T205-R230) was defined as semi-rigid with complex dynamics. A major bottleneck for filament structure determination has been structural heterogeneity, and some studies have even concluded that tau filaments inherently exist as an ensemble of structures with a consistent and well-defined structure only in the hexapeptide motifs. To reduce the sample heterogeneity due to transient long-range interactions, solid-state NMR experiments were performed with the truncated tau variants K18, K19, K32 and/or cysteine mutants, since it was determined to play a role in producing structural heterogeneity. This has allowed determination of locations of secondary structures, their involvement in the rigid core and relative protection from solvent for K19, K18 and K32 filaments. The first structural model of heparin-induced tau filaments was determined for the 0N4R construct owing to preparation of homogeneous filaments. The model comprised six β-strands covering the protein region from R2 hexapeptide to the beginning of R4, which only partially overlapped with the locations of β-strands in K18 and K19 filaments. The overall fold of the 0N4R fibril resembled a hairpin and the tau proteins had parallel-in-register intermolecular packing. Recently, another structural model of homogeneous tau 0N3R filaments was published. A major difference with respect to the 0N4R filaments was the inclusion of the C-terminus in the β-sheet core. In the model, which has an elongated C-shape, six β-strands show similar locations to 0N4R and four additional β-strands are located in the R′ and C-terminal parts. However, the heparin-induced tau filament models differ qualitatively from the structures of all patient-derived tau filaments determined by cryo-EM. Therefore, several recent studies have been performed with shorter constructs including full-length R3 peptide and R3R4 that aggregate spontaneously. Such constructs have been suggested as suitable model systems for *in vivo* tau filaments and could also serve as templates for aggregation of full-length tau. Another example is the solid-state NMR study of 2N4R isoform filaments prepared in the absence of heparin, which revealed a similar fibril core length and long-range contacts consistent with CBD patient-derived tau fibrils studied by cryo-EM. These recent results suggest that ssNMR should move towards studies of spontaneously formed aggregates as well as aggregates formed by seeding with shorter constructs or patient-derived material. Although cryo-EM has recently made significant progress in the structural characterization of patient-derived tau filaments, there are still several fundamental questions unanswered (e.g., what exactly determines the different tau filament structural signatures in individual human tauopathies). Recent technological advances such as proton detection at very fast MAS and DNP-NMR give additional confidence that some of those questions could be answered using NMR.

In conclusion, detailed understanding of the effects and significance of individual changes in the tau assembly pathway is required for selecting the best molecular species to target with new AD therapies. In this review, NMR studies of various monomeric, oligomeric and filamentous species in solution and in solids have been considered, which together allow the sketching of a plausible aggregation pathway. Nevertheless, further research is required, particularly to characterize on-pathway intermediate aggregates, which remain a black box in the mechanism of tau aggregation. Also, the interplay between tau truncation and phosphorylation in relation to aggregation behavior and the final filament structure as well as co-aggregation with (cross)-interaction partners are unaddressed problems. Many of the underlying questions are well suited for NMR, thus, we can expect significant contributions in the tau field from NMR studies in the future.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was funded from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 873127 and by Latvian Council of Science grant no. lzp-2019/1-0244. KK is supported by "Mikrotīkls doctoral scholarship in the field of exact and medical sciences", administered by the University of Latvia Foundation. RS acknowledges VEGA agency, grant numbers 2/0145/19 and 02/0163/19. NG and JH acknowledge the Ministry of Education, Youth, and Sport of the Czech Republic (MEYS CR), grant number LTAUSA18168(Inter-Excellence Inter-Action).

REFERENCES

Ahmadi, S., Zhu, S., Sharma, R., Wu, B., Soong, R., Dutta Majumdar, R., et al. (2019). Aggregation of Microtubule Binding Repeats of Tau Protein Is Promoted by Cu^{2+}. *ACS Omega* 4, 5356–5366. doi:10.1021/acsomega.8b03595

Ahuja, P., Cantrelle, F.-X., Huvent, I., Hanoulle, X., Lopez, J., Smet, C., et al. (2016). Proline Conformation in a Functional Tau Fragment. *J. Mol. Biol.* 428, 79–91. doi:10.1016/j.jmb.2015.11.023

Ait-Bouziad, N., Chiki, A., Limorenko, G., Xiao, S., Eliezer, D., and Lashuel, H. A. (2020). Phosphorylation of the Overlooked Tyrosine 310 Regulates the Structure, Aggregation, and Microtubule- and Lipid-Binding Properties of Tau. *J. Biol. Chem.* 295, 7905–7922. doi:10.1074/jbc.RA119.012517

Akoury, E., Gajda, M., Pickhardt, M., Biernat, J., Soraya, P., Griesinger, C., et al. (2013). Inhibition of Tau Filament Formation by Conformational Modulation. *J. Am. Chem. Soc.* 135, 2853–2862. doi:10.1021/ja312471h

Akoury, E., Mukrasch, M. D., Biernat, J., Tepper, K., Ozenne, V., Mandelkow, E., et al. (2016). Remodeling of the Conformational Ensemble of the Repeat Domain of Tau by an Aggregation Enhancer. *Protein Sci.* 25, 1010–1020. doi:10.1002/pro.2911

Al-Hilaly, Y. K., Foster, B. E., Biasetti, L., Lutter, L., Pollack, S. J., Rickard, J. E., et al. (2019). Tau (297-391) Forms Filaments that Structurally Mimic the Core of Paired Helical Filaments in Alzheimer's Disease Brain. *FEBS Lett.* 594, 944–950. doi:10.1002/1873-3468.13675

Al-Hilaly, Y. K., Pollack, S. J., Vadukul, D. M., Citossi, F., Rickard, J. E., Simpson, M., et al. (2017). Alzheimer's Disease-like Paired Helical Filament Assembly from Truncated Tau Protein Is Independent of Disulfide Crosslinking. *J. Mol. Biol.* 429, 3650–3665. doi:10.1016/j.jmb.2017.09.007

Alonso, A. d. C., Zaidi, T., Novak, M., Grundke-Iqbal, I., and Iqbal, K. (2001). Hyperphosphorylation Induces Self-Assembly of into Tangles of Paired Helical Filaments/straight Filaments. *Proc. Natl. Acad. Sci.* 98, 6923–6928. doi:10.1073/pnas.121119298

Amniai, L., Barbier, P., Sillen, A., Wieruszeski, J.-M., Peyrot, V., Lippens, G., et al. (2009). Alzheimer Disease Specific Phosphoepitopes of Tau Interfere with Assembly of Tubulin but Not Binding to Microtubules. *FASEB j.* 23, 1146–1152. doi:10.1096/fj.08-121590

Amniai, L., Lippens, G., and Landrieu, I. (2011). Characterization of the AT180 Epitope of Phosphorylated Tau Protein by a Combined Nuclear Magnetic Resonance and Fluorescence Spectroscopy Approach. *Biochem. Biophysical Res. Commun.* 412, 743–746. doi:10.1016/j.bbrc.2011.08.046

Andrei, S. A., Meijer, F. A., Neves, J. F., Brunsveld, L., Landrieu, I., Ottmann, C., et al. (2018). Inhibition of 14-3-3/Tau by Hybrid Small-Molecule Peptides Operating via Two Different Binding Modes. *ACS Chem. Neurosci.* 9, 2639–2654. doi:10.1021/acschemneuro.8b00118

Andronesi, O. C., Bergen, M. v., Biernat, J., Seidel, K., Griesinger, C., Mandelkow, E., et al. (2008). Characterization of Alzheimer's-like Paired Helical Filaments from the Core Domain of Tau Protein Using Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 130, 5922–5928. doi:10.1021/ja7100517

Berger, Z., Roder, H., Hanna, A., Carlson, A., Rangachari, V., Yue, M., et al. (2007). Accumulation of Pathological Tau Species and Memory Loss in a Conditional Model of Tauopathy. *J. Neurosci.* 27, 3650–3662. doi:10.1523/jneurosci.0587-07.2007

Bibow, S., Mukrasch, M. D., Chinnathambi, S., Biernat, J., Griesinger, C., Mandelkow, E., et al. (2011). The Dynamic Structure of Filamentous Tau. *Angew. Chem. Int. Ed.* 50, 11520–11524. doi:10.1002/anie.201105493

Bielska, A. A., and Zondlo, N. J. (2006). Hyperphosphorylation of Tau Induces Local Polyproline II helix. *Biochemistry* 45, 5527–5537. doi:10.1021/bi052662c

Biernat, J., Gustke, N., Drewes, G., Mandelkow, E., and Mandelkow, E. (1993). Phosphorylation of Ser262 Strongly Reduces Binding of Tau to Microtubules: Distinction between PHF-like Immunoreactivity and Microtubule Binding. *Neuron* 11, 153–163. doi:10.1016/0896-6273(93)90279-z

Braak, H., and Braak, E. (1991). Neuropathological Stageing of Alzheimer-Related Changes. *Acta Neuropathol.* 82, 239–259. doi:10.1007/bf00308809

Braak, H., Thal, D. R., Ghebremedhin, E., and Del Tredici, K. (2011). Stages of the Pathologic Process in Alzheimer Disease: Age Categories from 1 to 100 Years. *J. Neuropathol. Exp. Neurol.* 70, 960–969. doi:10.1097/nen.0b013e318232a379

Carlomagno, Y., Manne, S., DeTure, M., Prudencio, M., Zhang, Y.-J., Hanna Al-Shaikh, R., et al. (2021). The AD Tau Core Spontaneously Self-Assembles and Recruits Full-Length Tau to Filaments. *Cel Rep.* 34, 108843. doi:10.1016/j.celrep.2021.108843

Cavazzoni, P. (2021). FDA's Decision to Approve New Treatment for Alzheimer's Disease. FDA. Available at: https://www.fda.gov/drugs/news-events-human-drugs/fdas-decision-approve-new-treatment-alzheimers-disease (Accessed July 12, 2021).

Chakraborty, P., Rivière, G., Liu, S., de Opakua, A. I., Dervişoğlu, R., Hebestreit, A., et al. (2021). Co-factor-free Aggregation of Tau into Seeding-Competent RNA-Sequestering Amyloid Fibrils. *Nat. Commun.* 12, 4231. doi:10.1038/s41467-021-24362-8

Chen, Y., Chen, X., Yao, Z., Shi, Y., Xiong, J., Zhou, J., et al. (2019). 14-3-3/Tau Interaction and Tau Amyloidogenesis. *J. Mol. Neurosci.* 68, 620–630. doi:10.1007/s12031-019-01325-9

Congdon, E. E., and Sigurdsson, E. M. (2018). Tau-targeting Therapies for Alzheimer Disease. *Nat. Rev. Neurol.* 14, 399–415. doi:10.1038/s41582-018-0013-z

Cummings, J. L., Morstorf, T., and Zhong, K. (2014). Alzheimer's Disease Drug-Development Pipeline: Few Candidates, Frequent Failures. *Alzheimers Res. Ther.* 6, 37. doi:10.1186/alzrt269

Daebel, V., Chinnathambi, S., Biernat, J., Schwalbe, M., Habenstein, B., Loquet, A., et al. (2012). β-Sheet Core of Tau Paired Helical Filaments Revealed by Solid-State NMR. *J. Am. Chem. Soc.* 134, 13982–13989. doi:10.1021/ja305470p

Daly, N. L., Hoffmann, R., Otvos, L., and Craik, D. J. (2000). Role of Phosphorylation in the Conformation of τ Peptides Implicated in Alzheimer's Disease. *Biochemistry* 39, 9039–9046. doi:10.1021/bi0004807

de Calignon, A., Fox, L. M., Pitstick, R., Carlson, G. A., Bacskai, B. J., Spires-Jones, T. L., et al. (2010). Caspase Activation Precedes and Leads to Tangles. *Nature* 464, 1201–1204. doi:10.1038/nature08890

Despres, C., Byrne, C., Qi, H., Cantrelle, F.-X., Huvent, I., Chambraud, B., et al. (2017). Identification of the Tau Phosphorylation Pattern that Drives its Aggregation. *Proc. Natl. Acad. Sci. USA*. 114, 9080–9085. doi:10.1073/pnas.1708448114

Despres, C., Di, J., Cantrelle, F.-X., Li, Z., Huvent, I., Chambraud, B., et al. (2019). Major Differences between the Self-Assembly and Seeding Behavior of Heparin-Induced and *In Vitro* Phosphorylated Tau and Their Modulation by Potential Inhibitors. *ACS Chem. Biol.* 14, 1363–1379. doi:10.1021/acschembio.9b00325

Ding, F., Borreguero, J. M., Buldyrey, S. V., Stanley, H. E., and Dokholyan, N. V. (2003). Mechanism for the α-helix to β-hairpin Transition. *Proteins* 53, 220–228. doi:10.1002/prot.10468

Dregni, A. J., Duan, P., and Hong, M. (2020). Hydration and Dynamics of Full-Length Tau Amyloid Fibrils Investigated by Solid-State Nuclear Magnetic Resonance. *Biochemistry* 59, 2237–2248. doi:10.1021/acs.biochem.0c00342

Dregni, A. J., Mandala, V. S., Wu, H., Elkins, M. R., Wang, H. K., Hung, I., et al. (2019). *In Vitro* 0N4R Tau Fibrils Contain a Monomorphic β-sheet Core Enclosed by Dynamically Heterogeneous Fuzzy Coat Segments. *Proc. Natl. Acad. Sci. USA*. 116, 16357–16366. doi:10.1073/pnas.1906839116

Dregni, A. J., Wang, H. K., Wu, H., Duan, P., Jin, J., DeGrado, W. F., et al. (2021). Inclusion of the C-Terminal Domain in the β-Sheet Core of Heparin-Fibrillized Three-Repeat Tau Protein Revealed by Solid-State Nuclear Magnetic Resonance Spectroscopy. *J. Am. Chem. Soc.* 143, 7839–7851. doi:10.1021/jacs.1c03314

Eliezer, D., Barré, P., Kobaslija, M., Chan, D., Li, X., and Heend, L. (2005). Residual Structure in the Repeat Domain of Tau: Echoes of Microtubule Binding and Paired Helical Filament Formation. *Biochemistry* 44, 1026–1036. doi:10.1021/bi048953n

Falcon, B., Zhang, W., Murzin, A. G., Murshudov, G., Garringer, H. J., Vidal, R., et al. (2018). Structures of Filaments from Pick's Disease Reveal a Novel Tau Protein Fold. *Nature* 561, 137–140. doi:10.1038/s41586-018-0454-y

Fichou, Y., Vigers, M., Goring, A. K., Eschmann, N. A., and Han, S. (2018). Heparin-induced Tau Filaments Are Structurally Heterogeneous and Differ from Alzheimer's Disease Filaments. *Chem. Commun.* 54, 4573–4576. doi:10.1039/c8cc01355a

Fischer, D., Mukrasch, M. D., Biernat, J., Bibow, S., Blackledge, M., Griesinger, C., et al. (2009). Conformational Changes Specific for Pseudophosphorylation at Serine 262 Selectively Impair Binding of Tau to Microtubules. *Biochemistry* 48, 10047–10055. doi:10.1021/bi901090m

Fischer, D., Mukrasch, M. D., von Bergen, M., Klos-Witkowska, A., Biernat, J., Griesinger, C., et al. (2007). Structural and Microtubule Binding Properties of Tau Mutants of Frontotemporal Dementias. *Biochemistry* 46, 2574–2582. doi:10.1021/bi061318s

Fitzpatrick, A. W. P., Falcon, B., He, S., Murzin, A. G., Murshudov, G., Garringer, H. J., et al. (2017). Cryo-EM Structures of Tau Filaments from Alzheimer's Disease. *Nature* 547, 185–190. doi:10.1038/nature23002

Gandhi, N. S., Landrieu, I., Byrne, C., Kukic, P., Amniai, L., Cantrelle, F.-X., et al. (2015). A Phosphorylation-Induced Turn Defines the Alzheimer's Disease AT8 Antibody Epitope on the Tau Protein. *Angew. Chem. Int. Ed.* 54, 6819–6823. doi:10.1002/anie.201501898

Ghag, G., Bhatt, N., Cantu, D. V., Guerrero-Munoz, M. J., Ellsworth, A., Sengupta, U., et al. (2018). Soluble Tau Aggregates, Not Large Fibrils, Are the Toxic Species that Display Seeding and Cross-Seeding Behavior. *Protein Sci.* 27, 1901–1909. doi:10.1002/pro.3499

Goedert, M., Jakes, R., Spillantini, M. G., Hasegawa, M., Smith, M. J., and Crowther, R. A. (1996). Assembly of Microtubule-Associated Protein Tau into Alzheimer-like Filaments Induced by Sulphated Glycosaminoglycans. *Nature* 383, 550–553. doi:10.1038/383550a0

Goedert, M., Wischik, C. M., Crowther, R. A., Walker, J. E., and Klug, A. (1988). Cloning and Sequencing of the cDNA Encoding a Core Protein of the Paired Helical Filament of Alzheimer Disease: Identification as the Microtubule-Associated Protein Tau. *Proc. Natl. Acad. Sci.* 85, 4051–4055. doi:10.1073/pnas.85.11.4051

Gogl, G., Tugaeva, K. V., Eberling, P., Kostmann, C., Trave, G., and Sluchanko, N. N. (2021). Hierarchized Phosphotarget Binding by the Seven Human 14-3-3 Isoforms. *Nat. Commun.* 12, 1677. doi:10.1038/s41467-021-21908-8

Guo, J.-P., Arai, T., Miklossy, J., and McGeer, P. L. (2006). Abeta and Tau Form Soluble Complexes that May Promote Self Aggregation of Both into the Insoluble Forms Observed in Alzheimer's Disease. *Proc. Natl. Acad. Sci.* 103, 1953–1958. doi:10.1073/pnas.0509386103

Habchi, J., Tompa, P., Longhi, S., and Uversky, V. N. (2014). Introducing Protein Intrinsic Disorder. *Chem. Rev.* 114, 6561–6588. doi:10.1021/cr400514h

Haj-Yahya, M., Gopinath, P., Rajasekhar, K., Mirbaha, H., Diamond, M. I., and Lashuel, H. A. (2020). Site-Specific Hyperphosphorylation Inhibits, rather Than Promotes, Tau Fibrillization, Seeding Capacity, and its Microtubule Binding. *Angew. Chem. Int. Ed.* 59, 4059–4067. doi:10.1002/anie.201913001

Harbison, N. W., Bhattacharya, S., and Eliezer, D. (2012). Assigning Backbone NMR Resonances for Full Length Tau Isoforms: Efficient Compromise between Manual Assignments and Reduced Dimensionality. *PLoS ONE* 7, e34679. doi:10.1371/journal.pone.0034679

Hashiguchi, M., Sobue, K., and Paudel, H. K. (2000). 14-3-3ζ Is an Effector of Tau Protein Phosphorylation. *J. Biol. Chem.* 275, 25247–25254. doi:10.1074/jbc.m003738200

Hernández, F., Cuadros, R., and Avila, J. (2004). Zeta 14-3-3 Protein Favours the Formation of Human Tau Fibrillar Polymers. *Neurosci. Lett.* 357, 143–146. doi:10.1016/j.neulet.2003.12.049

Hojjatian, A., Dasari, A. K. R., Sengupta, U., Taylor, D., Daneshparvar, N., Yeganeh, F. A., et al. (2021). Tau Induces Formation of α-synuclein Filaments with Distinct Molecular Conformations. *Biochem. Biophysical Res. Commun.* 554, 145–150. doi:10.1016/j.bbrc.2021.03.091

Hou, G., Yan, S., Trébosc, J., Amoureux, J.-P., and Polenova, T. (2013). Broadband Homonuclear Correlation Spectroscopy Driven by Combined R2nv Sequences under Fast Magic Angle Spinning for NMR Structural Analysis of Organic and Biological Solids. *J. Magn. Reson.* 232, 18–30. doi:10.1016/j.jmr.2013.04.009

Huvent, I., Kamah, A., Cantrelle, F.-X., Barois, N., Slomianny, C., Smet-Nocca, C., et al. (2014). A Functional Fragment of Tau Forms Fibers without the Need for an Intermolecular Cysteine Bridge. *Biochem. Biophysical Res. Commun.* 445, 299–303. doi:10.1016/j.bbrc.2014.01.161

Jadhav, S., Avila, J., Schöll, M., Kovacs, G. G., Kövari, E., Skrabana, R., et al. (2019). A Walk through Tau Therapeutic Strategies. *Acta Neuropathol. Commun.* 7, 22. doi:10.1186/s40478-019-0664-z

Jayan, P., Vahid, A. A., Kizhakkeduth, S. T., Muhammed, S. O. H., Shibina, A. B., and Vijayan, V. (2021). Direct Observation of the Self-Aggregation of R3R4 Bi-repeat of Tau Protein. *ChemBioChem* 22, 2093–2097. doi:10.1002/cbic.202100013

Jiji, A. C., Arshad, A., Dhanya, S. R., Shabana, P. S., Mehjubin, C. K., and Vijayan, V. (2017). Zn^{2+} Interrupts R4-R3 Association Leading to Accelerated Aggregation of Tau Protein. *Chem. Eur. J.* 23, 16976–16979. doi:10.1002/chem.201704555

Jin, M., Shepardson, N., Yang, T., Chen, G., Walsh, D., and Selkoe, D. J. (2011). Soluble Amyloid -protein Dimers Isolated from Alzheimer Cortex Directly Induce Tau Hyperphosphorylation and Neuritic Degeneration. *Proc. Natl. Acad. Sci.* 108, 5819–5824. doi:10.1073/pnas.1017033108

Joo, Y., Schumacher, B., Landrieu, I., Bartel, M., Smet-Nocca, C., Jang, A., et al. (2015). Involvement of 14-3-3 in Tubulin Instability and Impaired Axon Development Is Mediated by Tau. *FASEB j.* 29, 4133–4144. doi:10.1096/fj.14-265009

Kadavath, H., Cabrales Fontela, Y., Jaremko, M., Jaremko, Ł., Overkamp, K., Biernat, J., et al. (2018). The Binding Mode of a Tau Peptide with Tubulin. *Angew. Chem. Int. Ed.* 57, 3246–3250. doi:10.1002/anie.201712089

Kadavath, H., Hofele, R. V., Biernat, J., Kumar, S., Tepper, K., Urlaub, H., et al. (2015a). Tau Stabilizes Microtubules by Binding at the Interface between Tubulin Heterodimers. *Proc. Natl. Acad. Sci. USA*. 112, 7501–7506. doi:10.1073/pnas.1504081112

Kadavath, H., Jaremko, M., Jaremko, Ł., Biernat, J., Mandelkow, E., and Zweckstetter, M. (2015b). Folding of the Tau Protein on Microtubules. *Angew. Chem. Int. Ed.* 54, 10347–10351. doi:10.1002/anie.201501714

Kellogg, E. H., Hejab, N. M. A., Poepsel, S., Downing, K. H., DiMaio, F., and Nogales, E. (2018). Near-atomic Model of Microtubule-Tau Interactions. *Science* 360, 1242–1246. doi:10.1126/science.aat1780

Kovacs, G. G. (2015). Invited Review: Neuropathology of Tauopathies: Principles and Practice. *Neuropathol. Appl. Neurobiol.* 41, 3–23. doi:10.1111/nan.12208

Landrieu, I., Lacosse, L., Leroy, A., Wieruszeski, J.-M., Trivelli, X., Sillen, A., et al. (2006). NMR Analysis of a Tau Phosphorylation Pattern. *J. Am. Chem. Soc.* 128, 3575–3583. doi:10.1021/ja054656+

Lasagna-Reeves, C. A., Castillo-Carranza, D. L., Sengupta, U., Sarmiento, J., Troncoso, J., Jackson, G. R., et al. (2012). Identification of Oligomers at Early Stages of Tau Aggregation in Alzheimer's Disease. *FASEB j.* 26, 1946–1959. doi:10.1096/fj.11-199851

Layfield, R., Fergusson, J., Aitken, A., Lowe, J., Landon, M., and Mayer, R. J. (1996). Neurofibrillary Tangles of Alzheimer's Disease Brains Contain 14-3-3 Proteins. *Neurosci. Lett.* 209, 57–60. doi:10.1016/0304-3940(96)12598-2

Li, C., and Götz, J. (2017). Somatodendritic Accumulation of Tau in Alzheimer's Disease Is Promoted by Fyn-mediated Local Protein Translation. *EMBO J.* 36, 3120–3138. doi:10.15252/embj.201797724

Li, T., and Paudel, H. K. (2016). 14-3-3ζ Mediates Tau Aggregation in Human Neuroblastoma M17 Cells. *PLoS ONE* 11, e0160635. doi:10.1371/journal.pone.0160635

Lindwall, G., and Cole, R. D. (1984). Phosphorylation Affects the Ability of Tau Protein to Promote Microtubule Assembly. *J. Biol. Chem.* 259, 5301–5305. Available at: https://pubmed.ncbi.nlm.nih.gov/6425287/ (Accessed August 18, 2021). doi:10.1016/s0021-9258(17)42989-9

Lippens, G., Sillen, A., Smet, C., Wieruszeski, J.-M., Leroy, A., Buée, L., et al. (2006). Studying the Natively Unfolded Neuronal Tau Protein by Solution NMR Spectroscopy. *Ppl* 13, 235–246. doi:10.2174/092986606775338461

Lippens, G., Wieruszeski, J.-M., Leroy, A., Smet, C., Sillen, A., Buée, L., et al. (2003). Proline-directed Random-Coil Chemical Shift Values as a Tool for the NMR Assignment of the Tau Phosphorylation Sites. *ChemBioChem.* 5, 73–78. doi:10.1002/cbic.200300763

Louša, P., Nedozrálová, H., Župa, E., Nováček, J., and Hritz, J. (2017). Phosphorylation of the Regulatory Domain of Human Tyrosine Hydroxylase 1 Monitored Using Non-uniformly Sampled NMR. *Biophysical Chem.* 223, 25–29. doi:10.1016/j.bpc.2017.01.003

Lu, J., Zhang, S., Ma, X., Jia, C., Liu, Z., Huang, C., et al. (2020). Structural Basis of the Interplay between α-synuclein and Tau in Regulating Pathological Amyloid Aggregation. *J. Biol. Chem.* 295, 7470–7480. doi:10.1074/jbc.RA119.012284

Luchinat, E., and Banci, L. (2018). In-cell NMR in Human Cells: Direct Protein Expression Allows Structural Studies of Protein Folding and Maturation. *Acc. Chem. Res.* 51, 1550–1557. doi:10.1021/acs.accounts.8b00147

Marsh, J. A., Singh, V. K., Jia, Z., and Forman-Kay, J. D. (2006). Sensitivity of Secondary Structure Propensities to Sequence Differences between α- and γ-synuclein: Implications for Fibrillation. *Protein Sci.* 15, 2795–2804. doi:10.1110/ps.062465306

Mayzel, M., Rosenlöw, J., Isaksson, L., and Orekhov, V. Y. (2014). Time-resolved Multidimensional NMR with Non-uniform Sampling. *J. Biomol. NMR* 58, 129–139. doi:10.1007/s10858-013-9811-1

Melková, K., Narasimhan, S., Jansen, S., Hritz, J., Škrabana, R., Zweckstetter, M., et al. (2019). Structure and Functions of Microtubule Associated Proteins Tau and MAP2c: Similarities and Differences. *Biomolecules* 9, 105. doi:10.3390/biom9030105

Minoura, K., Tomoo, K., Ishida, T., Hasegawa, H., Sasaki, M., and Taniguchi, T. (2002). Amphipathic Helical Behavior of the Third Repeat Fragment in the Tau Microtubule-Binding Domain, Studied by 1H NMR Spectroscopy. *Biochem. Biophysical Res. Commun.* 294, 210–214. doi:10.1016/S0006-291X(02)00457-6

Minoura, K., Tomoo, K., Ishida, T., Hasegawa, H., Sasaki, M., and Taniguchi, T. (2003). Solvent-Dependent Conformation of the Third Repeat Fragment in the Microtubule-Binding Domain of Tau Protein, Analyzed by1H-NMR Spectroscopy and Molecular Modeling Calculation. *Bcsj* 76, 1617–1624. doi:10.1246/bcsj.76.1617

Minoura, K., Yao, T.-M., Tomoo, K., Sumida, M., Sasaki, M., Taniguchi, T., et al. (2004). Different Associational and Conformational Behaviors between the Second and Third Repeat Fragments in the Tau Microtubule-Binding Domain. *Eur. J. Biochem.* 271, 545–552. doi:10.1046/j.1432-1033.2003.03956.x

Monteith, W. B., and Pielak, G. J. (2014). Residue Level Quantification of Protein Stability in Living Cells. *Proc. Natl. Acad. Sci.* 111, 11335–11340. doi:10.1073/pnas.1406845111

Mukrasch, M. D., Bibow, S., Korukottu, J., Jeganathan, S., Biernat, J., Griesinger, C., et al. (2009). Structural Polymorphism of 441-residue Tau at Single Residue Resolution. *Plos Biol.* 7, e1000034. doi:10.1371/journal.pbio.1000034

Mukrasch, M. D., Biernat, J., von Bergen, M., Griesinger, C., Mandelkow, E., and Zweckstetter, M. (2005). Sites of Tau Important for Aggregation Populate β-Structure and Bind to Microtubules and Polyanions. *J. Biol. Chem.* 280, 24978–24986. doi:10.1074/jbc.M501565200

Mukrasch, M. D., von Bergen, M., Biernat, J., Fischer, D., Griesinger, C., Mandelkow, E., et al. (2007). The "Jaws" of the Tau-Microtubule Interaction. *J. Biol. Chem.* 282, 12230–12239. doi:10.1074/jbc.M607159200

Narayanan, R. L., Dürr, U. H. N., Bibow, S., Biernat, J., Mandelkow, E., and Zweckstetter, M. (2010). Automatic Assignment of the Intrinsically Disordered Protein Tau with 441-residues. *J. Am. Chem. Soc.* 132, 11906–11907. doi:10.1021/ja105657f

Naruto, K., Minoura, K., Okuda, R., Taniguchi, T., In, Y., Ishida, T., et al. (2010). Interplay between I308 and Y310 Residues in the Third Repeat of Microtubule-Binding Domain Is Essential for Tau Filament Formation. *FEBS Lett.* 584, 4233–4236. doi:10.1016/j.febslet.2010.09.012

Novak, P., Cehlar, O., Skrabana, R., and Novak, M. (2018a). Tau Conformation as a Target for Disease-Modifying Therapy: The Role of Truncation. *Jad* 64, S535–S546. doi:10.3233/JAD-179942

Novak, P., Kontsekova, E., Zilka, N., and Novak, M. (2018b). Ten Years of Tau-Targeted Immunotherapy: The Path Walked and the Roads Ahead. *Front. Neurosci.* 12, 798. doi:10.3389/fnins.2018.00798

Novak, P., Kovacech, B., Katina, S., Schmidt, R., Scheltens, P., Kontsekova, E., et al. (2021). ADAMANT: a Placebo-Controlled Randomized Phase 2 Study of AADvac1, an Active Immunotherapy against Pathological Tau in Alzheimer's Disease. *Nat. Aging* 1, 521–534. doi:10.1038/s43587-021-00070-2

Paudel, H. K., and Li, W. (1999). Heparin-induced Conformational Change in Microtubule-Associated Protein Tau as Detected by Chemical Cross-Linking and Phosphopeptide Mapping. *J. Biol. Chem.* 274, 8029–8038. doi:10.1074/jbc.274.12.8029

Peterson, D. W., Zhou, H., Dahlquist, F. W., and Lew, J. (2008). A Soluble Oligomer of Tau Associated with Fiber Formation Analyzed by NMR. *Biochemistry* 47, 7393–7404. doi:10.1021/bi702466a

Qi, H., Prabakaran, S., Cantrelle, F.-X., Chambraud, B., Gunawardena, J., Lippens, G., et al. (2016). Characterization of Neuronal Tau Protein as a Target of Extracellular Signal-Regulated Kinase. *J. Biol. Chem.* 291, 7742–7753. doi:10.1074/jbc.M115.700914

Qureshi, H. Y., Li, T., MacDonald, R., Cho, C. M., Leclerc, N., and Paudel, H. K. (2013). Interaction of 14-3-3ζ with Microtubule-Associated Protein Tau within Alzheimer's Disease Neurofibrillary Tangles. *Biochemistry* 52, 6445–6455. doi:10.1021/bi400442d

Raz, Y., Adler, J., Vogel, A., Scheidt, H. A., Häupl, T., Abel, B., et al. (2014). The Influence of the ΔK280 Mutation and N- or C-Terminal Extensions on the Structure, Dynamics, and Fibril Morphology of the Tau R2 Repeat. *Phys. Chem. Chem. Phys.* 16, 7710–7717. doi:10.1039/C3CP54890B

Reitz, C., Brayne, C., and Mayeux, R. (2011). Epidemiology of Alzheimer Disease. *Nat. Rev. Neurol.* 7, 137–152. doi:10.1038/nrneurol.2011.2

Sadik, G., Tanaka, T., Kato, K., Yanagi, K., Kudo, T., and Takeda, M. (2009). Differential Interaction and Aggregation of 3-repeat and 4-repeat Tau Isoforms with 14-3-3ζ Protein. *Biochem. Biophysical Res. Commun.* 383, 37–41. doi:10.1016/j.bbrc.2009.03.107

Savastano, A., Jaipuria, G., Andreas, L., Mandelkow, E., and Zweckstetter, M. (2020). Solid-state NMR Investigation of the Involvement of the P2 Region in Tau Amyloid Fibrils. *Sci. Rep.* 10, 21210. doi:10.1038/s41598-020-78161-0

Schneider, A., Biernat, J., von Bergen, M., Mandelkow, E., and Mandelkow, E.-M. (1999). Phosphorylation that Detaches Tau Protein from Microtubules (Ser262, Ser214) Also Protects it against Aggregation into Alzheimer Paired Helical Filaments. *Biochemistry* 38, 3549–3558. doi:10.1021/bi981874p

Schwalbe, M., Kadavath, H., Biernat, J., Ozenne, V., Blackledge, M., Mandelkow, E., et al. (2015). Structural Impact of Tau Phosphorylation at Threonine 231. *Structure* 23, 1448–1458. doi:10.1016/j.str.2015.06.002

Schweers, O., Mandelkow, E. M., Biernat, J., and Mandelkow, E. (1995). Oxidation of Cysteine-322 in the Repeat Domain of Microtubule-Associated Protein Tau Controls the *In Vitro* Assembly of Paired Helical Filaments. *Proc. Natl. Acad. Sci.* 92, 8463–8467. doi:10.1073/pnas.92.18.8463

Scott, C. W., Spreen, R. C., Herman, J. L., Chow, F. P., Davison, M. D., Young, J., et al. (1993). Phosphorylation of Recombinant Tau by cAMP-dependent Protein Kinase. Identification of Phosphorylation Sites and Effect on Microtubule Assembly. *J. Biol. Chem.* 268, 1166–1173. doi:10.1016/S0021-9258(18)54055-2

Serber, Z., and Dötsch, V. (2001). In-cell NMR Spectroscopy. *Biochemistry* 40, 14317–14323. doi:10.1021/bi011751w

Sevigny, J., Chiao, P., Bussière, T., Weinreb, P. H., Williams, L., Maier, M., et al. (2016). The Antibody Aducanumab Reduces Aβ Plaques in Alzheimer's Disease. *Nature* 537, 50–56. doi:10.1038/nature19323

Shiraki, K., Nishikawa, K., and Goto, Y. (1995). Trifluoroethanol-induced Stabilization of the α-Helical Structure of β-Lactoglobulin: Implication for Non-hierarchical Protein Folding. *J. Mol. Biol.* 245, 180–194. doi:10.1006/jmbi.1994.0015

Sibille, N., Huvent, I., Fauquant, C., Verdegem, D., Amniai, L., Leroy, A., et al. (2012). Structural Characterization by Nuclear Magnetic Resonance of the Impact of Phosphorylation in the Proline-Rich Region of the Disordered Tau Protein. *Proteins* 80, 454–462. doi:10.1002/prot.23210

Sibille, N., Sillen, A., Leroy, A., Wieruszeski, J.-M., Mulloy, B., Landrieu, I., et al. (2006). Structural Impact of Heparin Binding to Full-Length Tau as Studied by NMR Spectroscopy. *Biochemistry* 45, 12560–12572. doi:10.1021/bi060964o

Sillen, A., Barbier, P., Landrieu, I., Lefebvre, S., Wieruszeski, J.-M., Leroy, A., et al. (2007). NMR Investigation of the Interaction between the Neuronal Protein Tau and the Microtubules. *Biochemistry* 46, 3055–3064. doi:10.1021/bi061920i

Sillen, A., Leroy, A., Wieruszeski, J.-M., Loyens, A., Beauvillain, J.-C., Buée, L., et al. (2005a). Regions of Tau Implicated in the Paired Helical Fragment Core as Defined by NMR. *Chembiochem* 6, 1849–1856. doi:10.1002/cbic.200400452

Sillen, A., Wieruszeski, J.-M., Leroy, A., Younes, A. B., Landrieu, I., and Lippens, G. (2005b). High-Resolution Magic Angle Spinning NMR of the Neuronal Tau Protein Integrated in Alzheimer's-like Paired Helical Fragments. *J. Am. Chem. Soc.* 127, 10138–10139. doi:10.1021/ja0516211

Sluchanko, N. N., and Bustos, D. M. (2019). Intrinsic Disorder Associated with 14-3-3 Proteins and Their Partners. *Prog. Mol. Biol. Transl. Sci.* 166, 19–61. doi:10.1016/bs.pmbts.2019.03.007

Sluchanko, N. N., Seit-Nebi, A. S., and Gusev, N. B. (2009). Phosphorylation of More Than One Site Is Required for Tight Interaction of Human Tau Protein with 14-3-3ζ. *FEBS Lett.* 583, 2739–2742. doi:10.1016/j.febslet.2009.07.043

Smet, C., Leroy, A., Sillen, A., Wieruszeski, J.-M., Landrieu, I., and Lippens, G. (2004). Accepting its Random Coil Nature Allows a Partial NMR Assignment of the Neuronal Tau Protein. *ChemBioChem.* 5, 1639–1646. doi:10.1002/cbic.200400145

Sogawa, K., Minoura, K., In, Y., Ishida, T., Taniguchi, T., and Tomoo, K. (2014). CH-π Interaction in VQIVYK Sequence Elucidated by NMR Spectroscopy Is Essential for PHF Formation of Tau. *Biopolymers* 102, 288–295. doi:10.1002/bip.22489

Soragni, A., Zambelli, B., Mukrasch, M. D., Biernat, J., Jeganathan, S., Griesinger, C., et al. (2008). Structural Characterization of Binding of Cu(II) to Tau Protein. *Biochemistry* 47, 10841–10851. doi:10.1021/bi8008856

Sotiropoulos, I., Galas, M.-C., Silva, J. M., Skoulakis, E., Wegmann, S., Maina, M. B., et al. (2017). Atypical, Non-standard Functions of the Microtubule Associated Tau Protein. *Acta Neuropathol. Commun.* 5, 91. doi:10.1186/s40478-017-0489-6

Stöhr, J., Wu, H., Nick, M., Wu, Y., Bhate, M., Condello, C., et al. (2017). A 31-residue Peptide Induces Aggregation of Tau's Microtubule-Binding Region in Cells. *Nat. Chem* 9, 874–881. doi:10.1038/nchem.2754

Theillet, F.-X., Binolfi, A., Bekei, B., Martorana, A., Rose, H. M., Stuiver, M., et al. (2016). Structural Disorder of Monomeric α-synuclein Persists in Mammalian Cells. *Nature* 530, 45–50. doi:10.1038/nature16531

Tomoo, K., Yao, T.-M., Minoura, K., Hiraoka, S., Sumida, M., Taniguchi, T., et al. (2005). Possible Role of Each Repeat Structure of the Microtubule-Binding Domain of the Tau Protein in *In Vitro* Aggregation. *J. Biochem.* 138, 413–423. doi:10.1093/jb/mvi142

Umahara, T., Uchihara, T., Tsuchiya, K., Nakamura, A., Iwamoto, T., Ikeda, K., et al. (2004). 14-3-3 Proteins and Zeta Isoform Containing Neurofibrillary Tangles in Patients with Alzheimer's Disease. *Acta Neuropathol.* 108, 279–286. doi:10.1007/s00401-004-0885-4

von Bergen, M., Barghorn, S., Biernat, J., Mandelkow, E.-M., and Mandelkow, E. (2005). Tau Aggregation Is Driven by a Transition from Random Coil to Beta Sheet Structure. *Biochim. Biophys. Acta (Bba) - Mol. Basis Dis.* 1739, 158–166. doi:10.1016/j.bbadis.2004.09.010

von Bergen, M., Friedhoff, P., Biernat, J., Heberle, J., Mandelkow, E.-M., and Mandelkow, E. (2000). Assembly of Tau Protein into Alzheimer Paired Helical Filaments Depends on a Local Sequence Motif (306VQIVYK311) Forming Beta Structure. *Proc. Natl. Acad. Sci.* 97, 5129–5134. doi:10.1073/pnas.97.10.5129

Wallin, C., Hiruma, Y., Wärmländer, S. K. T. S., Huvent, I., Jarvet, J., Abrahams, J. P., et al. (2018). The Neuronal Tau Protein Blocks *In Vitro* Fibrillation of the Amyloid-β (Aβ) Peptide at the Oligomeric Stage. *J. Am. Chem. Soc.* 140, 8138–8146. doi:10.1021/jacs.7b13623

Wegmann, S., Biernat, J., and Mandelkow, E. (2021). A Current View on Tau Protein Phosphorylation in Alzheimer's Disease. *Curr. Opin. Neurobiol.* 69, 131–138. doi:10.1016/j.conb.2021.03.003

Wegmann, S., Schöler, J., Bippes, C. A., Mandelkow, E., and Muller, D. J. (2011). Competing Interactions Stabilize Pro- and Anti-aggregant Conformations of Human Tau. *J. Biol. Chem.* 286, 20512–20524. doi:10.1074/jbc.m111.237875

Wimo, A., Jönsson, L., Gustavsson, A., McDaid, D., Ersek, K., Georges, J., et al. (2010). The Economic Impact of Dementia in Europe in 2008-cost Estimates from the Eurocode Project. *Int. J. Geriat. Psychiatry* 26, 825–832. doi:10.1002/gps.2610

Wischik, C. M., Novak, M., Edwards, P. C., Klug, A., Tichelaar, W., and Crowther, R. A. (1988). Structural Characterization of the Core of the Paired Helical Filament of Alzheimer Disease. *Proc. Natl. Acad. Sci.* 85, 4884–4888. doi:10.1073/pnas.85.13.4884

World Health Organisation (2020). Dementia. Who.int. Available at: https://www.who.int/news-room/fact-sheets/detail/dementia (Accessed July 1, 2021).

Xiang, S., Kulminskaya, N., Habenstein, B., Biernat, J., Tepper, K., Paulat, M., et al. (2017). A Two-Component Adhesive: Tau Fibrils Arise from a Combination of a Well-Defined Motif and Conformationally Flexible Interactions. *J. Am. Chem. Soc.* 139, 2639–2646. doi:10.1021/jacs.6b09619

Zhang, S., Wang, C., Lu, J., Ma, X., Liu, Z., Li, D., et al. (2018). In-cell NMR Study of Tau and MARK2 Phosphorylated Tau. *Ijms* 20, 90. doi:10.3390/ijms20010090

Zhang, W., Falcon, B., Murzin, A. G., Fan, J., Crowther, R. A., Goedert, M., et al. (2019). Heparin-induced Tau Filaments Are Polymorphic and Differ from Those in Alzheimer's and Pick's Diseases. *eLife* 8, e43584. doi:10.7554/elife.43584

Zhang, W., Tarutani, A., Newell, K. L., Murzin, A. G., Matsubara, T., Falcon, B., et al. (2020). Novel Tau Filament Fold in Corticobasal Degeneration. *Nature* 580, 283–287. doi:10.1038/s41586-020-2043-0

Zhao, J., Zhu, Y., Song, X., Xiao, Y., Su, G., Liu, X., et al. (2020). 3- O -Sulfation of Heparan Sulfate Enhances Tau Interaction and Cellular Uptake. *Angew. Chem. Int. Ed.* 59, 1818–1827. doi:10.1002/anie.201913029

Zhou, L.-X., Zeng, Z.-Y., Du, J.-T., Zhao, Y.-F., and Li, Y.-M. (2006). The Self-Assembly Ability of the First Microtubule-Binding Repeat from Tau and its Modulation by Phosphorylation. *Biochem. Biophysical Res. Commun.* 348, 637–642. doi:10.1016/j.bbrc.2006.07.099

Zilka, N., Kazmerova, Z., Jadhav, S., Neradil, P., Madari, A., Obetkova, D., et al. (2012). Who Fans the Flames of Alzheimer's Disease Brains? Misfolded Tau on the Crossroad of Neurodegenerative and Inflammatory Pathways. *J. Neuroinflammation* 9, 47. doi:10.1186/1742-2094-9-47

Conflict of Interest: Author RS was employed by AXON Neuroscience R&D Services SE.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 30 November 2021
doi: 10.3389/fmolb.2021.772855

Water Accessibility Refinement of the Extended Structure of KirBac1.1 in the Closed State

*Reza Amani[1], Charles D. Schwieters[2], Collin G. Borcik[1], Isaac R. Eason[1], Ruixian Han[3], Benjamin D. Harding[3,4] and Benjamin J. Wylie[1]**

[1]Texas Tech University, Department of Chemistry and Biochemistry, Lubbock, TX, United States, [2]Computational Biomolecular Magnetic Resonance Core, National Institutes of Digestive Diseases and Kidneys, NIH, Bethesda, MD, United States, [3]University of Wisconsin-Madison, Department of Biochemistry and Chemistry, Madison, WI, United States, [4]Biophysics Program, University of Wisconsin at Madison, Madison, WI, United States

NMR structures of membrane proteins are often hampered by poor chemical shift dispersion and internal dynamics which limit resolved distance restraints. However, the ordering and topology of these systems can be defined with site-specific water or lipid proximity. Membrane protein water accessibility surface area is often investigated as a topological function *via* solid-state NMR. Here we leverage water-edited solid-state NMR measurements in simulated annealing calculations to refine a membrane protein structure. This is demonstrated on the inward rectifier K^+ channel KirBac1.1 found in *Burkholderia pseudomallei*. KirBac1.1 is homologous to human Kir channels, sharing a nearly identical fold. Like many existing Kir channel crystal structures, the 1p7b crystal structure is incomplete, missing 85 out of 333 residues, including the N-terminus and C-terminus. We measure solid-state NMR water proximity information and use this for refinement of KirBac1.1 using the Xplor-NIH structure determination program. Along with predicted dihedral angles and sparse intra- and inter-subunit distances, we refined the residues 1–300 to atomic resolution. All structural quality metrics indicate these restraints are a powerful way forward to solve high quality structures of membrane proteins using NMR.

Keywords: solid state NMR, membrane protein, xplor-NIH, water-edited spectroscopy, structure refinement, potassium channel

OPEN ACCESS

Edited by:
Loren B. Andreas,
Max Planck Institute for Biophysical Chemistry, Germany

Reviewed by:
Kristaps Jaudzems,
Latvian Institute of Organic Synthesis (LAS), Latvia
Lynmarie Thompson,
University of Massachusetts Amherst, United States

***Correspondence:**
Benjamin J. Wylie
benjamin.j.wylie@ttu.edu

Specialty section:
This article was submitted to Structural Biology, a section of the journal Frontiers in Molecular Biosciences

Received: *08 September 2021*
Accepted: *08 November 2021*
Published: *30 November 2021*

Citation:
Amani R, Schwieters CD, Borcik CG, Eason IR, Han R, Harding BD and Wylie BJ (2021) Water Accessibility Refinement of the Extended Structure of KirBac1.1 in the Closed State. Front. Mol. Biosci. 8:772855. doi: 10.3389/fmolb.2021.772855

INTRODUCTION

Solid-state NMR (SSNMR) is essential to the structural and functional characterization of membrane proteins (MPs) (Schubeis et al., 2018; Radoicic et al., 2014; Wylie et al., 2016; Mandala et al., 2018; Tran et al., 2020). SSNMR can study MPs in native or native-like environments, allowing site-specific analysis of protein structure and activity. SSNMR is not inherently limited by the size of system, an issue for liquid-state NMR. SSNMR can thus access proteins in proteoliposomes and cellular envelopes (Renault et al., 2012). SSNMR does not require high salt concentrations, long-range order, or cryogenic temperatures, all required for X-ray crystallography. Over the past two decades the water accessible surface of MPs was actively quantified *via* SSNMR (Kumashiro et al., 1998; Ader et al., 2009; Li et al., 2010; Su et al., 2011; Hornig et al., 2013). Over this time, water-edited SSNMR examined the rearrangement of membrane proteins, molecular motion in deuterated samples, and determined membrane insertion topology (Najbauer et al., 2019). In pursuit of functional states of K^+ channels, Ader *et al.* used water-edited SSNMR spectroscopy to unambiguously uncover a dramatic

increase in water-accessible surface area between the closed/inactivated state and open/activated states of the KcsA-Kv1.3 chimera. Subsequently, Borcik *et al.* discovered site-specifically that water accessibly is diminished upon activation of KirBac1.1. This work proposed a key component of the KirBac1.1 gating mechanism, where C-terminal domains rotate and form electrostatic contacts to stabilize the activated state. Thus, relative solvent accessibility during the K^+ channel gating cycle may not be universal. Despite the utility and wide usage of water-edited SSNMR spectroscopy, site-specific solvent accessibility has never been utilized to solve or refine the structure of an MP. Here, we demonstrate the applicability of water-edited SSNMR spectroscopy as an experimental restraint to refine the structure of a MP within Xplor-NIH (Schwieters et al., 2018) simulated annealing calculations.

KirBac1.1 is a 149.02 kDa homotetrameric membrane protein native to *Burkholderia pseudomallei.* Like all inward-rectifier K^+ (Kir) channels, it favors inward potassium ion conductance through the membrane, helping to set the resting membrane potential (Kuo, 2003; Cheng et al., 2009; Wang et al., 2009; Linder et al., 2015). KirBac1.1 retains the characteristic TVGYG selectivity filter motif found in most K^+ channels. This facilitates K^+ conduction near the rate of free diffusion and is impermeable to Na^+ and smaller cations. Each KirBac1.1 monomer consists of two transmembrane (TM) helices, a slide helix, selectivity filter loop, pore helix, and a gating bundle. KirBac1.1 is activated by the association of anionic lipids to a large cationic binding pocket rich in arginine residues (Enkvetchakul et al., 2007; Clarke et al., 2010a; Wang S. et al., 2012; Borcik et al., 2020; van Aalst et al., 2020). Many regions of the protein are intimately tied to channel function and activity, including transmembrane helix 1 (TM1), transmembrane helix 2 (TM2), the slide helix, and the C-terminal gating bundle (Kuo et al., 2003; Enkvetchakul et al., 2004; Enkvetchakul et al., 2007; Paynter et al., 2010; Amani et al., 2020; Borcik et al., 2020). However, to uncover the complete structure-activity relationship of the gating cycle requires a more complete full-length structure. Unfortunately, the existing crystal structures 1p7b lacks the N-terminus (residues 1:35), several turns and coils in the gating bundle (residues 196:205, 290:295) and the C-terminus (residues 310:333) and the crystal structure 2wll lacks 5:37, 200:205, 290:295, 310:333 (Kuo, 2003; Clarke et al., 2010b). Thus, a full-length structure could provide needed information, including pivotal inter-subunit contacts between N-termini and the adjacent cytoplasmic subunit. In addition, it is known the orientations of these regions may change with lipid environment and may be sensitive to salt concentration as they are highly electrostatic. In our previous studies, we characterized the inactivated (closed) and activated (open) states of KirBac1.1 in great detail. We assigned the chemical shifts for both states in activating and inactivating bilayers (Amani et al., 2020) and identified domain motions correlating to both states (Borcik et al., 2020). We found that the water accessible surface of the Kir domain of the closed state was significantly greater than the activated state. These studies motivated this work, as we seek to leverage our acquired knowledge to probe distinct states of the channel structurally. The closed ground state of the channel is the logical starting point in the structural mapping of this Kir channel. It has a greater overall water accessible surface and is the starting point of the gating and thermodynamic cycle of the channel. Thus, full-length structures of KirBac1.1, and many other MPs, will benefit from SSNMR analysis and structure elucidation that recognize their unique topologies.

We refined the structure of KirBac1.1 from residues 1 to 301 using the following workflow: We first modelled in all missing regions of the 1p7b crystal structure using a ROSETTA remodel "quick and dirty" protocol (Huang et al., 2011). Our previously reported ^{15}N and ^{13}C chemical shift assignments for residues 1 to 301 (Amani et al., 2020; Borcik et al., 2020) for the closed-state of U-^{15}N,^{13}C-KirBac1.1 reconstituted into zwitterionic 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipid bilayers were used to generate dihedral angles in TALOSN (Shen and Bax, 2013). We then acquired water-edited SSNMR spectra of U-^{15}N,^{13}C-KirBac1.1 in POPC proteoliposomes and extensively site-specifically assigned a spectrum with a short $^1H_{water}$-$^1H_{protein}$ mixing time. To provide sparse distance restraints, we acquired a three-dimensional (3D) dipole-assisted rotational resonance (DARR) (Takegoshi et al., 2001) spectrum with 50 and 500 ms mixing times during the first and second mixing periods (Zhou et al., 2006). This spectrum yielded several key inter- and intra-subunit distances. We then utilized the TALOSN dihedral angles, sparse distances, and site-specific solvent accessibility measurements to refine the full-length model of KirBac1.1 within Xplor-NIH (Schwieters et al., 2018). Water-based paramagnetic resonance restraints had previously been used as solvent accessibility restraints. Here, Xplor-NIH's PSolPot term was used to fit SSNMR-style solvent accessibility water-accessible surface area data of KirBac1.1 This work represents one of the largest protein structures ever refined using SSNMR solvent accessible surfaces as a restraint.

MATERIALS AND METHODS

SSNMR Sample Preparation

U-^{15}N,^{13}C-labeled KirBac1.1 was expressed and purified as described previously (Amani et al., 2020; Borcik et al., 2020). Briefly, the protein was expressed from *E. coli* in M9 minimum media enriched with $^{15}NH_4Cl$, ^{13}C-glucose, and a 10 ml aliquot of 10X concentrated BioExpress (Cambridge Isotopes Laboratories, Tewksbury, MA 01876) (Bhate et al., 2013; Amani et al., 2020; Borcik et al., 2020). Protein overexpression was induced at an OD_{600} of 0.8 by adding isopropyl β-D-1-thiogalactopyronoside (IPTG) to a concentration of 1 mM. After 16 h of induction at 18°C, cells were harvested *via* centrifugation. Cells were lysed *via* homogenization at 10–15 kpsi. The protein was extracted by adding decyl-β-D-maltopyranoside (DM) to a 30 mM concentration and leaving the lysate on an orbital shaker for 4 h in the presence of Pierce™ Protease inhibitors tablets, EDTA-Free (Thermo Scientific). After extraction, supernatant was spun in an ultracentrifuge, sterile filtered, and loaded onto a 5 ml HisTrap (GE Healthcare Life Sciences) column. The sample was subsequently passed through a HiPrep 26/10 desalting column

(GE Healthcare Life Sciences), followed by a HiLoad 16/600 Superdex 200 size exclusion column (GE Healthcare Life Sciences). Purified protein was mixed with CHAPS solubilized POPC at a 1:1 ratio (w/w). The sample was then stepwise reconstituted *via* the slow addition of BioBeads-SM2 (Bio-Rad, Hercules, CA). BioBeads were then removed and the sample pelleted *via* centrifugation and packed into a 3.2 mm limited speed PENCIL rotor (Revolution NMR, Ft. Collins, CO).

NMR Spectroscopy

All SSNMR spectra were acquired at field strengths of either 17.6 T (750 MHz ^{1}H frequency) or 14.1 T (600 MHz ^{1}H frequency) on SSNMR spectrometers located at National Magnetic Resonance Facility at Madison (NMRFAM, University of Wisconsin, Madison, WI). The CCC 3D DARR spectrum (Zhou et al., 2006) was acquired with 50 and 500 ms of DARR mixing in the first and second mixing periods, respectively, at 750 MHz with a Varian (Fort Collins, CO) 3.2 mm Balun probe in double resonance mode ^{1}H-^{13}C mode. Magic-angle spinning (MAS) (Andrew et al., 1958; Lowe, 1959) was performed at 12.5 kHz with a variable temperature (VT) set point of −30°C and a flow rate of 40 lpm (calibrated to −15 ± 3°C). This temperature was chosen because it provided the greatest overall signal for this 3D experiment. 83 kHz of SPINAL-64 (Fung et al., 2000) ^{1}H decoupling was applied during all chemical shift evolution periods, and hard 90° pulses were 2.4 μs for ^{1}H and 3.05 μs ^{13}C. Polarization transfer was facilitated *via* adiabatic cross polarization (CP) (Pines et al., 1972) with a 1 ms contact time. During CP 1H power was set to 78 kHz and ^{13}C power set to 65 kHz. The recycle delay was set to 1.5 s. The three-dimensional (3D) data was acquired with non-uniform sampling of the indirect dimensions, with a 256 × 256 grid of acquired points with 12.5% points acquired corresponding to 35.4% sampled points in each dimension.

The water accessibility experiments were performed at a magnetic field strength of 14.1 T (600 MHz ^{1}H frequency). The rotor was placed in a 3.2 mm Varian (Fort Collins, CO) T3 HXY probe in double resonance mode, and spun at the magic angle at a spinning rate of 12.5 kHz. To ensure all water surrounding the protein was liquid, the VT was set to 10 C (sample temperature of 25 ± 3°C) for all water edited experiments with a flow rate of 40 lpm. The cross-peaks in these spectra were matched to similar 2D spectra acquired at −5°C and −15°C to confirm no major chemical shift differences. In our past work, KirBac1.1 was assigned over this temperature range to facilitate this process. Pulse widths of 2.7 and 2.55−μs were applied to 1H and 13C, respectively. A 1.5 s recycle delay was implemented for all water edited experiments. Water-edited spectra were acquired using an initial 1H T_2 filter of 1.5 ms, to eliminate ^{1}H polarization arising from protein and lipid signals. 1H to 13C transfer was mediated *via* cross polarization with spin lock fields of 65 and 84 kHz on ^{1}H and ^{13}C, respectively, for a contact duration of 1 ms. Additional parameters for the water edited spectra include a 50 ms DARR mixing followed by 70 kHz of ^{1}H SPINAL-64 decoupling. We assessed the water accessibility with ^{1}H-^{1}H spin diffusion times of 4 and 16 ms.

Structure Calculation *via* Xplor-NIH

Throughout all calculations, strict C4 symmetry was maintained using the symSimulation facility (Schwieters et al., 2018), and subunit backbone geometry of residues 36–301 was restrained to that of 1p7b using a non-crystallographic term allowing up to 1 Å of deviation with zero energy penalty. An additional NCS term was employed between the centroids of opposite subunits to prevent overall expansion. Energy terms employed during structure calculations included ^{13}C-^{13}C intra- and inter-subunit distances (NOE potential), TALOSN derived dihedral angles (CDIH) (Bermejo and Schwieters, 2018), the hydrogen bond potential of mean force (HBPot) (Schwieters et al., 2020), and either the EEFx (Tian et al., 2014; Tian et al., 2015) or EEFx with IMMx (Tian et al., 2017) terms which both model realistic non-bonded interactions within implicit solvent. In their current implementation the IMMx potential builds upon the EEFx potential by including terms explicitly defining the hydrophobic thickness of the bilayer and its dielectric properties. The bilayer dielectric is adjustable and can be scaled differently during initial structural solution and the final refinement. In each calculation, the backbone dihedral angles of residues 1–35 and 302–333 were randomized then relaxed into non-clashing conformations employing the repulsion-only RepelPot term (Schwieters et al., 2018) using gradient minimization, followed by 40 ps of high-temperature (3500 K) molecular dynamics. During this initial repulsion-only phase, EEFx and IMMx not enabled, as they are not stable in the presence of initial close-contacts. The nonbonded representation was then switched over to the implicit model and 30 ps of molecular dynamics was run. This was followed by annealing to 25 K using EEFx or EEFx with IMMx. Following initial calculations, refinement was performed including the PSolPot term (Wang Y. et al., 2012; Gong et al., 2018; Kooshapur et al., 2018) representing site-specific protein-water interactions along the other restraints in a procedure identical to that above with the exception that there is no torsion angle randomization. We performed three PSolPot calculations. In the first we only used completely unambiguously assigned solvent-accessible residues. This generated an ensemble of structures with improved overall structural resolution. These structures were then used to aid in assigning ambiguous water-proximal resonances. The complete set of water-accessible restraints were then used to refine the ensemble of structures. At the end we ran another simulated annealing structural refinement with several Ramachandran outliers deleted from the PSolPot table. The result of last set of PSolPot calculation showed small improvement in some cases. Structure quality was assessed by MolProbity (Williams et al., 2018). All RMSDs were measured *via* VMD-XPLOR (Schwieters and Clore, 2001).

RESULTS AND DISCUSSION

SSNMR Data

NMR structures are solved by including distance measurements and other structural restraints as pseudopotentials into simulated annealing calculations. However, as proteins grow larger, spectral

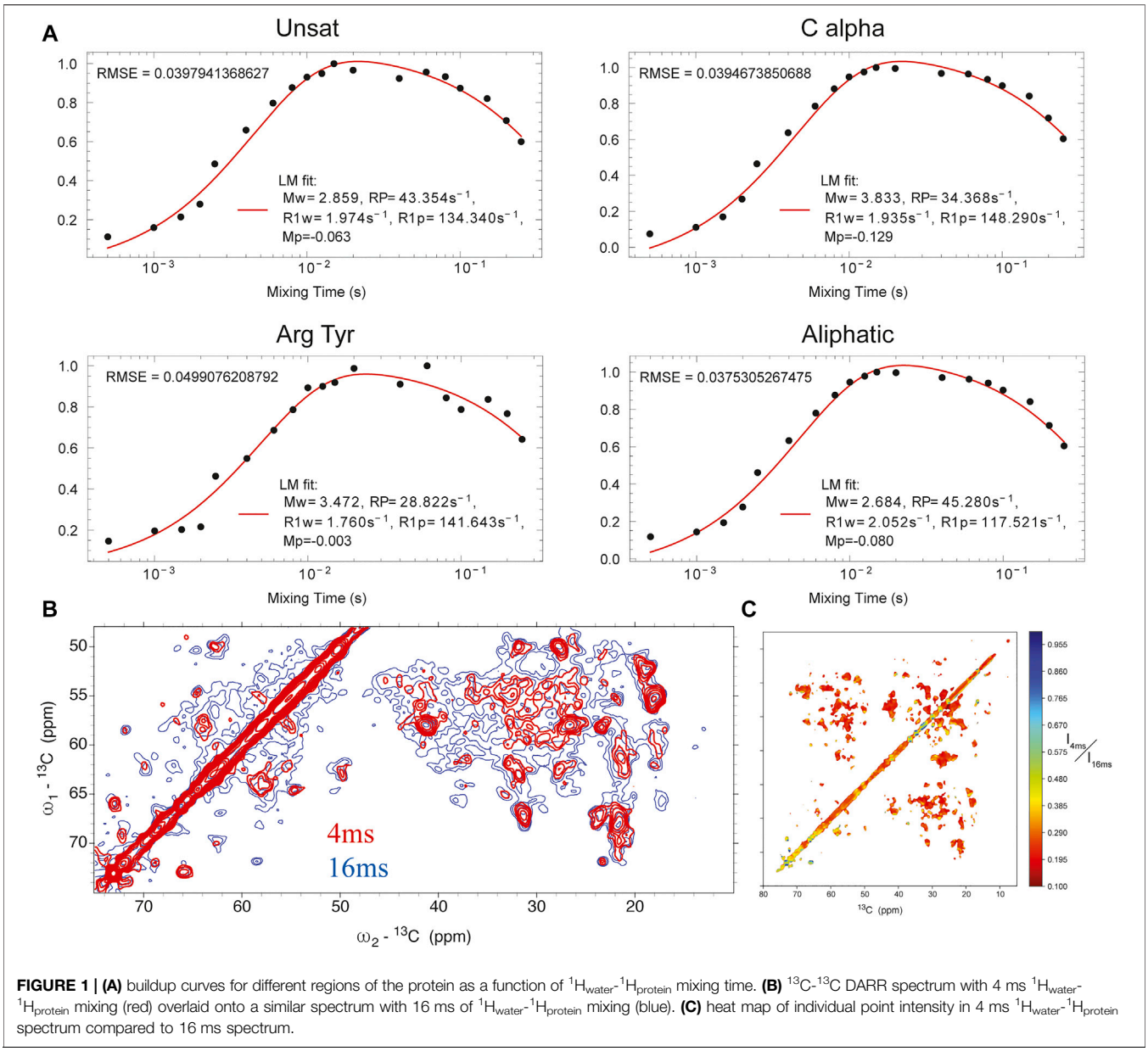


FIGURE 1 | (A) buildup curves for different regions of the protein as a function of $^{1}H_{water}$-$^{1}H_{protein}$ mixing time. **(B)** ^{13}C-^{13}C DARR spectrum with 4 ms $^{1}H_{water}$-$^{1}H_{protein}$ mixing (red) overlaid onto a similar spectrum with 16 ms of $^{1}H_{water}$-$^{1}H_{protein}$ mixing (blue). **(C)** heat map of individual point intensity in 4 ms $^{1}H_{water}$-$^{1}H_{protein}$ spectrum compared to 16 ms spectrum.

crowding will occur. This is compounded when the protein structure is dominated by a single type of secondary structure, as often occurs in α-helical membrane proteins. Thus, as more distances are measured more peaks appear leading to greater information at the cost of reduced site-specific resolution. However, following observations reported by several groups and within our own laboratory, we found that large domain motions may be mapped by the observable solvent accessible surface (Borcik et al., 2020).

Water-edited SSNMR. We measured the solvent-accessible surface of the closed state of the I131C mutant of KirBac1.1 in POPC bilayers using water-edited SSNMR. These water-edited spectra of U-^{15}N,^{13}C-KirBac1.1 are similar to spectra described previously, but they probe the native closed state rather than the closed state of the R49/151/153/Q mutant (Borcik et al., 2020). This technique capitalizes on the great disparity in ^{1}H transverse relaxation times (T_2) between water and protons within the protein, where ^{1}H signal persists for a much longer time within the water bath. Thus, using a T_2 filter we can actively select the ^{1}H signal originating from the surrounding water. This signal is transferred to the protein *via* spin diffusion. This polarization transfer follows a characteristic buildup curve obeying a rate equation we adapted previously (Borcik et al., 2020; Luo and Hong, 2010; Najbauer et al., 2019). Representative buildup curves are depicted in **Figure 1A**. These buildup curves exhibit a good overall fit to our derived rate equation (**Eq. 1**) (Borcik et al., 2020). As depicted in **Figure 1B** in red, 4 ms of $^{1}H_{water}$-$^{1}H_{protein}$ mixing is a good representation of solvent-exposed residues. To better understand the water accessible surface, we also acquired a spectrum with 16 ms of ^{1}H-^{1}H

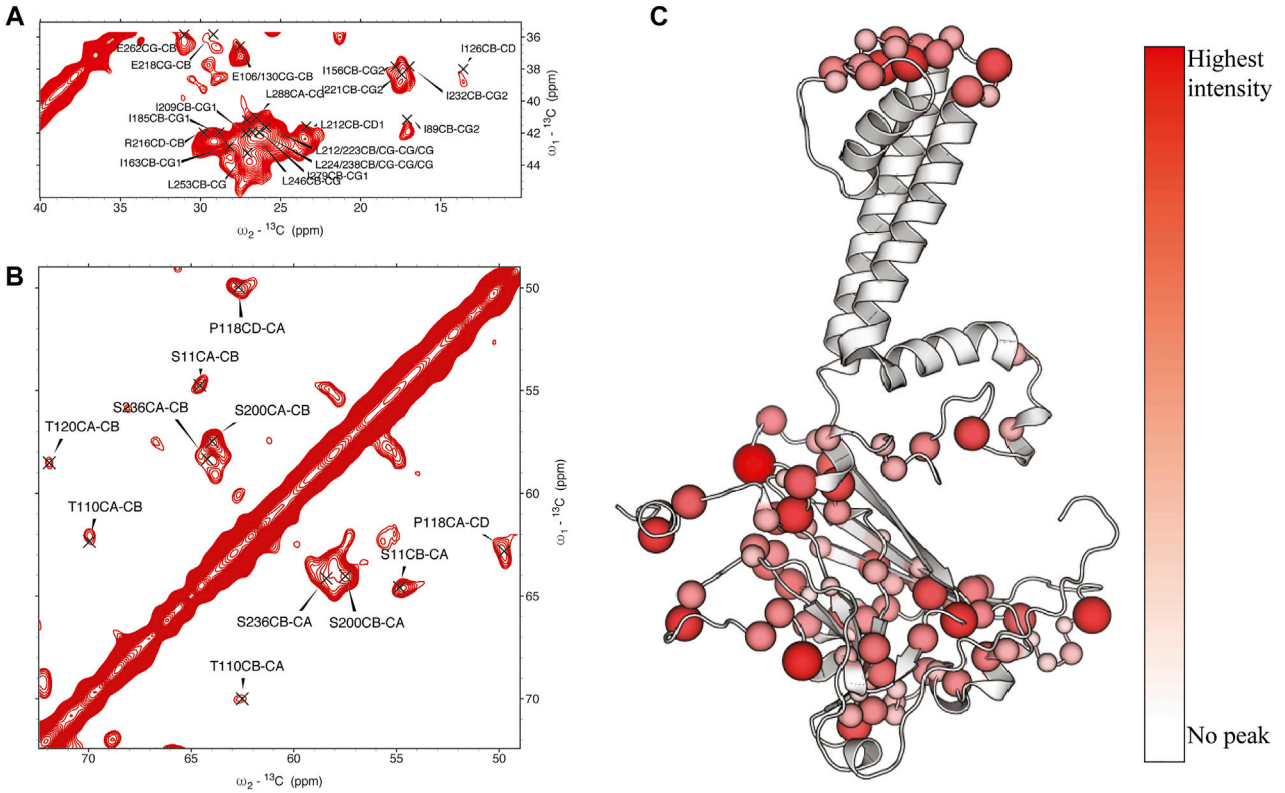


FIGURE 2 | (A, (B) assigned regions of ^{13}C-^{13}C water-edited spectrum with 4 ms $^{1}H_{water}$-$^{1}H_{protein}$ mixing, **(C)** position of assigned residues on the structure of KirBac1.1 with relative intensity represented in color and size of spheres.

spin diffusion. The contoured difference in these spectra is depicted in **Figure 1C**. At 16 ms of $^{1}H_{water}$-$^{1}H_{protein}$ spin-diffusion more embedded parts of protein appear in the spectra, consistent with fit buildup curves presented in **Figure 1A**. With 16 ms of $^{1}H_{water}$-$^{1}H_{protein}$ mixing we observe most resonances in a standard DARR spectrum without a T_2 filter, further indicating large spin-diffusion coverage (**Supplementary Figure S1**).

$$M_p(t_m) = M_w\left(\frac{2R_p}{R_{1p} + 2R_p - R_{1w}}\right)\left(e^{-R_{1w}t_m} - e^{-(R_{1p}+2R_p)t_m}\right) \quad (1)$$

In **Eq. 1** the p index specifies protein and w specifies water. M is magnetization on the specified chemical species at mixing time t_m, and R is the rate of longitudinal cross relaxation for the specified species.

In KirBac1.1, we consistently found the best $^{1}H_{water}$-$^{1}H_{protein}$ mixing time for surface residues to be 4 ms. We were able to assign many sites in these spectra (**Figure 2**). Initially, 51 unambiguous water-edited peaks were assigned based upon our chemical shift assignments for this state of the protein. After multiple iterations of structure refinement, the initial structures helped us to assign an additional 187 ambiguous peaks for a total of 238 solvent-accessibility restraints (**Supplementary Figure S2**) as described below. However, many solvent accessible peaks, especially methyl groups, remained too degenerate for reasonable assignment. However, three- and four-dimensional versions of these pulse sequences may resolve this ambiguity in even more challenging membrane protein systems.

^{13}C-^{13}C-^{13}C 3D spectrum. We obtained sparse distance restraints for tertiary and quaternary structure from a CCC 3D spectrum with 50 and 500 ms of DARR mixing during the first and second mixing periods. The 3D cross peaks were assigned based upon our reported 3D chemical shift assignments. Only completely unambiguous cross peaks were assigned, providing 54 intra subunit distances and 3 inter subunit distances. Given that the sample was uniformly ^{13}C enriched, this limited number of distances was expected. More extensive distance assignments would require significantly less isotopic enrichment to provide the needed resolution.

Initial structural calculations. We started our structure refinement process by generating structures using dihedral angles and distance information in Xplor-NIH. The protocol described above in which the PSolPot term was not used in the initial phase was necessitated by difficulties the term has in representing the very extended structures obtained during initial randomization. Using our previously reported ^{13}C chemical shift assignments, we determined backbone dihedral angles in TALOSN. The prediction resulted in 502 dihedral angle restraints (φ, ψ). In the first step of structural refinement, 100 structures were generated using Xplor-NIH version 3.2.9 with 502

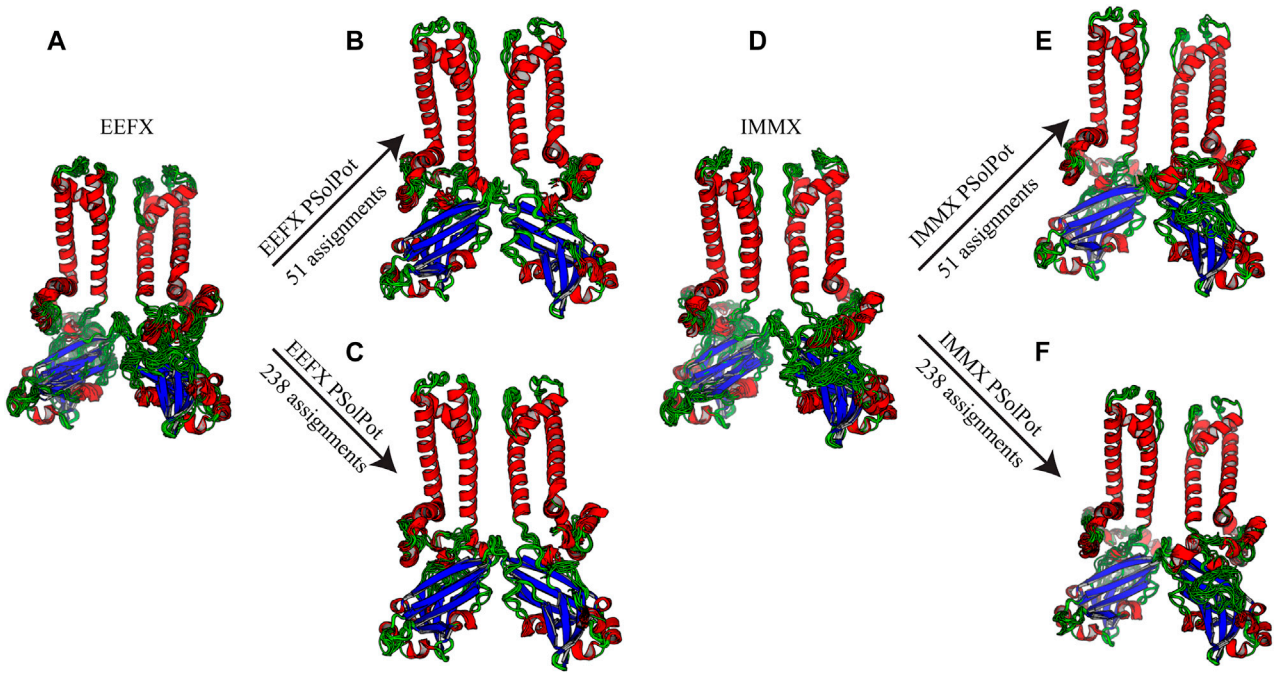


FIGURE 3 | Ensemble of ten lowest structure in each step of calculation **(A)** No water-edited restraints with EEFx potential term, **(B)** 51 unambiguous water-edited restraints with EEFx potential, **(C)** 238 ambiguous and unambiguous water-edited restraints with EEFx potential term, **(D)** No water-edited restraints with IMMx potential term, **(E)** 51 unambiguous water-edited restraints with IMMx potential term, **(F)** 238 ambiguous and unambiguous water-edited restraints with IMMx potential term.

TABLE 1 | Structural statistics for calculations with EEFx potential.

Structure statistics	No PSolPot term	51 PSolPot restraints	238 PSolPot restraints	223 PSolPot restraints
Violations (mean ± σ)				
Bond lengths (A°)	0.0101 ± 0.0003	0.009 ± 0.0	0.009 ± 0.0	0.009 ± 0.0
Bond angles (°)	1.21 ± 0.04	1.196 ± 0.008	1.199 ± 0.007	1.195 ± 0.008
Improper (°)	1.21 ± 0.06	1.05 ± 0.01	1.06 ± 0.01	1.06 ± 0.01
Pairwise r.m.s.d. (A°)				
Heavy atoms (1–301)	3.1 ± 0.8	2 ± 0.2	2 ± 0.1	1.9 ± 0.1
Backbone (1–301)	2.4 ± 1	0.9 ± 0.2	0.9 ± 0.2	0.8 ± 0.2
Heavy atoms (40–282)	3.1 ± 0.7	1.8 ± 0.2	1.9 ± 0.2	1.8 ± 0.1
Backbone (40–282)	2.2 ± 1	0.7 ± 0.3	0.8 ± 0.2	0.7 ± 0.2
Ensemble backbone to crystal structure	3.1 ± 1.3	1.66 ± 0.05	1.68 ± 0.05	1.66 ± 0.05

TABLE 2 | Structural statistics for IMMx potential.

Structure statistics	No PSolPot term	51 PSolPot restraints	238 PSolPot restraints	223 PSolPot restraints
Violations (mean ± σ)				
Bond lengths (A°)	0.0101 ± 0.0003	0.009 ± 0.0	0.009 ± 0.0	0.009 ± 0.0
Bond angles (°)	1.23 ± 0.03	1.24 ± 0.01	1.240 ± 0.006	1.241 ± 0.007
Improper (°)	1.16 ± 0.06	1.12 ± 0.04	1.12 ± 0.03	1.10 ± 0.02
Pairwise r.m.s.d. (A°)				
Heavy atoms (1–301)	3.1 ± 0.4	2 ± 0.1	2.09 ± 0.09	2.1 ± 0.1
Backbone (1–301)	2.2 ± 0.5	0.9 ± 0.2	1.0 ± 0.1	1.0 ± 0.2
Heavy atoms (40–282)	2.9 ± 0.4	1.80 ± 0.08	1.84 ± 0.09	1.83 ± 0.08
Backbone (40–282)	2.0 ± 0.5	0.7 ± 0.1	0.7 ± 0.1	0.7 ± 0.1
Ensemble backbone to crystal structure	2.1 ± 0.1	1.71 ± 0.05	1.75 ± 0.06	1.75 ± 0.07

SSNMR dihedral angles, 54 intra- and 3 inter-subunit SSNMR unambiguous distances, the hydrogen bond potential of mean force (HBPot), and the EEFx potential adapted from CHARMM22 (MacKerell et al., 1998; Tian et al., 2015). The ensemble of the 10 lowest energy structures is presented in **Figure 3A**. An additional 100 structures were generated with identical restraints, but with the addition of the IMMx function to the EEFx potential with membrane parameters of 27.0 for POPC membrane thickness, profileN set to 2, and the delectric screening value or A parameter to 0.85 (calculations that lack the IMMx function are simply called EEFx and the calculations with IMMx added to EEFx potential function are called IMMx). This ensemble is presented in **Figure 3D**. As shown in **Tables 1**, **2**, the calculated pairwise RMSD (pwRMSD) *via* VMD-Xplor for the first step of this structure calculation without water-edited restraints, are 2.4 ± 1 Å and 2.2 ± 0.5 Å for backbone residues 1 to 301 of the EEFx and IMMx calculations respectively. The pwRMSD for the backbone of the well-ordered regions of the protein (residues 40–282) improves to 2.2 ± 1 and 2 ± 0.5 for the EEFx and IMMx calculations respectively. Although these pwRMSD are acceptable for this level of experimental dihedral angle and distance restraints, there is significant room for improvement. This improvement in the quality of the calculated structures shows the importance of water-edited restraint usage for the structure calculation.

Water-accessibility restraints. Previous applications of solvent accessibility as a structure refinement tool utilized solution based paramagnetic relaxation enhancement (sPRE). Generally, paramagnetic relaxation enhancement (PRE) results from the coupling between a magnetically active nucleus and an unpaired electron. The electron may be a stable radical or metal. This unpaired electronic spin may be bound to the protein or free in solution. This interaction has r^{-6} range dependence. Recently, soluble paramagnetic probes gained popularity. When these moieties contact the surface of the protein, they introduce the sPRE which can thus be tied to solvent accessible surface (S_{Acc}). Surface accessibility restraints were initially incorporated in Xplor-NIH using an empirical expression involving distances to neighboring nuclei, and it was shown to qualitatively represent water-protein interactions in solution and solvent PRE data (Wang et al., 2012). Wang *et al.* found that S_{Acc} can be calculated with a linear equation, where the slope and intercept is a unique function of a specific protein's topology. More recently (Gong et al., 2018; Kooshapur et al., 2018), a more quantitative representation of solvent PRE data has been developed, where the observable is represented by **Eq. 2**. For sPREs this expression is approximate, with the quantitative relationship between **Eq. 2** and solvent PRE being somewhat more complicated (Okuno et al., 2020), and yet this formulation has been employed with some success. In this vein, our residue-based water-edited SSNMR-derived surface area data are fit to values computed from molecular structure using **Eq. 2**. In keeping with the qualitative nature of the representation, the corresponding Xplor-NIH energy term depends only on the correlation between the two quantities (Gong et al., 2018). Gong *et al.* and Kooshapur *et al* proposed a grid search algorithm to determine the accessible surface. This included a protein surface integral that can be written in form of a tessellation composed of triangular patches (**Eq. 2**). In **Eq. 2**, k' is a constant prefactor, n is the outward-facing distance normal surface, and r is the distance from this surface to a nucleus:

$$\Gamma_{sPRE} = \frac{-k'}{9}\sum_i a_i n_i . \frac{r_i}{|r_i|^6} \tag{2}$$

They incorporated these concepts into sPRE module and energy potential (PSolPot) to include sPRE data in Xplor-NIH simulated annealing calculations. This potential was shown to be quite effective in direct structure refinement (Gong et al., 2018; Kooshapur et al., 2018).

We now present a new application of the PSolPot potential function to refine protein structures using water-edited SSNMR spectroscopy derived restraints. Water-edited SSNMR identifies the accessible surface of the protein with a similar r^{-6} distance dependence. As described above, previous studies found that the overall surface area of the water-protein interface can be expressed by **Eq. 3**

$$S_{Acc} = V_P\sqrt{\frac{\pi}{D_{eff}t_m^s}} \tag{3}$$

Where S_{Acc} is the surface area of the water-protein interface, t_m^s is the time of mixing until saturation, V_P is the volume of the protein, and D_{eff} is the effective diffusion parameter. This equation provides a global picture rather than a site-specific view of the water-protein interface. Following Andreas et al. (Najbauer et al., 2019), we previously found the water-to-protein polarization transfer could be defined by a longitudinal cross relaxation-dependent rate equation stated above (**Eq. 1**).

It has been long known (Bloembergen et al., 1948) that the relaxation term for longitudinal cross relaxation depends on the ^{1}H-^{1}H dipolar coupling that has the form of **equation 4** from dipolar alphabet.

$$\langle H_{loc}^2 \rangle_{Av} = \frac{1}{3}\gamma^2\hbar^2 I(I+1)\sum_j \left(1 - 3\cos^2\theta_{ij}\right)^2 r_{ij}^{-6} \tag{4}$$

Thus, because of the similar r^{-6} dependence, we found the PSolPot potential could accommodate our restraints after modification.

As shown in **Figure 2**, our extensive chemical shift assignments of water-edited spectra provide restraints for nearly half the protein (full assignments of the aliphatic region are shown in **Supplementary Figure S2**). Because PSolPot is a correlation function that fits the water accessible surface, the relative signal intensity of each peak can be used as the data input for structural refinement. The chemical shift assignments of the water-edit spectra were performed in two rounds. In the first round, the integrated intensity of resolved unambiguous peaks were used for structure refinement. These 51 assignments were used to refine two sets of 100 structures starting from the 10 lowest energy structures of the EEFx (**Figure 3B**) and IMMx (**Figure 3E**) calculations respectively. In the second round, we included the integrated intensity of all possible assignments, corresponding to 238 total PSolPot restraints. This provided two additional sets of 100 structures starting from the same

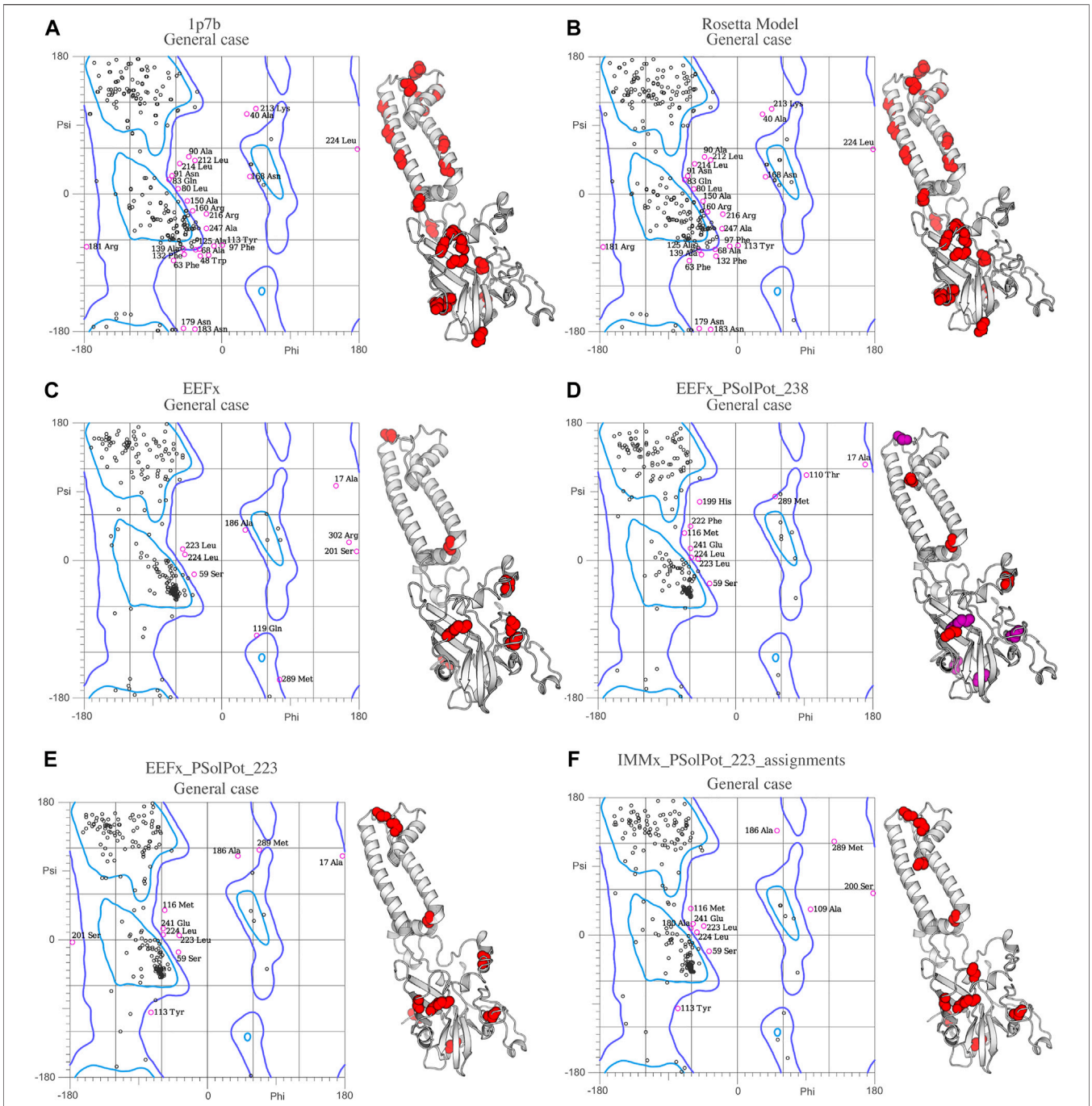


FIGURE 4 | Ramachandran space of **(A)** Crystal structure 1p7b, **(B)** Initial Rosetta model, **(C)** Lowest structure of EEFx run, **(D)** The lowest energy structure of PSolPot, residues in magenta are the outlier residues in Ramachandran space that has been deleted in the last round of structure calculation via PSolPot. **(E)** The lowest energy structure without Ramachandran outlier in PSolPot list with EEFx potential term. **(F)** The lowest energy structure without Ramachandran outlier in PSolPot list with IMMx potential term.

PSolPot-free EEFx (**Figure 3C**) and IMXx (**Figure 3F**) ensembles. We then determined the bbRMSD and judged the quality in MolProbity. This indicated the overall structural quality slightly diminished relative to the PSolPot ensemble with only unambiguous restraints (**Table 3** and **Figure 4**). We then deleted 6 Ramachandran outliers and their neighboring residues and repeated the calculations with only 223 PSolPot restraints. This slightly improved the overall structural quality.

We judged the internal consistency of all four final structural ensembles using heavy-atom RMSD, and backbone RMSD. We also judged their objective quality *via* rotameric, Ramachandran, and Z-score analysis in MolProbity. The statistical summary of all

TABLE 3 | Comparison of structure quality performed on MolProbity for the generated structures to initial structures.

	Favored rotamers (%)	Ramachandran favored (%)	Rama distribution Z-score (less than 2 is the goal)
Crystal structure 1p7b	66.2	60.9	−6.7 ± 0.3
Rosetta Model	74.3	66.5	−6.0 ± 0.2
Lowest generated EFFx structure	82.9	85.2	−1.7 ± 0.2
Lowest EEFx-PSolPot 51 assignments	88.1	87.6	−0.8 ± 0.2
Lowest EEFx-PSolPot 238 Assignments	87.0	86.4	−1.4 ± 0.2
Lowest EEFx-PSolPot 223 Assignments	91.5	87.3	−0.8 ± 0.2
Lowest IMMx structure	85.5	84.6	−1.6 ± 0.2
Lowest IMMx-PSolPot 51 assignments	89.6	87.6	−0.7 ± 0.2
Lowest IMMx-PSolPot 238 Assignments	91.1	85.2	−0.5 ± 0.2
Lowest IMMx-PSolPot 223 Assignments	92.6	85.5	−0.7 ± 0.2

four stages of simulated annealing with EEFx are provided in **Table 1**. These calculations include the ensemble with only EEFx, the addition of 51 unambiguous assignments, 238 assignments (ambiguous and unambiguous), and finally 223 assignments (ambiguous and unambiguous without Ramachandran outliers). In **Table 2** the same statistics are listed after IMMx is included in the calculation. As mentioned above, the addition of the unambiguous water-edited restraints in structure calculation improved the structural quality dramatically. Using the EEFx forcefield and the 51 unambiguous restraints, the pwRMSD for the protein backbone improved from 2.4 ± 1 Å to 0.9 ± 0.2 Å (**Table 1**) for residues 1 to 301. In the well-ordered regions, residues 40 to 282, the pwRMSD improved from 2.2 ± 1 Å to 0.7 ± 0.3 Å. When the IMMx membrane potential is added to the calculation the pwRMSD improved from 2.2 ± 0.5 Å to 0.9 ± 0.2 Å for the first 301 residues, and from 2.0 ± 0.5 Å to 0.7 ± 0.1 Å for the well-ordered regions (**Table2**). In most cases the utilization of IMMx produces improvements relative to EEFx within the hydrophobic region of the protein. The addition of unambiguous water-edited restraints did not result in a significant improvement in the structure, where the pwRMSD slightly diminished. Deleting identified Ramachandran violators only improved the pwRMSD slightly.

Overall, the objective structural quality improved with the addition of EEFx, IMMx, and PSolPot restraints as judged by MolProbity. As stated above, the 1p7b crystal structure lacks 75 residues (22.5% of the residues in WT-KirBac1.1 sequence). Our initial Rosetta model includes all 75 residues missing from 1p7b. Out of the 75 residues missing in the X-ray crystal structure, we have experimental SSNMR restraints for 51 residues. The final 24 residues of the protein remain unrestrained. Despite the incomplete information on end of the C-terminus of protein, the best structural ensembles possess up to 92.6% favored rotameric scores and up to 87.6% of residues occupy most favored regions of Ramachandran space. Ramachandran plots and the location of Ramachandran outliers is depicted in **Figure 4**. As shown in **Table 3**, the inclusion of EEFx and the 51 unambiguous PSolPot restraints dramatically improved the structural quality compared to both the 1p7b X-ray structure and the starting Rosetta model. The population of residues in favored Ramachandran and rotameric space increases from 66.5 to 74.3% to 85.2 and 82.9% when EEFx terms are included. The addition of 51 PSolPot restraints further improves these statistics to 87.6 and 88.1% favored occupancy. As further depicted in **Table 3** and **Figure 4**, the inclusion of all possible PSolPot restraints does improve upon structural ensembles without these restraints, but slightly deteriorates, indicating that PSolPot is very capable of improving structures with a set of high-quality but relatively sparse restraints, compared to a much longer list of lower-quality information. However, after fully analyzing the structures solved with 238 restraints, and comparing Ramachandran violators to overlapped regions of solvent accessible spectra, we deleted several Ramachandran space violators. This improved the favored rotamer percentage and favored Ramachandran percentage to 91.5 and 87.6%. This further indicates that PSolPot is best implemented with high-quality rather than high-quantity restraints. After significant data quality control, only marginal improvement over the unambiguous structure is observed.

CONCLUSION

Structural refinement of KirBac1.1 was performed using predicted dihedral angles from SSNMR chemical shift assignments, unambiguous distance restraints, and SSNMR water-edited spectroscopy. Calculations were carried out using Xplor-NIH version 3.2.9. The statistical comparison of the 10 lowest energy structures solved with water-edited SSNMR restraints supplied to the PSolPot potential improved both the backbone and all heavy-atom RMSDs relative to ensembles without these restraints. The pair-wise bbRMSD improved from 2.4 Å to 0.9 Å after including PSolPot in the calculation, which is a 62.5% improvement. However, it is clear, that given the nature of the grid search matrix algorithm inherent to PSolPot, relatively sparse but high-quality restraints can create a significant improvement in protein quality. However, including of less-resolved sites in the protein will require significant further analysis. Yet, it is clear even incomplete water-accessibility, and perhaps lipid accessibility, can be a powerful means to structural improvement. Given the difficulty and complexity in solving the structures of transmembrane proteins by NMR, this technique provides a new powerful means to solve and refine the structures of these proteins, which are fundamental to human health. Given the wide application of water-edited SSNMR, this technique could soon reach wide acceptance. In addition, it is clear this method is compatible with the implicit lipid and water

models within Xplor-NIH. It was previously shown that EEFx (Tian et al., 2014; Tian et al., 2015) and IMMx (Tian et al., 2017) were powerful means for *de novo* solution of monomeric membrane proteins. Based upon our results, these forcefields are also applicable to membrane protein oligomers provided the appropriate parameter adjustments during annealing and refinement respectively. In future studies the application of this technique for *de novo* structure determination can be tested.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://doi.org/10.13018/BMR50135, 50135.

AUTHOR CONTRIBUTIONS

RA, CB, RH, BH, and BW designed and performed experiments. CS provided a customized build of Xplor-NIH. RA, IE, and CS performed all simulated annealing calculations. RA, BW, and CS drafted the article.

FUNDING

This research was supported by the National Institutes of Health (Maximizing Investigators' Research Award (MIRA, R35, 1R35GM124979). CS was supported by the Intramural Program of the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health. This study made use of the National Magnetic Resonance Facility at Madison (NMRFAM), which is supported by NIH grant P41GM136463. The 600 MHz solid-state NMR spectrometer at NMRFAM was installed with financial support of the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison, and Chad M. Rienstra collected the data. This study made use of NMRbox: National Center for Biomolecular NMR Data Processing and Analysis, a Biomedical Technology Research Resource (BTRR), which is supported by NIH grant P41GM111135 (NIGMS).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.772855/full#supplementary-material

REFERENCES

Ader, C., Schneider, R., Seidel, K., Etzkorn, M., Becker, S., and Baldus, M. (2009). Structural Rearrangements of Membrane Proteins Probed by Water-Edited Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 131, 170–176. doi:10.1021/ja806306e

Amani, R., Borcik, C. G., Khan, N. H., Versteeg, D. B., Yekefallah, M., Do, H. Q., et al. (2020). Conformational Changes upon Gating of KirBac1.1 into an Open-Activated State Revealed by Solid-State NMR and Functional Assays. *Proc. Natl. Acad. Sci. USA* 117, 2938–2947. doi:10.1073/pnas.1915010117

Andrew, E. R., Bradbury, A., and Eades, R. G. (1958). Nuclear Magnetic Resonance Spectra from a Crystal Rotated at High Speed. *Nature* 182, 1659. doi:10.1038/1821659a0

Bermejo, G. A., and Schwieters, C. D. (2018). Protein Structure Elucidation from NMR Data with the Program Xplor-NIH. *Methods Mol. Biol.* 1688, 311–340. doi:10.1007/978-1-4939-7386-6_14

Bhate, M. P., Wylie, B. J., Thompson, A., Tian, L., Nimigean, C., and McDermott, A. E. (2013). Preparation of Uniformly Isotope Labeled KcsA for Solid State NMR: Expression, Purification, Reconstitution into Liposomes and Functional Assay. *Protein Expr. Purif.* 91, 119–124. doi:10.1016/j.pep.2013.07.013

Bloembergen, N., Purcell, E. M., and Pound, R. V. (1948). RELAXATION EFFECTS IN NUCLEAR MAGNETIC RESONANCE ABSORPTION. *Phys. Rev.* 73, 679–712. doi:10.1103/PhysRev.73.679

Borcik, C. G., Versteeg, D. B., Amani, R., Yekefallah, M., Khan, N. H., and Wylie, B. J. (2020). The Lipid Activation Mechanism of a Transmembrane Potassium Channel. *J. Am. Chem. Soc.* 142, 14102–14116. doi:10.1021/jacs.0c01991

Cheng, W. W. L., Enkvetchakul, D., and Nichols, C. G. (2009). KirBac1.1: It's an Inward Rectifying Potassium Channel. *J. Gen. Physiol.* 133, 295–305. doi:10.1085/jgp.200810125

Clarke, O. B., Caputo, A. T., Hill, A. P., Vandenberg, J. I., Smith, B. J., and Gulbis, J. M. (2010a). Domain Reorientation and Rotation of an Intracellular Assembly Regulate Conduction in Kir Potassium Channels. *Cell* 141, 1018–1029. doi:10.1016/j.cell.2010.05.003

Clarke, O. B., Caputo, A. T., Hill, A. P., Vandenberg, J. I., Smith, B. J., and Gulbis, J. M. (2010b). Domain Reorientation and Rotation of an Intracellular Assembly Regulate Conduction in Kir Potassium Channels. *Cell* 141, 1018–1029. doi:10.1016/j.cell.2010.05.003

Enkvetchakul, D., Bhattacharyya, J., Jeliazkova, I., Groesbeck, D. K., Cukras, C. A., and Nichols, C. G. (2004). Functional Characterization of a Prokaryotic Kir Channel. *J. Biol. Chem.* 279, 47076–47080. doi:10.1074/jbc.C400417200

Enkvetchakul, D., Jeliazkova, I., Bhattacharyya, J., and Nichols, C. G. (2007). Control of Inward Rectifier K Channel Activity by Lipid Tethering of Cytoplasmic Domains. *J. Gen. Physiol.* 130, 329–334. doi:10.1085/jgp.200709764

Fung, B. M., Khitrin, A. K., and Ermolaev, K. (2000). An Improved Broadband Decoupling Sequence for Liquid Crystals and Solids. *J. Magn. Reson.* 142, 97–101. doi:10.1006/jmre.1999.1896

Gong, Z., Schwieters, C. D., and Tang, C. (2018). Theory and Practice of Using Solvent Paramagnetic Relaxation Enhancement to Characterize Protein Conformational Dynamics. *Methods* 148, 48–56. doi:10.1016/j.ymeth.2018.04.006

Hornig, S., Ohmert, I., Trauner, D., Ader, C., Baldus, M., and Pongs, O. (2013). Tetraphenylporphyrin Derivative Specifically Blocks Members of the Voltage-Gated Potassium Channel Subfamily Kv1. *Channels* 7, 473–482. doi:10.4161/chan.25848

Huang, P.-S., Ban, Y.-E. A., Richter, F., Andre, I., Vernon, R., Schief, W. R., et al. (2011). RosettaRemodel: a Generalized Framework for Flexible Backbone Protein Design. *PLoS One* 6, e24109. doi:10.1371/journal.pone.0024109

Kooshapur, H., Schwieters, C. D., and Tjandra, N. (2018). Conformational Ensemble of Disordered Proteins Probed by Solvent Paramagnetic Relaxation Enhancement (sPRE). *Angew. Chem. Int. Ed.* 57, 13519–13522. doi:10.1002/anie.201807365

Kumashiro, K. K., Schmidt-Rohr, K., Murphy, O. J., Ouellette, K. L., Cramer, W. A., and Thompson, L. K. (1998). A Novel Tool for Probing Membrane Protein Structure: Solid-State NMR with Proton Spin Diffusion and X-Nucleus Detection. *J. Am. Chem. Soc.* 120, 5043–5051. doi:10.1021/ja972655e

Kuo, A., Gulbis, J. M., Antcliff, J. F., Rahman, T., Lowe, E. D., Zimmer, J., et al. (2003). Crystal Structure of the Potassium Channel KirBac1.1 in the Closed State. *Science* 300, 1922–1926. doi:10.1126/science.1085028

Li, S., Su, Y., Luo, W., and Hong, M. (2010). Water–Protein Interactions of an Arginine-Rich Membrane Peptide in Lipid Bilayers Investigated by Solid-State

Nuclear Magnetic Resonance Spectroscopy. *J. Phys. Chem. B* 114, 4063–4069. doi:10.1021/jp912283r

Linder, T., Wang, S., Zangerl-Plessl, E.-M., Nichols, C. G., and Stary-Weinzinger, A. (2015). Molecular Dynamics Simulations of KirBac1.1 Mutants Reveal Global Gating Changes of Kir Channels. *J. Chem. Inf. Model.* 55, 814–822. doi:10.1021/acs.jcim.5b00010

Lowe, I. J. (1959). Free Induction Decays of Rotating Solids. *Phys. Rev. Lett.* 2, 285–287. doi:10.1103/physrevlett.2.285

Luo, W., and Hong, M. (2010). Conformational Changes of an Ion Channel Detected through Water–Protein Interactions Using Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 132, 2378–2384. doi:10.1021/ja9096219

Maciejewski, M. W., Schuyler, A. D., Gryk, M. R., Moraru, I. I., Romero, P. R., Ulrich, E. L., et al. (2017). NMRbox: A Resource for Biomolecular NMR Computation. *Biophysical J.* 112, 1529–1534. doi:10.1016/j.bpj.2017.03.011

MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., et al. (1998). All-atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* 102, 3586–3616. doi:10.1021/jp973084f

Mandala, V. S., Williams, J. K., and Hong, M. (2018). Structure and Dynamics of Membrane Proteins from Solid-State NMR. *Annu. Rev. Biophys.* 47, 201–222. doi:10.1146/annurev-biophys-070816-033712

Najbauer, E. E., Movellan, K. T., Schubeis, T., Schwarzer, T., Castiglione, K., Giller, K., et al. (2019). Probing Membrane Protein Insertion into Lipid Bilayers by Solid-State NMR. *ChemPhysChem* 20, 302–310. doi:10.1002/cphc.201800793

Nilges, M., Clore, G. M., and Gronenborn, A. M. (1988). Determination of Three-Dimensional Structures of Proteins from Interproton Distance Data by Hybrid Distance Geometry-Dynamical Simulated Annealing Calculations. *FEBS Lett.* 229, 317–324. doi:10.1016/0014-5793(88)81148-7

Okuno, Y., Szabo, A., and Clore, G. M. (2020). Quantitative Interpretation of Solvent Paramagnetic Relaxation for Probing Protein-Cosolute Interactions. *J. Am. Chem. Soc.* 142, 8281–8290. doi:10.1021/jacs.0c00747

Paynter, J. J., Andres-Enguix, I., Fowler, P. W., Tottey, S., Cheng, W., Enkvetchakul, D., et al. (2010). Functional Complementation and Genetic Deletion Studies of KirBac Channels. *J. Biol. Chem.* 285, 40754–40761. doi:10.1074/jbc.M110.175687

Pines, A., Gibby, M. G., and Waugh, J. S. (1972). Proton-Enhanced Nuclear Induction Spectroscopy. A Method for High Resolution NMR of Dilute Spins in Solids. *J. Chem. Phys.* 56, 1776–1777. doi:10.1063/1.1677439

Radoicic, J., Lu, G. J., and Opella, S. J. (2014). NMR Structures of Membrane Proteins in Phospholipid Bilayers. *Quart. Rev. Biophys.* 47, 249–283. doi:10.1017/S0033583514000080

Renault, M., Tommassen-van Boxtel, R., Bos, M. P., Post, J. A., Tommassen, J., and Baldus, M. (2012). Cellular Solid-State Nuclear Magnetic Resonance Spectroscopy. *Proc. Natl. Acad. Sci.* 109, 4863–4868. doi:10.1073/pnas.1116478109

Schubeis, T., Le Marchand, T., Andreas, L. B., and Pintacuda, G. (2018). H Magic-Angle Spinning NMR Evolves as a Powerful New Tool for Membrane Proteins. *J. Magn. Reson.* 287, 140–152. doi:10.1016/j.jmr.2017.11.014

Schwieters, C. D., Bermejo, G. A., and Clore, G. M. (2020). A Three-Dimensional Potential of Mean Force to Improve Backbone and Sidechain Hydrogen Bond Geometry in Xplor-NIH Protein Structure Determination. *Protein Sci.* 29, 100–110. doi:10.1002/pro.3745

Schwieters, C. D., Bermejo, G. A., and Clore, G. M. (2018). Xplor-NIH for Molecular Structure Determination from NMR and Other Data Sources. *Protein Sci.* 27, 26–40. doi:10.1002/pro.3248

Schwieters, C. D., and Clore, G. M. (2001). The VMD-XPLOR Visualization Package for NMR Structure Refinement. *J. Magn. Reson.* 149, 239–244. doi:10.1006/jmre.2001.2300

Shen, Y., and Bax, A. (2013). Protein Backbone and Sidechain Torsion Angles Predicted from NMR Chemical Shifts Using Artificial Neural Networks. *J. Biomol. NMR* 56, 227–241. doi:10.1007/s10858-013-9741-y

Su, Y., Waring, A. J., Ruchala, P., and Hong, M. (2011). Structures of β-Hairpin Antimicrobial Protegrin Peptides in Lipopolysaccharide Membranes: Mechanism of Gram Selectivity Obtained from Solid-State Nuclear Magnetic Resonance. *Biochemistry* 50, 2072–2083. doi:10.1021/bi101975v

Takegoshi, K., Nakamura, S., and Terao, T. (2001). - Dipolar-Assisted Rotational Resonance in Magic-Angle Spinning NMR. *Chem. Phys. Lett.* 344, 631–637. doi:10.1016/S0009-2614(01)00791-6

Tian, Y., Schwieters, C. D., Opella, S. J., and Marassi, F. M. (2015). A Practical Implicit Membrane Potential for NMR Structure Calculations of Membrane Proteins. *Biophysical J.* 109, 574–585. doi:10.1016/j.bpj.2015.06.047

Tian, Y., Schwieters, C. D., Opella, S. J., and Marassi, F. M. (2014). A Practical Implicit Solvent Potential for NMR Structure Calculation. *J. Magn. Reson.* 243, 54–64. doi:10.1016/j.jmr.2014.03.011

Tian, Y., Schwieters, C. D., Opella, S. J., and Marassi, F. M. (2017). High Quality NMR Structures: a New Force Field with Implicit Water and Membrane Solvation for Xplor-NIH. *J. Biomol. NMR* 67, 35–49. doi:10.1007/s10858-016-0082-5

Tran, N. T., Mentink-Vigier, F., and Long, J. R. (2020). Dynamic Nuclear Polarization of Biomembrane Assemblies. *Biomolecules* 10, 1246. doi:10.3390/biom10091246

van Aalst, E., Yekefallah, M., Mehta, A. K., Eason, I., and Wylie, B. (2020). Codon Harmonization of a Kir3.1-KirBac1.3 Chimera for Structural Study Optimization. *Biomolecules* 10, 430. doi:10.3390/biom10030430

Wang, S., Alimi, Y., Tong, A., Nichols, C. G., and Enkvetchakul, D. (2009). Differential Roles of Blocking Ions in KirBac1.1 Tetramer Stability. *J. Biol. Chem.* 284, 2854–2860. doi:10.1074/jbc.M807474200

Wang, S., Lee, S.-J., Heyman, S., Enkvetchakul, D., and Nichols, C. G. (2012a). Structural Rearrangements Underlying Ligand-Gating in Kir Channels. *Nat. Commun.* 3, 617. doi:10.1038/ncomms1625

Wang, Y., Schwieters, C. D., and Tjandra, N. (2012b). Parameterization of Solvent-Protein Interaction and its Use on NMR Protein Structure Determination. *J. Magn. Reson.* 221, 76–84. doi:10.1016/j.jmr.2012.05.020

Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., et al. (2018). MolProbity: More and Better Reference Data for Improved All-Atom Structure Validation. *Protein Sci.* 27, 293–315. doi:10.1002/pro.3330

Wylie, B. J., Do, H. Q., Borcik, C. G., and Hardy, E. P. (2016). Advances in Solid-State NMR of Membrane Proteins. *Mol. Phys.* 114, 3598–3609. doi:10.1080/00268976.2016.1252470

Zhou, D. H., Kloepper, K. D., Winter, K. A., and Rienstra, C. M. (2006). Band-selective 13C Homonuclear 3D Spectroscopy for Solid Proteins at High Field with Rotor-Synchronized Soft Pulses. *J. Biomol. NMR* 34, 245–257. doi:10.1007/s10858-006-0026-6

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

REVIEW
published: 06 December 2021
doi: 10.3389/fmolb.2021.791090

Dihedral Angle Measurements for Structure Determination by Biomolecular Solid-State NMR Spectroscopy

*Patrick C. A. van der Wel**

Solid-state NMR Group, Zernike Institute for Advanced Materials, University of Groningen, Groningen, Netherlands

In structural studies of immobilized, aggregated and self-assembled biomolecules, solid-state NMR (ssNMR) spectroscopy can provide valuable high-resolution structural information. Among the structural restraints provided by magic angle spinning (MAS) ssNMR the canonical focus is on inter-atomic distance measurements. In the current review, we examine the utility of ssNMR measurements of angular constraints, as a complement to distance-based structure determination. The focus is on direct measurements of angular restraints via the judicious recoupling of multiple anisotropic ssNMR parameters, such as dipolar couplings and chemical shift anisotropies. Recent applications are highlighted, with a focus on studies of nanocrystalline polypeptides, aggregated peptides and proteins, receptor-substrate interactions, and small molecule interactions with amyloid protein fibrils. The review also examines considerations of when and where ssNMR torsion angle experiments are (most) effective, and discusses challenges and opportunities for future applications.

Keywords: solid-state NMR, structural biology, amyloid, polyglutamine, protein structures

OPEN ACCESS

Edited by:
Amir Goldbourt,
Tel Aviv University, Israel

Reviewed by:
Guangjin Hou,
Dalian Institute of Chemical Physics (CAS), China
Bernd Reif,
Technical University of Munich, Germany

***Correspondence:**
Patrick C. A. van der Wel
p.c.a.van.der.wel@rug.nl

Specialty section:
This article was submitted to Structural Biology, a section of the journal Frontiers in Molecular Biosciences

Received: *07 October 2021*
Accepted: *22 November 2021*
Published: *06 December 2021*

Citation:
van der Wel PCA (2021) Dihedral Angle Measurements for Structure Determination by Biomolecular Solid-State NMR Spectroscopy. Front. Mol. Biosci. 8:791090. doi: 10.3389/fmolb.2021.791090

INTRODUCTION

In modern integrative structural biology, complementary structure determination methods provide insights into different aspects of the structure, dynamics and co-assembly of biomolecules. Tremendous advances in solid-state NMR (ssNMR) and in particular magic-angle-spinning (MAS) NMR spectroscopy have widened the contribution of this technique to our understanding of protein aggregates and assemblies (Van der Wel, 2018). Descriptions of the process of structure determination via ssNMR spectroscopy often focus on the role of inter-atomic (or rather inter-nuclear) distances. However, NMR-based insights into the local geometry of protein backbones, in the form of dihedral angles, can also be a powerful tool in the NMR-based structural biology arsenal. In this review article we examine MAS ssNMR techniques that allow direct insights into angular constraints that define the structures of proteins and other biomolecules. Readers are also referred to earlier reviews in the mid 2000s (Hong, 2006; Hong and Wi, 2006; Ladizhansky, 2009), and the current review will focus on more recent applications since 2006. We shall also examine the place of these techniques in current and future biomolecular MAS ssNMR. Examples show how these restraints can be effective and even essential tools for structure determination of specific kinds of biological structures and assemblies.

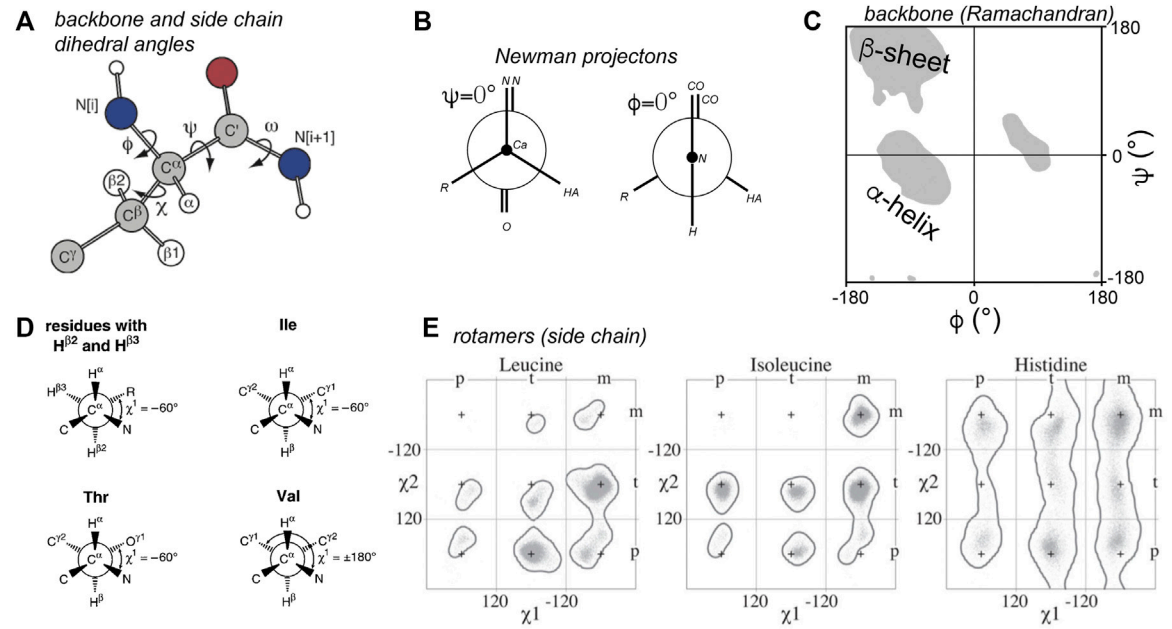


FIGURE 1 | Torsion angles in proteins. **(A)** Dihedral angles defining the protein backbone and side chain: ϕ, ψ, and ω, and χ. **(B)** Newman projections of the ϕ and ψ angles. The zero degree configuration of each is shown. **(C)** Ramachandran plot of ϕ and ψ angles. Grey areas indicate the classic secondary structure regions that are most populated. Note that the ϕ = 0° from panel B is not usually observed due to steric hindrance. **(D)** Definitions of the χ_1 side chain angle. **(E)** Rotamer states for χ_1 and χ_2 for three residue types. Panel A is adapted with permission from (Hong et al., 1997a), **(D)** is adapted with permission from (Markley et al., 1998), and panel E from (Hintze et al., 2016).

TORSION ANGLES IN PROTEINS

By these dihedral or torsion angles we refer to the angle of two neighboring chemical bonds with each other (**Figure 1**). In protein structural biology the torsion angles that define the backbones of proteins are most commonly discussed, and they have a standard nomenclature (IUPAC-IUB Commission on Biochemical Nomenclature., 1970; Markley et al., 1998). These angles are defined and illustrated in **Figures 1A,B**. Among these the ω angle is not very variable, with it typically adopting a value of ~180° (trans). The cis configuration is rare, but when it occurs it can have notable biological consequences, for instance Pro cis/trans isomerization is implicated in the aggregation process of the β2-microglobulin protein (Torbeev and Hilvert, 2013; Mukhopadhyay et al., 2018). Of more common interest are the ϕ and ψ backbone angles, which are visualized in Ramachandran plots (**Figure 1C**) (Ramachandran and Sasisekharan, 1968; Hovmöller et al., 2002). Only particular regions of Ramachandran space tend to be accessible, with the highly populated clusters representing secondary structure elements. The side chains of amino acids are also defined by named torsion angles, called χ_1, χ_2 etc., following an accepted nomenclature (**Figures 1D,E**) (IUPAC-IUB Commission on Biochemical Nomenclature., 1970; Markley et al., 1998; Lovell et al., 2000). These angles have preferred regions of geometric space, known as the preferred rotamer states that are nicely captured in rotamer libraries of evolving sophistication (**Figures 1E,F**) (Lovell et al., 2000; Dunbrack, 2002; Hintze et al., 2016).

SSNMR TORSION ANGLE MEASUREMENTS—NCCN AS AN EXAMPLE

Technical Implementation and Analysis

The general principle of torsion angle measurements by ssNMR is based on the measurement of the relative orientations of anisotropic ssNMR parameters, which in turn can be correlated to bond angles or orientations. The most straightforward and most common example involves the relative orientations of two different dipolar couplings, since these dipolar interactions are conveniently aligned with chemical bonds. For example, one can measure the relative orientation of the dipolar coupling vector (between ^{13}C and ^{15}N) of one C-N bond to the dipolar vector associated with a second C-N bond. When these two vectors are associated with the Cα-N and C′-N bonds in a polypeptide, the angle would be equivalent to the ψ angle (**Figure 1A**). This is the approach behind the "NCCN" torsion angle experiments introduced in the late 1990s (Feng et al., 1997a; Costa et al., 1997).

The original implementation of NCCN measurements (Rienstra et al., 2002a, 2002b; Ladizhansky et al., 2003; Jaroniec et al., 2004; Bajaj et al., 2009; Barnes et al., 2009; van der Wel et al., 2010; Hoop et al., 2016) is as follows: a double quantum (DQ) state is generated between the directly bonded Cα and C′ carbons, e.g., via SPC5 dipolar recoupling (**Figure 2A**; red boxes). This DQ state is then submitted to ^{15}N-^{13}C recoupling (commonly via the REDOR approach; **Figure 2A**, blue box). Since each carbon has a directly attached ^{15}N, they both

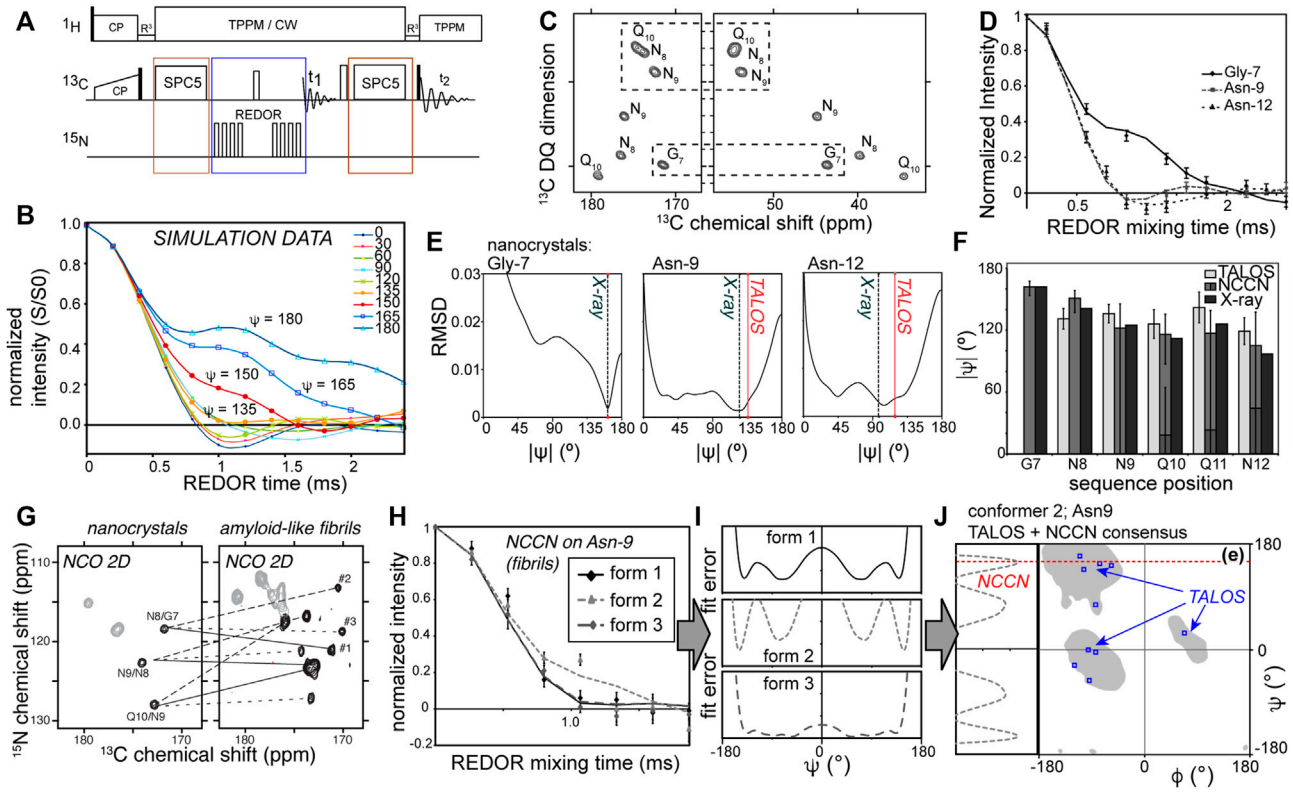


FIGURE 2 | Measuring backbone angle ψ via NCCN experiments. **(A)** Pulse sequence for a 2D NCCN experiment, in which 2D ^{13}C-^{13}C (SQ-DQ) spectra are acquired for varying lengths of ^{13}C-^{15}N dipolar mixing times. **(B)** Simulated NCCN dephasing curves, showing the S/S0 intensity ratios as a function of the REDOR mix time. **(C)** Example 2D SQ-DQ spectrum acquired as part of a NCCN measurement of amyloid-like peptide nanocrystals, along with **(D)** experimental NCCN curves for three residues. **(E)** RMSD fit results of the experimental data, showing the minima. The X-ray and TALOS results are marked as well. **(F)** Bar graph of the NCCN best-fit results compared to the TALOS chemical shift analysis and the known X-ray structure data. The NCCN data match the X-ray data, with a better accuracy than TALOS analysis alone. **(G)** 2D NCO spectra for nanocrystalline and fibrillar GNNQQNY peptide (uniformly ^{13}C,^{15}N-labeled in the first four residues, marked as residues G7 through Q10 of the parent Sup35p protein). Lines show how each residue gives one signal in the crystals, but three peaks in the fibrils, which complicates distance measurements. **(H)** A NCCN experiment results probing the backbone angle ψ for Asn-9 in the three co-existing fibril conformers. Two conformers have similar dephasing curves, but conformer #2 diverges. **(I)** RMSD curves between simulation and experiment for the full range of ψ angles, for each conformer. Note that the minima occur at multiple angles, and sometimes are broad, providing ambiguous constraints. **(J)** To resolve this one can combine chemical shift analysis (TALOS) with the NCCN constraints. In this example the 10 best TALOS "solutions" are not self-consistent, but by considering the NCCN angle, we pinpoint the optimal result in the β-sheet region near ψ = 150°. Panels C-J are reprinted with permission from ref. (van der Wel et al., 2010) copyright 2010 American Chemical Society.

experience rapid dephasing in a manner that is dominated by the directly-bonded ^{15}N. Whilst the orientation dependence of a single ^{15}N-^{13}C REDOR recoupling experiment is masked in a typical MAS ssNMR study, these two ^{15}N-^{13}C recoupling curves display an interference effect that results in behavior that is sensitive to the N-Cα-C′-N dihedral angle (**Figure 2B**).

Practically speaking, a typical REDOR approach is used to measure a relaxation-corrected dephasing curve (**Figure 2B**). A series of datasets with varying REDOR mixing times is acquired (keeping the DQ excitation time fixed). The relaxation correction involves the measurement (at each REDOR time) for each peak of interest of a "S" signal (*with* active ^{15}N REDOR pulses) and "S0" signal (without ^{15}N REDOR pulses). The S/S0 peak intensity ratio is then plotted, yielding angle-dependent curves as shown in **Figure 2B**. At each time point either a 1D or 2D spectrum is acquired, with the 2D implementation shown in **Figures 2A,C** (Ladizhansky et al., 2003; van der Wel et al., 2010). The main benefit of the 2D version is that it allows one to resolve many signals at once (Ladizhansky et al., 2003). Naturally, it comes at the expense of signal-to-noise (per unit time), which is superior in the 1D versions. It is worth noting that the signal-to-noise can be a challenge in these experiments. This stems from the fact that a clear distinction of the different dihedral angles only occurs in those time points where already extensive dephasing has occurred (i.e., much of the signal is lost; see **Figure 2B**). Moreover, in this region, the differences between some of the dihedral angles can be quite small, such that a very good signal to noise may be required to narrow down a specific angle. Achieving a good signal-to-noise level is inherently a challenge due to the polarization losses associated with the multiple recoupling steps (both the ^{13}C DQ recoupling and filtering, and also the heteronuclear recoupling), both due to losses inherent in the DQ filtering and relaxation processes.

Analysis and interpretation of the obtained data is done by comparing experimental data points (i.e., the S/S0 ratios) to reference curves such as the data in **Figure 2B**. The latter can

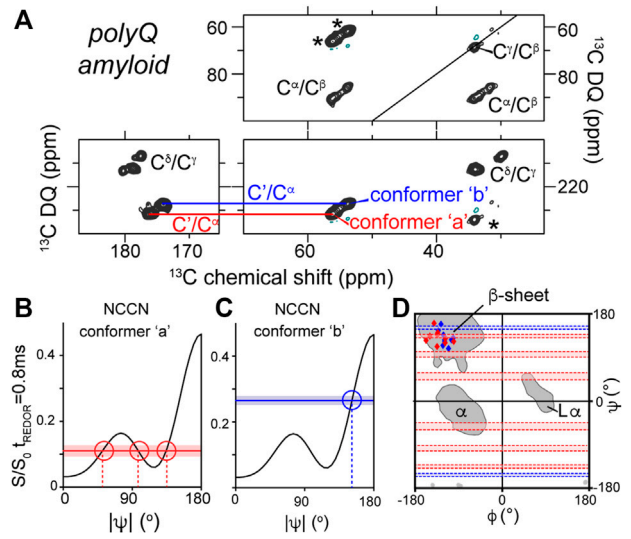


FIGURE 3 | NCCN analysis of polyQ amyloid structure. **(A)** 2D SQ-DQ spectrum from a NCCN measurement on polyQ amyloid fibrils, showing the main peaks of the two conformers called 'a' (red) and 'b' (blue) that make up the polyQ amyloid core. **(B,C)** Comparison of experimental S/S_0 values (horizontal lines) to simulated ψ-dependence S/S_0 ratios (black lines) for each conformer. **(D)** Ramachandran plot showing the integration of TALOS and NCCN analysis of the structures of polyQ amyloid conformers a and b. See also **Figure 4** below. Figure adapted with permission from ref. (Hoop et al., 2016).

be derived from analytical descriptions of the experiment. In our own analyses we typically employ primarily numerical simulations to generate the curves. A relevant spin system is modelled for the range of candidate dihedral angles (or molecular structures), resulting in a library of reference curves. For some types of torsion angle experiments it may also be necessary to incorporate in this "library" relaxation effects, the impact of the chemical shift anisotropy (CSA), and dynamic modulation (van der Wel et al., 2009; Hou et al., 2010; Hoop et al., 2014). Naturally, these additional free parameters increase the complexity of the analysis and reduce the expected precision of obtained results.

Additionally, it is worth noting that a dipolar recoupling experiment like REDOR reflects a through-space interaction and can therefore invoke multispin interactions in extensively or fully isotope-labeled samples. Such multispin interactions would not invalidate the dihedral angle experiment, as it is essentially based on seeing a dependence of the observed signal on the dihedral angle in question. This dependency would be modulated, but not be negated by multispin interactions, unless non-local interactions would dominate over the local interactions which are supposed to be recoupled. However, in this case the effect of "dipolar truncation", in which strong dipolar interactions suppress or truncate the contributions from longer range interactions (Bayro et al., 2009), is actually beneficial. Directly-bonded ^{13}C-^{15}N interactions will effectively truncate the effect of long-range inter-residue and inter-molecular interactions. It is also worth noting that we benefit here from the relatively low density of ^{15}N sites in polypeptides, which allows for the trains of ^{15}N REDOR pulses without detrimental ^{15}N-^{15}N recoupling (illustrated in **Figure 2A**). Another important observation is that these experiments can be applied to polypeptides outfitted with uniform ^{13}C and ^{15}N labeling, without need for site- or residue-specific isotope labeling as may be needed for certain other types of torsion angle experiments. The examples in **Figures 2**, **3** are obtained with synthetic peptides where short stretches of residues were labeled, but the same experiments can be applied to fully labeled proteins (Ladizhansky et al., 2003).

Example Application to Nanocrystalline Peptides With Known Structures

The data in **Figures 2C–F** represent validation experiments in an amyloid-like, but nanocrystalline, peptide assembly. Integrated peak volumes from the SQ-DQ 2D spectrum (**Figure 2C**) are plotted as S/S0 ratios as a function of the REDOR recoupling time. The example curves in **Figure 2D** show the variable differences between residues, with some of the angles hard to distinguish (see also **Figure 2B** for $\psi < 130^{\circ}$). **Figure 2E** shows the results of fitting the experimental curves, yielding in some cases multiple minima. This illustrates one common challenge of torsion angle experiments, which is that they often have regions of angular space where the obtained dephasing curves are essentially indistinguishable. As reviewed in earlier work (Hong, 2006; Hong and Wi, 2006; Ladizhansky, 2009), this issue can be addressed by combining multiple types of dihedral angle measurements with complementary sensitivities. Notably, the NCCN experiment minima are close to both the X-ray structure angles and the results of NCCN-based chemical shift analysis. This is illustrated in the bar graphs of **Figure 2F**. These results show that the NCCN experiment gives results that match the structure as known from X-ray microcrystallography (Nelson et al., 2005; van der Wel et al., 2007, 2010), consistent with earlier studies (Ladizhansky et al., 2003).

Studies of Unknown Structures in Amyloid Fibrils

Several studies have deployed NCCN experiments to study unknown structures of amyloid-like fibrils (Jaroniec et al., 2004; van der Wel et al., 2010; Hoop et al., 2016). For the fibrils formed by the abovementioned prion-derived model peptide GNNQQNY, these experiments detected the anticipated β-sheet structures typical of amyloids, but also non-β structure as an integral part of a composite fibril structure (van der Wel et al., 2010; Lewandowski et al., 2011). Notably this feature was present in one of three conformers (monomer structures) that composed the "composite" fiber architecture (which manifest as three peaks per residue; **Figure 2G**). This feature made the reliance on distance measurements more difficult, due to substantial peak overlap. Thus, the complementary torsion angle measurements were particularly helpful and valuable. The three conformers' distinct structures are clear in the torsion angle data, as illustrated for residue Asn-9 in **Figure 2H**. The presence of a

non-β kink or turn was at first surprising, as it is absent from the nanocrystals. However, nowadays this finding is reminiscent of the common presence of kinks, turns or bends in amyloid architectures (between quite short β-strand segments) (van der Wel, 2017; Sawaya et al., 2021). **Figures 2I,J** shows how the experimental ssNMR data were translated to dihedral angles. The fit between the experimental data and simulated curves was evaluated as a function of the simulated dihedral angle (**Figure 2I**). In some cases this shows consistency (i.e., low RMSD between the experimental data and simulated data) with multiple possible angles. One approach to overcome this ambiguity, shown in **Figure 2J**, is to combine the optimal NCCN angles with TALOS-based chemical shift analysis, in order to obtain a unique solution not accessible through either approach alone.

A similar structural complexity was subsequently detected in ssNMR studies of polyglutamine (polyQ) amyloid structure, which was found to consistently contain two distinct types of peptide conformations (**Figure 3A**) (Schneider et al., 2011; Sivanandam et al., 2011; Hoop et al., 2014, Hoop et al., 2016; Isas et al., 2015; Matlahov and van der Wel, 2019). Proteins with expanded polyQ domains are of biomedical interest as they represent the molecular basis of a series of CAG repeat expansion disorders, which remain as-yet incurable and untreatable (Wetzel, 2012). The disease-associated mutant proteins are prone to aggregation and form toxic aggregates, which include fibrillar structures with all the hallmarks of typical amyloid fibrils. Unlike several other amyloid proteins (van der Wel, 2017; Sawaya et al., 2021), the 3D atomic structure of none of the pathogenic polyQ protein aggregates is as yet known (Matlahov and van der Wel, 2019). The degenerate nature of the primary sequence of these polypeptides greatly complicates structural analysis by ssNMR, which has made the deployment of torsion angle measurements valuable and indeed essential. Structural study by distance measurements is further complicated by the already noted composite nature of the amyloid core structure, presenting as two peaks for each residue (marked "a" and "b" in **Figure 3**). By combining NCCN measurements and TALOS-based chemical shift analysis, it was shown that these two conformers reflect two types of β-strand structures with distinct backbone conformations (**Figures 3A–D**) (Hoop et al., 2016). The origin of this characteristic doubled-peaks signature stems from the presence of antiparallel β-sheets in these fibrils, which contain β-hairpin motifs for the longer polyQ expansion lengths (more about polyQ structures below) (Matlahov and van der Wel, 2019).

BEYOND THE PROTEIN BACKBONE—HCCH EXPERIMENTS

The first examples of direct MAS ssNMR torsion angle measurements were the HCCH experiments (Feng et al., 1996; Feng et al., 1997b), whose principles were also discussed in some detail in prior reviews (Hong and Wi, 2006; Ladizhansky, 2009). Briefly, again a DQ state is generated between two directly bonded ^{13}C carbons, but now it is combined with controlled recoupling of the C-H dipolar interaction (**Figure 4A**). Keeping the DQ excitation time fixed, the C-H recoupling time is varied. This yields a time-dependent decrease in the DQ signal, with the shape of the curve dependent on the geometry (i.e., torsion angle) of the HC-CH spin system (**Figure 4B**) (Feng et al., 1997b; Bajaj et al., 2009; Edwards et al., 2010; Hoop et al., 2016). Whilst initial applications focused on organic molecules and retinal structures, recent applications have used this technique to probe the side chain dihedral angles of amino acids (Rienstra et al., 2002b; Bajaj et al., 2009; Edwards et al., 2010; Hoop et al., 2016). Like the NCCN experiment, these HCCH measurements can be applied to uniformly ^{13}C labeled residues and polypeptides. In contrast to backbone torsion angles, the side chain angles are (at this time) not accessible via the kind of chemical shift analysis applied to polypeptide backbones (e.g., TALOS; **Figure 2**).

Polyglutamine Amyloid Steric Zippers

One recent application of the HCCH experiment was also in our own work on aggregated polyglutamine (polyQ) proteins. As noted above, these protein aggregates are hard to study by distance measurements alone, requiring the application of dihedral angle measurements for structure determination. The glutamine residues feature an extended aliphatic side chain, with two methylene (CH_2) groups for the Cβ and Cγ atoms. This permitted the use of HC-CH dihedral angle measurements of the χ_1 and χ_2 angles of these amino acids. **Figure 4B** shows how the χ_2 angles of 60 and 180° give clearly different HCCH dephasing curves, with the dephasing at the mid-point (1/2 rotor period) varying with the χ_2 angle (**Figure 4C**). For a practical experiment, measuring up to this mid-point is sufficient, as shown in **Figures 4D,E**. These results identified the χ_2 angle for the glutamine residues in polyQ amyloid to be near 180°, in contradiction to certain prior structural models based on (low-resolution) X-ray diffraction (Sharma et al., 2005; Hoop et al., 2016). Similarly, the same experiments can be used to probe the χ_1 angle (**Figure 4F**), adding further structural constraints on the polyQ amyloid core structure (**Figure 4G**). As summarized in **Figure 4H**, one unfortunate aspect of the χ_1 measurements is that they were able to *exclude* various conformations, but did not result in a completely unambiguous single angle value. As noted above, this is not uncommon for dihedral angle measurements, which can provide a unique solution in some cases ($\chi_2 \sim 180°$) but only partially constrain the angle in other cases. The model in **Figure 4G** represents a visualization of the best-fit backbone and side chain angle results. This model shows how the two thus-obtained conformers (β-strand types a and b) are conformationally compatible, and explain the co-assembly of the polyQ amyloid core. Notably, prior structural models derived from other techniques were not consistent with the obtained ssNMR results (Hoop et al., 2016).

Receptor-Substrate Interactions

Another notable use of the HCCH experiments is a nice study of a small molecule substrate (glutamate) bound to a receptor protein, which predated our work on polyQ (Edwards et al., 2010). An isotopically labeled substrate was bound to the ionotropic glutamate receptor 2, which was itself unlabeled (**Figure 5**).

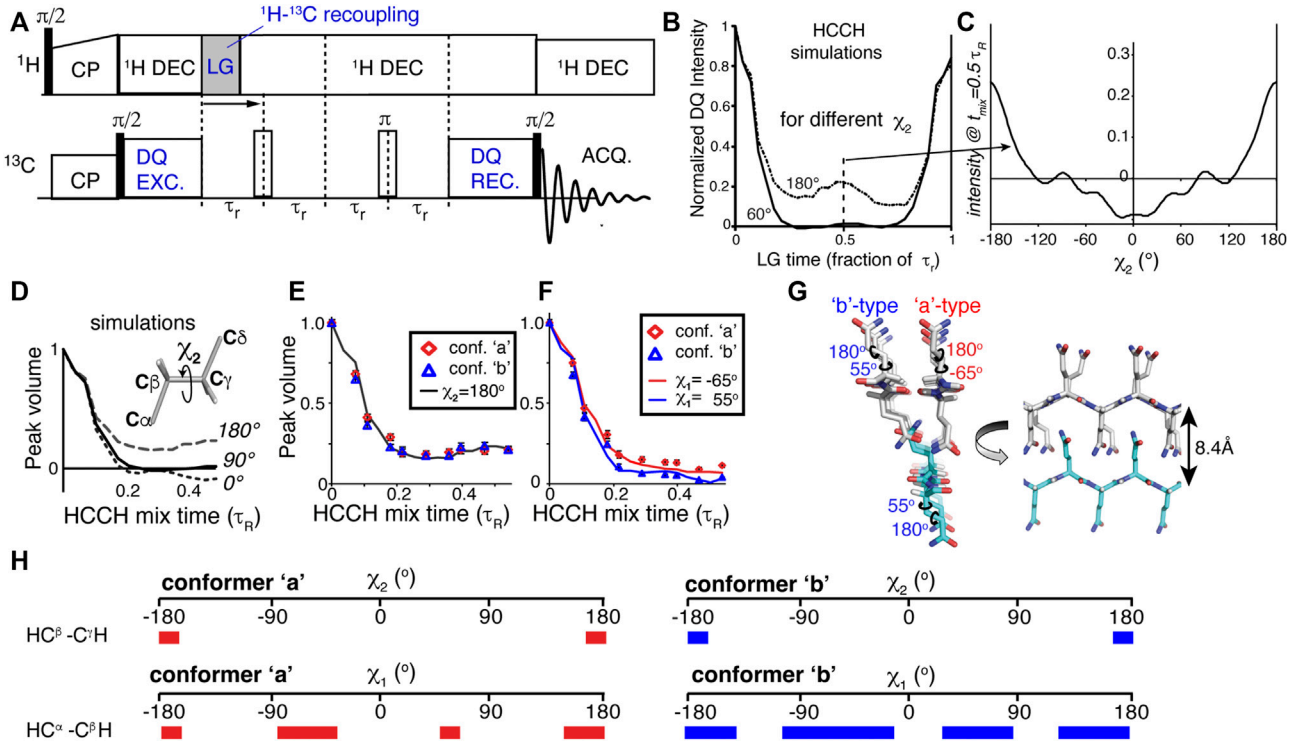


FIGURE 4 | HCCH experiments for measuring side chain angles. **(A)** Pulse program schematic for a HCCH experiment. **(B)** Simulated HCCH curves for measurement of the χ_2 angle in glutamine residues (a $C\beta H_2$-$C\gamma H_2$ spin system), for χ_2 = 60 or 180°. **(C)** Variation of the mid-rotor-period intensity as a function of the χ_2 angle. **(D,E)** Simulated and experimental data measuring the χ_2 angle of glutamine residues in the core of polyglutamine amyloid. **(F)** Application to the χ_1 angle in polyglutamine amyloid. **(G)** Visualized structural model for how the polyglutamine amyloid core is structured, based on the best-fit dihedral angle solutions. **(H)** Overview of the ambiguity in some, but not all, of the measured χ angles. The colored bars indicate those angles that are consistent with the observed torsion angle curves. The red bars on the left apply to conformer 'a', and the blue ones (right) to conformer 'b'. Figure adapted from ref. (Hoop et al., 2016) with permission.

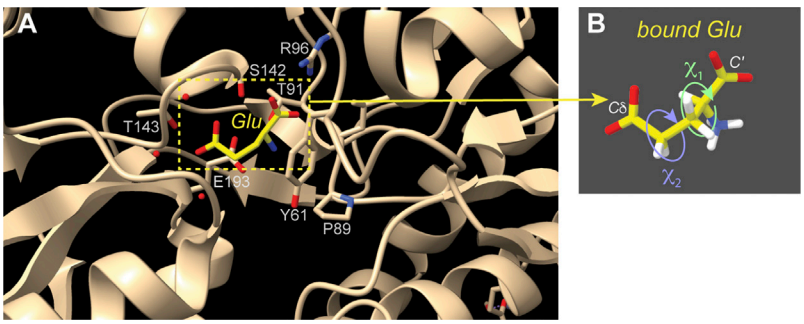


FIGURE 5 | Receptor-bound substrate studied by torsion angle ssNMR measurements. **(A)** This image shows the structure of the receptor-bound Glu amino acid, based on X-ray crystallography (PDB 1FTJ) (Armstrong and Gouaux, 2000). Selected residues surrounding the bound substrate are indicated. **(B)** HCCH ssNMR analysis of χ_1 and χ_2 side chain angles enabled the measurement of the Glu residue structure while it was bound to the crystalline unlabeled protein (Edwards et al., 2010). This figure was prepared with UCSF ChimeraX (Pettersen et al., 2021).

Two torsion angles, defining χ_1 and χ_2 of the Glu (**Figure 5B**), were measured with HCCH-type experiments. When combined with REDOR-based distance constraints, these data defined the conformation of the receptor-bound substrate. The reliability of the method was validated by comparison to known crystal structures (**Figure 5A**), with the proof-of-principle ssNMR study being applied to the crystalline receptor-substrate complex, in which the glutamate substrate is uniformly labeled. The HCCH measurements represented analogous experiments to our own subsequent studies of the polyQ amyloid structure (**Figure 4**). The individual dihedral angle measurements were again consistent with multiple distinct

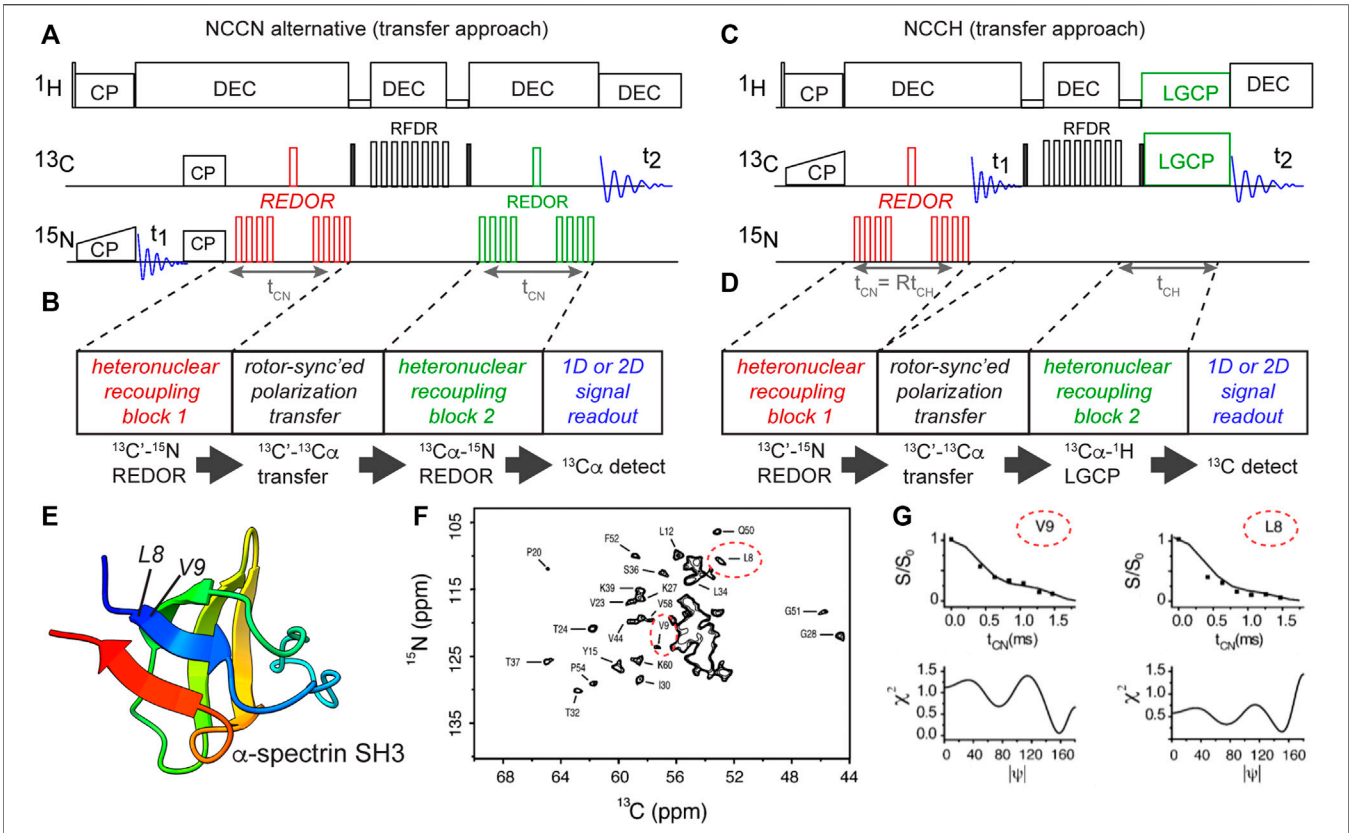


FIGURE 6 | Transfer-based torsion angle measurements of backbone angle ψ. **(A,B)** NCCN pulse sequence implemented without ^{13}C DQ generation, using rotor-synchronized ^{13}C-^{13}C transfer instead (Ladizhansky et al., 2003). Carbonyl (C′) polarization is created via a double CP scheme, followed by REDOR-based ^{13}C-^{15}N recoupling (time t_{CN}). The residual ^{13}C′ signal is transferred to neighboring ^{13}Cα via a short RFDR block, after which the ^{13}Cα-^{15}N recoupling is done (also for t_{CN}). Finally, the remaining ^{13}C signal (as a function of t_{CN}) is recorded in a series of 1D or 2D N(CO)CA spectra (see panel F). **(C,D)** NCCH-type measurements of the ϕ backbone angle based on the same design principle (Ladizhansky et al., 2002). Here, ^{13}C-^{1}H dipolar recoupling occurred during the LGCP period. Note that here the REDOR time (t_{CN}) and LGCP time (t_{CH}) are both incremented in parallel. Panels B and D illustrate the generic design of these experiments, with color-coding of the dipolar recoupling blocks. **(E)** Structure of α-spectrin SH3, with residues L8 and V9 in the first β-strand marked (Musacchio et al., 1992), prepared with UCSF ChimeraX (Pettersen et al., 2021). **(F)** 2D N(CO)CA spectrum for uniformly labeled α-spectrin SH3, showing peaks for ^{13}Cα$_i$-^{15}N$_{i+1}$ correlations between residues i and i+1. Peaks for the Cα of residues L8 and V9 are indicated with red ovals. **(G)** Example N(CO)CA NCCN data curves for L8 and V9, along with matching values of the ψ angle. Panels **(F-G)** are adapted from (Ladizhansky et al., 2003) with permission, copyright American Chemical Society 2003.

solutions, but by combining the different data with a REDOR-based distance measurement, a unique structural solution was obtained. The obtained conformation matched X-ray based structures previously determined, as illustrated in **Figure 5** (Armstrong and Gouaux, 2000).

BACK TO THE BACKBONE–ALTERNATIVE IMPLEMENTATIONS

In our hands, the abovementioned DQ-based approach is particularly powerful and robust. However, also alternative implementations of analogous torsion angle measurements have been demonstrated. This is illustrated via a different implementation of the NCCN-type measurement (**Figures 6A,B**) (Ladizhansky et al., 2003). This experiment follows the general approach shown in panel B: having two distinct recoupling blocks sandwiching a polarization transfer block. A key aspect of the pulse sequence is that each of these blocks should ideally be deployed in a rotor-synchronized (and/or constant-time) manner, such that individual molecules (crystallites) are recoupled at the same orientation relative to the magnetic field. This method has the benefit that it can be quite flexibly deployed, for example in the form of the HCCN experiment shown in **Figures 6C,D** (Ladizhansky et al., 2002). This HCCN experiment can be used to determine the ψ backbone angle, especially in α-helical secondary structures, thus complementing the NCCN experiment in terms of its applicability.

An illustration of the application of the abovementioned NCCN experiment from **Figure 6A** is shown in **Figures 6E,F**, from a study on crystalline uniformly ^{13}C,^{15}N-labeled α-spectrin protein (**Figure 6E**) (Ladizhansky et al., 2003). Integrating peaks in a series of 2D N(CO)CA spectra (**Figure 6F**), the REDOR dephasing curves for individual residues were measured. The β-sheet residues 8 and 9 are marked in the spectrum, with their

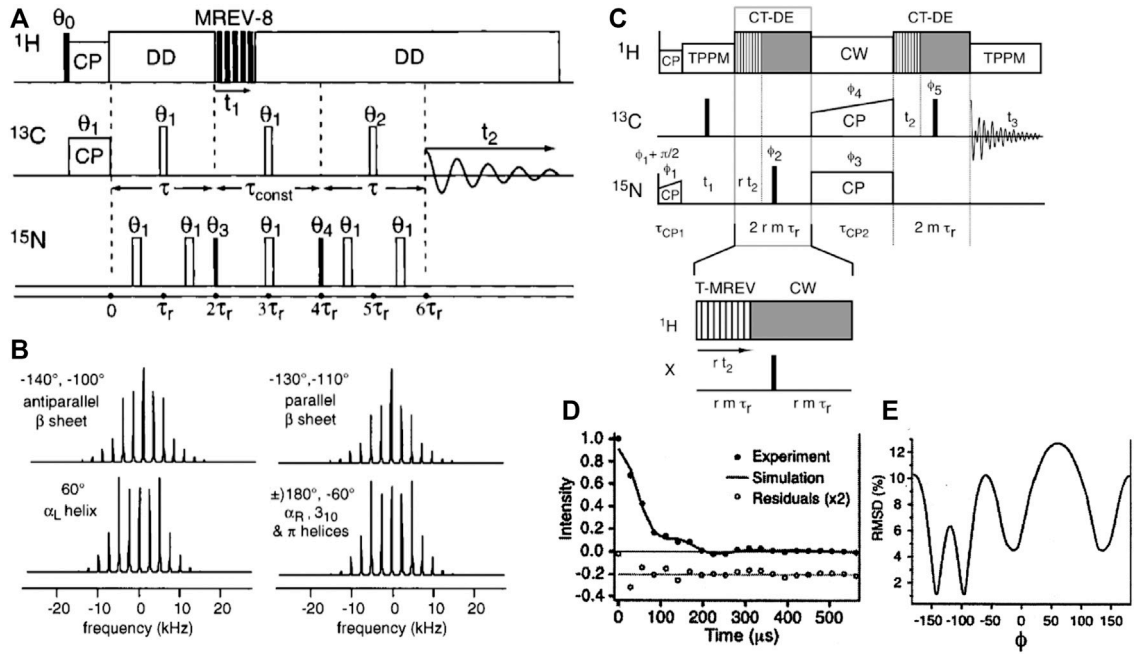


FIGURE 7 | HCNH measurement of the ϕ backbone angle. **(A)** Pulse sequence for the HCNH experiment, where MREV-8 is used to achieve ^{1}H-^{13}C/^{15}N dipolar recoupling of ^{13}C-^{15}N correlations. **(B)** Example dipolar lineshapes for four different ϕ angles, based on simulations. Note that one lineshape often represents two distinct ϕ angles, as indicated. The figure also lists the matching secondary structures, showing that each type of secondary structure has a characteristic spectral lineshape. **(C)** Pulse sequence of a HCNH experiment based on a transfer-like implementation similar to those shown in **Figure 6**. **(D,E)** Representative results showing how also this variant points to multiple minima. Adapted with permission from **(A,B)** ref. (Hong et al., 1997a) and **(C–E)** (Rienstra et al., 2002a), Copyright 1997 and 2002 American Chemical Society.

NCCN curves shown in **Figure 6G**. Once again, multiple minima can be observed, e.g. for residue L8.

To measure the ϕ backbone angle one can use the HNCH ssNMR experiment (Hong et al., 1997a, Hong et al., 1997b; Hong, 1999; Rienstra et al., 2002a). Also this experiment can be implemented in different ways (**Figure 7**), either via coherence generation or a rotor-synchronized transfer approach (Rienstra et al., 2002a). The corresponding pulse sequences are illustrated in **Figure 7**, along with selected results from the literature. Once more these data illustrate how a single torsion angle measurement commonly is consistent with multiple best-fit minima. As noted above, this ambiguity can be resolved by performing multiple different torsion angle measurements, integration with chemical-shift-based analysis, consideration of accessible Ramachandran/rotamer space, and the measurement of relevant distance measurements. The reader is referred to prior review articles for a more in-depth discussion of these techniques (Hong and Wi, 2006; Ladizhansky, 2009).

USING CSA TENSORS FOR MEASURING ω AND MORE

The most intuitive types of dihedral angle measurements are arguably those in which one recouples dipolar interactions that align nicely with chemical bonds, as discussed above. However, also the anisotropy of the chemical shift (CSA) can be deployed to good effect, either by recoupling two CSAs to each other, or by combining CSA recoupling with dipolar recoupling. Examples of these approaches were also introduced in the late 1990s (Ishii et al., 1996; Weliky and Tycko, 1996; Bower et al., 1999). The experiment in its simplest form involved simply slow MAS along with ^{13}C-^{13}C mixing, but this depended on ^{13}CO-only labeling (Weliky and Tycko, 1996). It would not work in this form for a uniformly ^{13}C labeled sample. An interesting recent application of this experiment was used to examine the conformation of the amyloid-binding fluorescent dye congo red, in its fibril-bound state (**Figure 8**). The CSA-based torsion angle measurement determined the central bond in the amyloid-bound dye, complementing other structural studies on the dye-fibril interactions by the same research groups (Schütz et al., 2011; Gowda et al., 2017). In this case, site-selective ^{13}C labeling of the congo red was used to label two aromatic carbons. These aromatic sites have large ^{13}C CSAs, with a well-understood orientation of the CSA tensor relative to the molecular frame. Thus, by determining the relative orientations of the two CSA tensors, one can determine the rotational angle marked in **Figure 8A**. Free rotation around this bond causes the unbound dye to display low fluorescence. Upon binding to amyloid fiber surfaces, the immobilization of the dye boost the fluorescence and permits thereby the selective detection of amyloid-like structures (Schütz et al., 2011; Yakupova et al., 2019). The

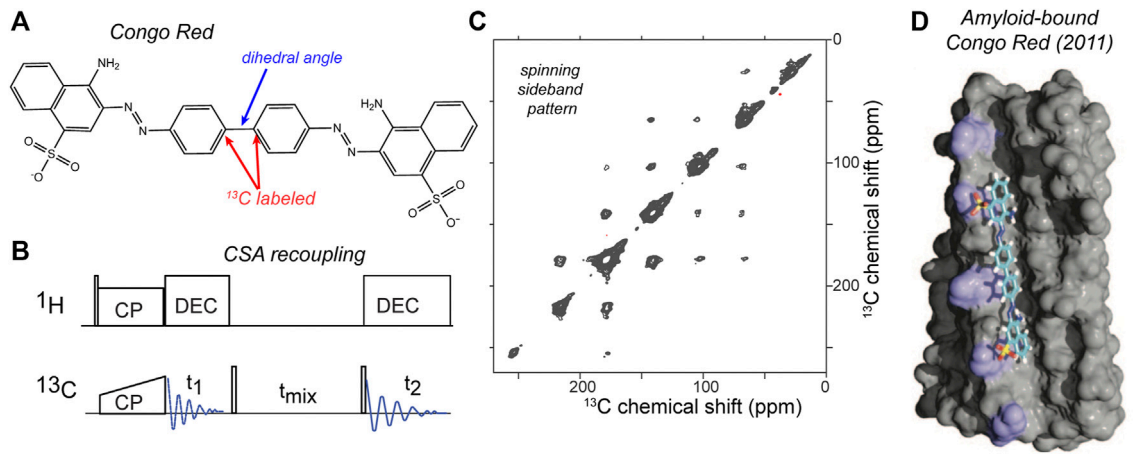


FIGURE 8 | CSA-recoupling with selective labeling: amyloid dye Congo Red. **(A)** Chemical structure of congo red, which is widely used to detect amyloid structures. For the CSA recoupling experiments the 1 and 1′ carbons are ^{13}C labelled (marked red), in order to measure the marked dihedral angle (blue). **(B)** Pulse sequence used for the CSA recoupling experiment used by (Gowda et al., 2017) to measure the dihedral angle marked in **(A)**. **(C)** 2D ^{13}C-^{13}C tensor correlation spectrum at 8 kHz MAS on a 850-MHz (^{1}H) spectrometer, using the indicated pulse sequence with a 50-ms rotor-synchronized ^{13}C-^{13}C transfer time. The sample contained HET-s amyloid fibrils with the selectively labeled congo red bound. For more details see ref. (Gowda et al., 2017). **(D)** Structural model of amyloid-bound congo red from prior ssNMR studies by the same groups, adapted with permission from ref. (Schütz et al., 2011).

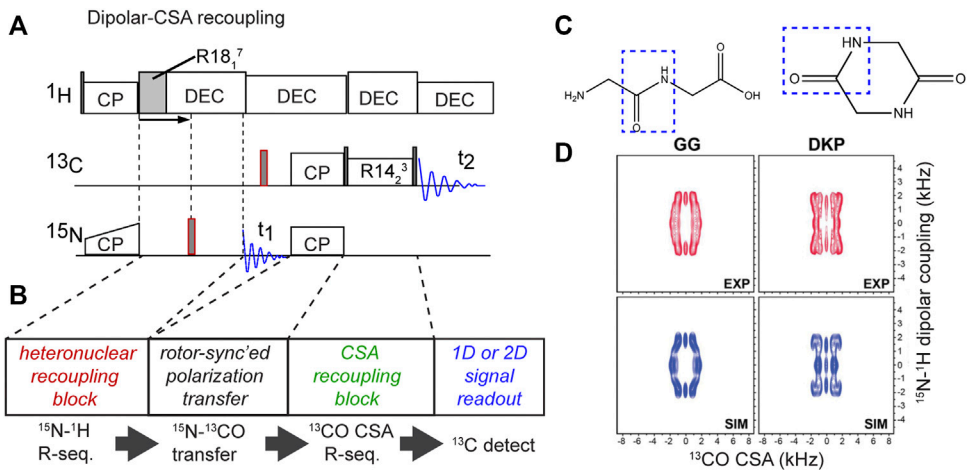


FIGURE 9 | Hybrid CSA/dipolar recoupling experiment. **(A)** Pulse sequence and **(B)** its building blocks. **(C)** Chemical structures of test molecules with different molecular configurations mimicking the cis or trans peptide bond angle ω. **(D)** 2D CSA-dipolar correlation spectra for the two reference compounds representing the distinct ω angles. Panel (D) was adapted with permission from ref. (Mukhopadhyay et al., 2018).

experiment was performed with a fairly straightforward pulse sequence (**Figure 8B**), that allowed the measurement of a 2D ^{13}C-^{13}C spectrum that at slow MAS rates (here 8 kHz) showed a spinning side band pattern as illustrated in **Figure 8C**. Analysis of the peak pattern allowed the determination of the abovementioned torsion angle, which could then be used to refine the structural model of the congo red bound to the fiber surface grooves (**Figure 8D**). This represents a nice example of how one can target a torsion angle measurement to a particular (biological) question.

Moreover, hybrid methods can also determine the relative orientations of a CSA tensor and a dipolar coupling (Ishii et al., 1996; Fujiwara et al., 1997; Hong et al., 1998; Chan and Tycko, 2003; Hou et al., 2010; Mukhopadhyay et al., 2018). In these MAS ssNMR experiments, different kinds of pulse sequence elements are used to actively recouple the CSA under medium/fast MAS (>10 kHz), unlike the slow-MAS (<10 kHz) CSA measurement mentioned above. For example, recent studies have used either ROCSA and R-sequence-based CSA recoupling techniques (Chan and Tycko, 2003; Hou et al., 2010; Mukhopadhyay et al., 2018). This approach also makes it feasible to apply these experiments without selective labeling strategies, enabling studies of fully ^{13}C-labeled proteins. **Figure 9** shows a recent example in which R-sequences were used for both CSA

recoupling and dipolar recoupling, in order to enable measurement of the peptide bond angle ω (Mukhopadhyay et al., 2018). The figure shows the implementation and results for two model compounds (**Figure 9C**), but the original paper includes application to crystalline and fibrillar protein samples as well. This experiment and figure also illustrate a few relevant concepts, common to dihedral angle measurements. **Figure 9D** shows a 2D spectrum in which none of the axes show the (isotropic) chemical shift, but rather the anisotropic parameters: the ^{13}C CSA and the ^{1}H-^{15}N dipolar coupling strength. Prior work has proposed the term Relayed Anisotropy Correlation spectra (RACO) (Ishii et al., 1996). In this experiment, these CSA and dipolar recoupling periods are both independently incremented, resulting in these typical 2D data. This figure illustrates how both of the (recoupling) time domain periods can be processed (Fourier transformed) in order to obtain a dipolar or CSA-based 2D lineshape. Although information-rich, this 2D approach does take considerable time. In line with prior work, this paper (Mukhopadhyay et al., 2018) discussed and demonstrated an "accordion" approach in which multiple distinct recoupling periods are incremented in synchrony. This yields a faster "1D" experiment that is must more time efficient than the 2D RACO spectra on uniformly labeled test compounds in **Figure 9D**. For more details, and the application to uniformly ^{13}C,^{15}N-labeled proteins, the reader is referred to the original work (Mukhopadhyay et al., 2018).

LONGER-RANGE VECTOR ANGLE MEASUREMENTS

Most of the above examples are based on recoupling anisotropic interactions of neighboring (directly-bonded) atoms. However, useful angular constraints can also be derived from non-local interactions, involving atoms that are not directly bonded. Some CSA-CSA recoupling experiments fall into this category, but there are also dipolar based variants. This includes the recoupling of N-H dipolar vectors for ^{15}N nuclei in neighboring amino acids, in HNNH-type experiments (Reif et al., 2000; Rienstra et al., 2002a, Rienstra et al., 2002b; Franks et al., 2006, Franks et al., 2008). The utility of such vector angle (VEAN) constraints was nicely demonstrated in the high-resolution structure determination of the model protein GB1 (Franks et al., 2006, Franks et al., 2008; Wylie et al., 2011). The obtained N-H, N-H VEAN depends on several standard protein dihedral angles, as well as bond lengths and bond angles. To use these constraints, the authors directly incorporated the VEAN angle into the structure calculation routine, rather than attempt to back-calculate the individual $\phi/\psi/\chi$ angles. It is worth noting that these "remote" angle constraints can also take other forms, outside the HNNH variant, with any orientational constraint being potential valuable for structure determination. For example, the relative orientation (or projection angle θ) between $(N\text{-}H)_{i+1}$ and $(C\alpha\text{-}H)_i$ have been used to constrain amyloid structure of a fragment of the transthyretin protein (Jaroniec et al., 2004).

OTHER SSNMR PROBES OF DIHEDRAL ANGLES

Distance-Based Constraints

Although not a focus of the current review, it is worth noting that many ssNMR distance measurements directly or indirectly constrain dihedral angles. In some cases, specific ssNMR experiments were designed with the explicit goal to determine dihedral angles via precise measurements of specific internuclear distances (Sinha and Hong, 2003; Wi and Spano, 2011; Hu et al., 2012; Pope et al., 2018). This includes for instance the so-called BARE (Backbone Recoupling) experiments that measure the distances between backbone nitrogens and carbonyls, with implications for the intervening backbone torsion angles (Hu et al., 2012).

Isotropic Chemical Shifts as Torsion Angle Constraints

From the above it is clear that there is an impressive library of ssNMR torsion angle measurements, many of which were developed and demonstrated in the late 1990s and early 2000s. Nonentheless, since then many ssNMR protein structures were determined without use of these types of constraints. Instead, most structural ssNMR studies currently deploy isotropic chemical shifts to estimate the residue-specific backbone torsion angles. This is facilitated by the observation that backbone chemical shifts (along with the Cβ shift) are reliable indicators of local secondary structure, and can even be used for quantitative determinations of backbone dihedral angles (Cornilescu et al., 1999; Shen and Bax, 2007; Shen et al., 2009). Although developed for (and from) solution NMR protein structures, these algorithms have been shown to be similarly effective for ssNMR chemical shifts. These results have been sufficiently reliable that several papers deploying explicit torsion angle measurements note the high degree of consistency between the two approaches (Jaroniec et al., 2004; van der Wel et al., 2010). That said, both methods have their strengths and weaknesses, and can used as complementary tools (van der Wel et al., 2010; Hoop et al., 2016). For example, chemical shift-based estimates can be used to resolve ambiguities inherent in direct dihedral angle measurements (and *vice versa*).

PROSPECTS FOR TORSION ANGLE APPLICATIONS

These recent studies give a chance to consider the question of when and why to deploy torsion angle measurements. These experiments are in principle powerful, and they been used in recent years, but clearly not as widely as other structural ssNMR measurements. There are several reasons of this. One reason is that protein chemical shifts themselves give a lot of insight into the backbone torsion angles (i.e. secondary structure), with programs like TALOS allowing for a semi-quantitative determination of the backbone torsion angles

(Shen et al., 2009; van der Wel et al., 2010). Most recent ssNMR-based structures are based on combinations of distance measurements along with such chemical-shift-based backbone angles. Although (typically) neither the distances nor the angles are extremely precise, their combination in sufficient numbers can yield good atomic structures of proteins (Loquet et al., 2008). Dihedral angle measurements may be used to improve the resolution of a structure derived from a combination of distance constraints and chemical shift information. This has been demonstrated for example in work by the Rienstra group on crystals of the model protein GB1 (Franks et al., 2008; Wylie et al., 2011). This enables higher resolution structures than otherwise accessible, but the question may arise whether the improvement in the structure quality warrants the amount of work (NMR time, but also data analysis). One possible approach would be to deploy these experiments in a targeted fashion: to enhance our understanding of especially important parts of protein structures, such as enzyme active sites, ion channel selectivity pores and similar (van der Wel et al., 2009; Caulkins et al., 2014; Wylie et al., 2014). Still, the implementation, execution, and interpretation of torsion angle measurements can be challenging, more so due to more limited prior experience with these methods in the overall ssNMR community (compared to canonical distance measurements). Thus, it may seem unclear when one would deploy dihedral angle measurements. We will examine some considerations or conditions that favor their use.

Extensive Intermolecular Interactions

Distance-based ssNMR structure determination borrows heavily from methods perfected in solution NMR structural biology. In dissolved or crystalline globular proteins, or in membrane-associated proteins, one can assume that most NMR-detected distance constraints reflect interactions within an individual protein (Loquet et al., 2008). However, this is not always a safe assumption, for example when studying amyloid fibrils in which the predominant residue-residue interactions may be inter- rather than intra-molecular. SSNMR studies of amyloid structures have addressed this by diluting labeled monomers in an excess of unlabeled monomers, prior to the assembly process. This approach allows for the suppression of intermolecular interactions among (labeled) atoms, thus revealing intramolecular interactions that define the monomer structure within the fibril. Unfortunately, it also implies a drastic loss of sensitivity, as the sample is now only partly filled with "visible" labeled monomers. Torsion angle measurements are designed to probe the very local environment of the (bond) angle of interest. This means that they deploy relatively modest dipolar recoupling times and are largely insensitive to the presence or absence of intermolecular interactions, and that they can always be used to probe the monomer structure even in fully labeled samples. Aside from the already mentioned signal-to-noise benefits, this may also be important for (biological) assemblies that are difficult or impossible to (re)assemble *in vitro* from monomers, and thus are either fully labeled or fully unlabeled.

Assemblies With Inherent or Internal Polymorphism

Another challenge faced in our recent studies of amyloid fibrils is that some fibrils contain the same monomer in two or three distinct configurations or conformations, as part of a complex or composite fiber architecture (van der Wel et al., 2010; Lewandowski et al., 2011; Hoop et al., 2016). This means that any single atom (or residue) yields multiple signals, which further complicates distance-based structural measurements (E.g., **Figures 3**, **4** show examples). Extensive signal overlap results that cannot be resolved by site-specific isotope labeling or spin-system-based spectroscopic editing. Moreover, the effective signal to noise is decreased, as you are effectively studying a system that behaves as if it is twice or thrice as large (in terms of signal to noise). Yet, unlike a larger protein sequence, this challenge cannot be resolved by residue-specific labeling, truncation of the sequence, mutations, or segmental labeling. In such a case, distance measurements become less powerful, and torsion angle measurements can become an essential tool.

Cases Where Chemical Shift Analysis Falls Short

It was noted above that backbone dihedral angles can be probed via the chemical shift assignments of proteins, with the results sometimes being not much "worse" than much more time-consuming torsion angle measurements (**Figure 2F**). However, even if the detection of extended β-strands and α-helices is quite reliable, some non-standard motifs can be more challenging. In such cases chemical shift-based analysis alone may fail to resolve a reliable backbone conformation for a combination of observed shift values (see e.g., **Figure 2J**). Thus, it may be helpful to deploy targeted torsion angle measurements in such cases (Franks et al., 2008; van der Wel et al., 2010). Moreover, chemical shift analysis is limited to protein backbones, while ssNMR torsion angle measurements can be applied to side chains and non-protein biomolecules.

Repetitive Sequences and Structures

Our work on polyQ amyloid structure illustrates one interesting class of proteins where distance constraints fail to provide a complete answer. Highly repetitive sequences such as the polyQ proteins render distance constraints more difficult to obtain, or at least interpret. Biology features many repetitive protein sequences. Protein aggregation diseases often feature proteins with low complexity sequences, which are sometimes defined as prion-like sequence elements (Franzmann and Alberti, 2019). Besides several distinct polyQ disease proteins and highly sequence-repetitive prions, much interest extends also to the repetitive dipeptide-repeats associated with ALS disease (Odeh and Shorter, 2020; Schmitz et al., 2021). The aggregated and phase-separated states of these proteins remain as yet poorly understood, with studies by ssNMR likely being important to understanding their structure, dynamics and phase behavior. However, repetitive sequences go well beyond amyloidogenic proteins. A different example is collagen, which is an essential

component of the extracellular matrix (ECM), where it helps define the structural characteristics of tissues. The (mechanical) properties of the ECM are of substantial biological interest, for instance in the context of cancer research and treatments (Venning et al., 2015). Already a topic of significant ssNMR studies (Goldberga et al., 2018), collagen's repetitive structure yields highly challenging spectra with highly overlapping signals, akin to the polyQ case study above. Dihedral angle measurements may be similarly useful for nonetheless probing the local structure in atomic detail. Similarly, the silk proteins produced by spiders and other animals are highly repetitive in sequence and have interesting structural properties that have been studied by ssNMR (van Beek and Meier, 2006; Holland et al., 2008). Also here dihedral angle measurements can be useful to probe their structures and structural transitions (van Beek and Meier, 2006).

Applications Beyond Polypeptides

Mostly we have focused on the study of protein samples. However, the use of dihedral angle measurements is also of interest for the study of other biomolecules (or non-biological samples). Besides repetitive protein structures, we may also consider applications to non-protein macromolecules of biological, biomedical and bioengineering interest. For instance (high-molecular weight) polysaccharides, RNA and DNA are increasingly studied by ssNMR, but present new challenges in terms of structure determination (Marchanka et al., 2015; El Hariri El Nokab and van der Wel, 2020).

However, also in small molecules (or short peptides), it can be difficult to gain sufficient information from distance constraints alone. In such a case, dihedral angle measurements can play an important role. In the context of peptides, this has now been well demonstrated (Rienstra et al., 2002b; Jaroniec et al., 2004; Bajaj et al., 2009; Barnes et al., 2009; van der Wel et al., 2010). Nice examples can also be found in the earlier literature, for example in studies of the retinal of the membrane protein rhodopsin (Feng et al., 1997b). The recent literature offers several other interesting case studies, as we have already seen above, in which torsion angle measurements probe the structure of small molecules bound to proteins, rather than the protein itself. This includes the example of a small molecule substrate bound to a receptor protein (**Figure 5**) (Edwards et al., 2010) and the amyloid-specific fluorescent dye congo red bound to HET-s fibrils (Gowda et al., 2017). It is likely that further applications like this can be expected in future work on, e.g., drug-protein, substrate-enzyme and other such interactions.

TECHNICAL CHALLENGES AND OPPORTUNITIES

As with any (ssNMR) experiment, the torsion angle measurements offer both unique strengths and also specific challenges. This section notes a few specific challenges, but follows this with perspectives on how they can be overcome (partly with the help of modern MAS ssNMR equipment). Like with other structural ssNMR studies reliant on anisotropic interactions, any dynamics in the molecular system can interfere with the execution and/or analysis of torsion angles. Since dynamics modulate both dipolar interactions and CSA tensors, this would naturally cause problems. These dynamics can be important and relevant, since biological contexts often require dynamics (e.g., in enzymes or ion channels) to achieve proper function. On the one hand, it may be possible to account for certain types of (anisotropic) motion in the data analysis, as exemplified in prior studies that characterized such motion by ssNMR (Hu et al., 2010; Li and Hong, 2011). Another workaround could be the use of low-temperature experiments, which are increasingly accessible with the enhanced availability of low-temperature dynamic nuclear polarization (DNP) equipment (Lilly Thankamony et al., 2017), and can permit the execution of dihedral angle measurements at reduced temperatures where molecular motion is suppressed (Bajaj et al., 2009).

One challenge independent of motion is the inherently low sensitivity of the experiments. Whilst applications to crystalline peptides have been quite effective and convincing, applications to more complex systems are fairly demanding. As already discussed above, a combination of factors contribute to this challenge. On the one hand, the nature of the pulse sequences is such the signals are purposely decreased, and we monitor the differential degree of dephasing to distinguish the differences in structure. Crucially, for several types of measurements certain ranges of angles are close together in terms of their dephasing curves. This places significant demands on the signal to noise. The good news is the techniques and equipment available for MAS ssNMR have greatly improved since the dihedral angle measurements were first developed. On the one hand, we already discussed DNP for its low temperature capabilities. This technique also offers substantial signal enhancements, which even permit applications to natural abundance proteins and materials within reach. Notably, recent DNP studies of unlabeled proteins and other organic (bio)molecules (Märker et al., 2017; Smith et al., 2018) already depend heavily on the kinds of DQ experiments at the heart of several types of dihedral angle measurements, as reviewed above.

In addition to DNP technology, we also see the gains in the application of high field ssNMR and especially also (^{1}H-detected) fast MAS (>60 kHz). Most of the example data discussed in this review article were obtained with MAS rates between 10 and 15 kHz (reflecting typically 3.2 and 4 mm MAS rotor diameters). High MAS rates that exceed even 100 kHz are now available, permitting ^{1}H detection and other new pulse sequences (Barbet-Massin et al., 2014; Zhang et al., 2017; Ji et al., 2021). This can enable new types of experiments and thereby new structural insights, often with improved sensitivity and time-savings. However at the same time, some traditional pulse sequence elements become difficult or impossible to implement, for instance due to the overly short rotor period lengths or the high RF power requirements. This for instance means that REDOR-based methods may not work, requiring the deployment of new schemes. This has opened up new developments of various sorts of MAS ssNMR experiments, such as tailored assignment and relaxation measurements (Barbet-Massin et al., 2014; Zhang et al., 2017; Ji et al., 2021),

but the development of (new) dihedral angle experiments has lagged behind (Hou et al., 2010).

A final practical challenge relates to the implementation and analysis of torsion angle experiments. Fewer groups have hands-on experience with dihedral angle experiments, compared to the more widely used distance measurements. Moreover, the configuration and implementation of these experiments is arguably more involved than, e.g., a traditional DARR-based distance measurement. Also the interpretation of the obtained results is perhaps less intuitive than looking for the presence or absence of cross peaks in typical distance measurements. One helpful development is the availability of various efficient and flexible numerical simulation packages, which can facilitate both test simulations to better understand the use of these experiments, and can also enable efficient analysis and interpretation of obtained results (Bak et al., 2000; Veshtort and Griffin, 2006).

With all these technical and hardware enhancements, it seems likely that many new and useful dihedral angle measurements may be introduced in the future. And it also seems plausible that existing techniques may find wider adoption and application to suitable systems, whether biological or non-biological in nature.

DISCUSSION

In this article we have examined several recent applications and methodologies of ssNMR dihedral angle measurements, focusing on those techniques based on the recoupling of anisotropic dipolar and/or CSA tensors. With the advent of highly productive distance and shift-based structure determination techniques for ssNMR-based structural biology, these direct dihedral angle measurements have been a bit left outside the mainstream. Yet, we have seen how they can be powerful and essential for tackling various biologically important questions, ranging from amyloid fiber structure determination to receptor-substrate interactions. It is furthermore anticipated that there is significant room for enhancing the utility of these techniques, as future studies will surely integrate these techniques with modern MAS ssNMR techniques such as ^{1}H-detected fast MAS and DNP. Thus, we foresee an expansion of the role of these methods toward broader application in the ssNMR community, with valuable roles in studies of biological and non-biological systems, both with and without stable-isotope labeling.

AUTHOR CONTRIBUTIONS

PvdW wrote and edited the article.

FUNDING

Our studies of polyglutamine amyloid structure have been funded by the NIH NIGMS R01 GM112678 (United States) and the Campagne Team Huntington foundation in Netherlands.

ACKNOWLEDGMENTS

The author thanks members of the Van der Wel group for their contributions to the discussed work and comments on the article draft. I also thank Matthias Ernst for kindly providing the data plotted in **Figure 8C**. Molecular graphics were made with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.

REFERENCES

Armstrong, N., and Gouaux, E. (2000). Mechanisms for Activation and Antagonism of an AMPA-Sensitive Glutamate Receptor. *Neuron* 28, 165–181. doi:10.1016/S0896-6273(00)00094-5

Bajaj, V. S., van der Wel, P. C. A., and Griffin, R. G. (2009). Observation of a Low-Temperature, Dynamically Driven Structural Transition in a Polypeptide by Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 131, 118–128. doi:10.1021/ja8045926

Bak, M., Rasmussen, J. T., and Nielsen, N. C. (2000). SIMPSON: a General Simulation Program for Solid-State NMR Spectroscopy. *J. Magn. Reson.* 147, 296–330. doi:10.1006/jmre.2000.2179

Barbet-Massin, E., Pell, A. J., Retel, J. S., Andreas, L. B., Jaudzems, K., Franks, W. T., et al. (2014). Rapid Proton-Detected NMR Assignment for Proteins with Fast Magic Angle Spinning. *J. Am. Chem. Soc.* 136, 12489–12497. doi:10.1021/ja507382j

Barnes, A. B., Andreas, L. B., Huber, M., Ramachandran, R., van der Wel, P. C. A., Veshtort, M., et al. (2009). High-resolution Solid-State NMR Structure of Alanyl-Prolyl-glycine. *J. Magn. Reson.* 200, 95–100. doi:10.1016/j.jmr.2009.06.009

Bayro, M. J., Huber, M., Ramachandran, R., Davenport, T. C., Meier, B. H., Ernst, M., et al. (2009). Dipolar Truncation in Magic-Angle Spinning NMR Recoupling Experiments. *J. Chem. Phys.* 130, 114506. doi:10.1063/1.3089370

Bower, P. V., Oyler, N., Mehta, M. A., Long, J. R., Stayton, P. S., and Drobny, G. P. (1999). Determination of Torsion Angles in Proteins and Peptides Using Solid State NMR. *J. Am. Chem. Soc.* 121, 8373–8375. doi:10.1021/ja991330q

Caulkins, B. G., Bastin, B., Yang, C., Neubauer, T. J., Young, R. P., Hilario, E., et al. (2014). Protonation States of the Tryptophan Synthase Internal Aldimine Active Site from Solid-State NMR Spectroscopy: Direct Observation of the Protonated Schiff Base Linkage to Pyridoxal-5′-Phosphate. *J. Am. Chem. Soc.* 136, 12824–12827. doi:10.1021/ja506267d

Chan, J. C. C., and Tycko, R. (2003). Solid-State NMR Spectroscopy Method for Determination of the Backbone Torsion Angle ψ in Peptides with Isolated Uniformly Labeled Residues. *J. Am. Chem. Soc.* 125, 11828–11829. doi:10.1021/ja0369820

Cornilescu, G., Delaglio, F., and Bax, A. (1999). Protein Backbone Angle Restraints from Searching a Database for Chemical Shift and Sequence Homology. *J. Biomol. NMR* 13, 289–302. doi:10.1023/a:1008392405740

Costa, P. R., Gross, J. D., Hong, M., and Griffin, R. G. (1997). Solid-state NMR Measurement of Ψ in Peptides: a NCCN 2Q-Heteronuclear Local Field experiment. *Chem. Phys. Lett.* 280, 95–103. doi:10.1016/s0009-2614(97)01107-x

Dunbrack, R. L. (2002). Rotamer Libraries in the 21st century. *Curr. Opin. Struct. Biol.* 12, 431–440. doi:10.1016/s0959-440x(02)00344-5

Edwards, R., Madine, J., Fielding, L., and Middleton, D. A. (2010). Measurement of Multiple Torsional Angles from One-Dimensional Solid-State NMR Spectra: Application to the Conformational Analysis of a Ligand in its Biological Receptor Site. *Phys. Chem. Chem. Phys.* 12, 13999–14008. doi:10.1039/c0cp00326c

El Hariri El Nokab, M., and van der Wel, P. C. A. (2020). Use of Solid-State NMR Spectroscopy for Investigating Polysaccharide-Based Hydrogels: A Review. *Carbohydr. Polym.* 240, 116276. doi:10.1016/j.carbpol.2020.116276

Feng, X., Edén, M., Brinkmann, A., Luthman, H., Eriksson, L., Gräslund, A., et al. (1997a). Direct Determination of a Peptide Torsional Angle ψ by Double-Quantum Solid-State NMR. *J. Am. Chem. Soc.* 119, 12006–12007. doi:10.1021/ja972252e

Feng, X., Lee, Y. K., Sandström, D., Edén, M., Maisel, H., Sebald, A., et al. (1996). Direct Determination of a Molecular Torsional Angle by Solid-State NMR. *Chem. Phys. Lett.* 257, 314–320. doi:10.1016/0009-2614(96)00558-1

Feng, X., Verdegem, P. J. E., Lee, Y. K., Sandström, D., Edén, M., Bovee-Geurts, P., et al. (1997b). Direct Determination of a Molecular Torsional Angle in the Membrane Protein Rhodopsin by Solid-State NMR. *J. Am. Chem. Soc.* 119, 6853–6857. doi:10.1021/ja970710d

Franks, W. T., Wylie, B. J., Schmidt, H. L. F., Nieuwkoop, A. J., Mayrhofer, R.-M., Shah, G. J., et al. (2008). Dipole Tensor-Based Atomic-Resolution Structure Determination of a Nanocrystalline Protein by Solid-State NMR. *Proc. Natl. Acad. Sci.* 105, 4621–4626. doi:10.1073/pnas.0712393105

Franks, W. T., Wylie, B. J., Stellfox, S. A., and Rienstra, C. M. (2006). Backbone Conformational Constraints in a Microcrystalline U-15n-Labeled Protein by 3D Dipolar-Shift Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 128, 3154–3155. doi:10.1021/ja058292x

Franzmann, T. M., and Alberti, S. (2019). Prion-like Low-Complexity Sequences: Key Regulators of Protein Solubility and Phase Behavior. *J. Biol. Chem.* 294, 7128–7136. doi:10.1074/jbc.TM118.001190

Fujiwara, T., Shimomura, T., and Akutsu, H. (1997). Multidimensional Solid-State Nuclear Magnetic Resonance for Correlating Anisotropic Interactions under Magic-Angle Spinning Conditions. *J. Magn. Reson.* 124, 147–153. doi:10.1006/jmre.1996.1022

Goldberga, I., Li, R., and Duer, M. J. (2018). Collagen Structure-Function Relationships from Solid-State NMR Spectroscopy. *Acc. Chem. Res.* 51, 1621–1629. doi:10.1021/acs.accounts.8b00092

Gowda, C., Zandomeneghi, G., Zimmermann, H., Schütz, A. K., Böckmann, A., Ernst, M., et al. (2017). The Conformation of the Congo-red Ligand Bound to Amyloid Fibrils HET-S(218-289): a Solid-State NMR Study. *J. Biomol. NMR* 69, 207–213. doi:10.1007/s10858-017-0148-z

Hintze, B. J., Lewis, S. M., Richardson, J. S., and Richardson, D. C. (2016). Molprobity's Ultimate Rotamer-Library Distributions for Model Validation. *Proteins* 84, 1177–1189. doi:10.1002/prot.25039

Holland, G. P., Creager, M. S., Jenkins, J. E., Lewis, R. V., and Yarger, J. L. (2008). Determining Secondary Structure in Spider Dragline Silk by Carbon–Carbon Correlation Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 130, 9871–9877. doi:10.1021/ja8021208

Hong, M. (2006). "Torsion Angle Determination by Solid-State NMR," in *Modern Magnetic Resonance* (Dordrecht: Springer), 723–729.

Hong, M. (1999). Determination of Multiple φ-Torsion Angles in Proteins by Selective and Extensive 13C Labeling and Two-Dimensional Solid-State NMR. *J. Magn. Reson.* 139, 389–401. doi:10.1006/jmre.1999.1805

Hong, M., Gross, J. D., and Griffin, R. G. (1997a). Site-Resolved Determination of Peptide Torsion Angle φ from the Relative Orientations of Backbone N–H and C–H Bonds by Solid-State NMR. *J. Phys. Chem. B* 101, 5869–5874. doi:10.1021/jp970887u

Hong, M., Gross, J. D., Hu, W., and Griffin, R. G. (1998). Determination of the Peptide Torsion Angle φ by15N Chemical Shift and13Cα-1HαDipolar Tensor Correlation in Solid-State MAS NMR. *J. Magn. Reson.* 135, 169–177. doi:10.1006/jmre.1998.1573

Hong, M., Gross, J. D., Rienstra, C. M., Griffin, R. G., Kumashiro, K. K., and Schmidt-Rohr, K. (1997b). Coupling Amplification in 2D MAS NMR and its Application to Torsion Angle Determination in Peptides. *J. Magn. Reson.* 129, 85–92. doi:10.1006/jmre.1997.1242

Hong, M., and Wi, S. (2006). "Torsion Angle Determination in Biological Solids by Solid-State Nuclear Magnetic Resonance," in *NMR Spectroscopy of Biological Solids*. Editor A. Ramamoorthy (Boca Raton, FL: CRC Press), 88–122.

Hoop, C. L., Lin, H.-K., Kar, K., Hou, Z., Poirier, M. A., Wetzel, R., et al. (2014). Polyglutamine Amyloid Core Boundaries and Flanking Domain Dynamics in Huntingtin Fragment Fibrils Determined by Solid-State Nuclear Magnetic Resonance. *Biochemistry* 53, 6653–6666. doi:10.1021/bi501010q

Hoop, C. L., Lin, H.-K., Kar, K., Magyarfalvi, G., Lamley, J. M., Boatz, J. C., et al. (2016). Huntingtin Exon 1 Fibrils Feature an Interdigitated β-hairpin-based Polyglutamine Core. *Proc. Natl. Acad. Sci. USA* 113, 1546–1551. doi:10.1073/pnas.1521933113

Hou, G., Paramasivam, S., Byeon, I.-J. L., Gronenborn, A. M., and Polenova, T. (2010). Determination of Relative Tensor Orientations by γ-encoded Chemical Shift Anisotropy/heteronuclear Dipolar Coupling 3D NMR Spectroscopy in Biological Solids. *Phys. Chem. Chem. Phys.* 12, 14873–14883. doi:10.1039/c0cp00795a

Hovmöller, S., Zhou, T., and Ohlson, T. (2002). Conformations of Amino Acids in Proteins. *Acta Crystallogr. D Biol. Cryst.* 58, 768–776. doi:10.1107/S0907444902003359

Hu, F., Luo, W., and Hong, M. (2010). Mechanisms of Proton Conduction and Gating in Influenza M2 Proton Channels from Solid-State NMR. *Science* 330, 505–508. doi:10.1126/science.1191714

Hu, K.-N., Qiang, W., Bermejo, G. A., Schwieters, C. D., and Tycko, R. (2012). Restraints on Backbone Conformations in Solid State NMR Studies of Uniformly Labeled Proteins from Quantitative Amide 15N-15N and Carbonyl 13C-13C Dipolar Recoupling Data. *J. Magn. Reson.* 218, 115–127. doi:10.1016/j.jmr.2012.03.001

Isas, J. M., Langen, R., and Siemer, A. B. (2015). Solid-State Nuclear Magnetic Resonance on the Static and Dynamic Domains of Huntingtin Exon-1 Fibrils. *Biochemistry* 54, 3942–3949. doi:10.1021/acs.biochem.5b00281

Ishii, Y., Terao, T., and Kainosho, M. (1996). Relayed Anisotropy Correlation NMR: Determination of Dihedral Angles in Solids. *Chem. Phys. Lett.* 256, 133–140. doi:10.1016/0009-2614(96)00426-5

IUPAC-IUB Commission on Biochemical Nomenclature (1970). IUPAC-IUB Commission on Biochemical Nomenclature. Abbreviations and Symbols for the Description of the Conformation of Polypeptide Chains. *J. Mol. Biol.* 52, 1–17. doi:10.1016/0022-2836(70)90173-7

Jaroniec, C. P., MacPhee, C. E., Bajaj, V. S., McMahon, M. T., Dobson, C. M., and Griffin, R. G. (2004). High-resolution Molecular Structure of a Peptide in an Amyloid Fibril Determined by Magic Angle Spinning NMR Spectroscopy. *Proc. Natl. Acad. Sci.* 101, 711–716. doi:10.1073/pnas.0304849101

Ji, Y., Liang, L., Bao, X., and Hou, G. (2021). Recent Progress in Dipolar Recoupling Techniques under Fast MAS in Solid-State NMR Spectroscopy. *Solid State. Nucl. Magn. Reson.* 112, 101711. doi:10.1016/j.ssnmr.2020.101711

Ladizhansky, V. (2009). "Dipolar-Based Torsion Angle Measurements for Protein Structure Determination," in *Encyclopedia of Magnetic Resonance*. Editors R. K. Harris and R. Wasylishen (Chichester, United Kingdom: John Wiley). doi:10.1002/9780470034590.emrstm1153

Ladizhansky, V., Jaroniec, C. P., Diehl, A., Oschkinat, H., and Griffin, R. G. (2003). Measurement of Multiple ψ Torsion Angles in Uniformly 13C,15N-Labeled α-Spectrin SH3 Domain Using 3D 15N−13C−13C−15N MAS Dipolar-Chemical Shift Correlation Spectroscopy. *J. Am. Chem. Soc.* 125, 6827–6833. doi:10.1021/ja029082c

Ladizhansky, V., Veshtort, M., and Griffin, R. G. (2002). NMR Determination of the Torsion Angle Ψ in α-Helical Peptides and Proteins: The HCCN Dipolar Correlation Experiment. *J. Magn. Reson.* 154, 317–324. doi:10.1006/jmre.2001.2488

Lewandowski, J. R., van der Wel, P. C. A., Rigney, M., Grigorieff, N., and Griffin, R. G. (2011). Structural Complexity of a Composite Amyloid Fibril. *J. Am. Chem. Soc.* 133, 14686–14698. doi:10.1021/ja203736z

Li, S., and Hong, M. (2011). Protonation, Tautomerization, and Rotameric Structure of Histidine: a Comprehensive Study by Magic-Angle-Spinning Solid-State NMR. *J. Am. Chem. Soc.* 133, 1534–1544. doi:10.1021/ja108943n

Lilly Thankamony, A. S., Kaushik, M., Corzilius, B., and Corzilius, B. (2017). Dynamic Nuclear Polarization for Sensitivity Enhancement in Modern Solid-State NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* 102-103, 120–195. doi:10.1016/j.pnmrs.2017.06.002

Loquet, A., Bardiaux, B., Gardiennet, C., Blanchet, C., Baldus, M., Nilges, M., et al. (2008). 3D Structure Determination of the Crh Protein from Highly Ambiguous Solid-State NMR Restraints. *J. Am. Chem. Soc.* 130, 3579–3589. doi:10.1021/ja078014t

Lovell, S. C., Word, J. M., Richardson, J. S., and Richardson, D. C. (2000). The Penultimate Rotamer Library. *Proteins* 40, 389–408. doi:10.1002/1097-0134(20000815)40:3<389:aid-prot50>3.0.co;2-2

Marchanka, A., Simon, B., Althoff-Ospelt, G., and Carlomagno, T. (2015). RNA Structure Determination by Solid-State NMR Spectroscopy. *Nat. Commun.* 6, 7024. doi:10.1038/ncomms8024

Märker, K., Paul, S., Fernández-de-Alba, C., Lee, D., Mouesca, J.-M., Hediger, S., et al. (2017). Welcoming Natural Isotopic Abundance in Solid-State NMR: Probing π-stacking and Supramolecular Structure of Organic Nanoassemblies Using DNP. *Chem. Sci.* 8, 974–987. doi:10.1039/C6SC02709A

Markley, J. L., Bax, A., Arata, Y., Hilbers, C. W., Kaptein, R., Sykes, B. D., et al. (1998). Recommendations for the Presentation of NMR Structures of Proteins and Nucleic Acids (IUPAC Recommendations 1998). *Pure Appl. Chem.* 70, 117–142. doi:10.1351/pac199870010117

Matlahov, I., and van der Wel, P. C. (2019). Conformational Studies of Pathogenic Expanded Polyglutamine Protein Deposits from Huntington's Disease. *Exp. Biol. Med. (Maywood)* 244, 1584–1595. doi:10.1177/1535370219856620

Mukhopadhyay, D., Gupta, C., Theint, T., and Jaroniec, C. P. (2018). Peptide Bond Conformation in Peptides and Proteins Probed by Dipolar Coupling-Chemical Shift Tensor Correlation Solid-State NMR. *J. Magn. Reson.* 297, 152–160. doi:10.1016/j.jmr.2018.10.015

Musacchio, A., Noble, M., Pauptit, R., Wierenga, R., and Saraste, M. (1992). Crystal Structure of a Src-Homology 3 (SH3) Domain. *Nature* 359, 851–855. doi:10.1038/359851a0

Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. Ø., Riekel, C., Grothe, R., et al. (2005). Structure of the Cross-β Spine of Amyloid-like Fibrils. *Nature* 435, 773–778. doi:10.1038/nature03680

Odeh, H. M., and Shorter, J. (2020). Arginine-rich Dipeptide-Repeat Proteins as Phase Disruptors in C9-ALS/FTD. *Emerging Top. Life Sci.* 4, 293–305. doi:10.1042/ETLS20190167

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., et al. (2021). UCSF ChimeraX : Structure Visualization for Researchers, Educators, and Developers. *Protein Sci.* 30, 70–82. doi:10.1002/pro.3943

Pope, G. M., Hung, I., Gan, Z., Mobarak, H., Widmalm, G., and Harper, J. K. (2018). Exploiting 13C/14N Solid-State NMR Distance Measurements to Assign Dihedral Angles and Locate Neighboring Molecules. *Chem. Commun.* 54, 6376–6379. doi:10.1039/c8cc02597e

Ramachandran, G. N., and Sasisekharan, V. (1968). "Conformation of Polypeptides and Proteins," in *Advances in Protein Chemistry* (Elsevier), 283–437. doi:10.1016/S0065-3233(08)60402-7

Reif, B., Hohwy, M., Jaroniec, C. P., Rienstra, C. M., and Griffin, R. G. (2000). NH-NH Vector Correlation in Peptides by Solid-State NMR. *J. Magn. Reson.* 145, 132–141. doi:10.1006/jmre.2000.2067

Rienstra, C. M., Hohwy, M., Mueller, L. J., Jaroniec, C. P., Reif, B., and Griffin, R. G. (2002a). Determination of Multiple Torsion-Angle Constraints in U−13C,15N-Labeled Peptides: 3D 1H−15N−13C−1H Dipolar Chemical Shift NMR Spectroscopy in Rotating Solids. *J. Am. Chem. Soc.* 124, 11908–11922. doi:10.1021/ja020802p

Rienstra, C. M., Tucker-Kellogg, L., Jaroniec, C. P., Hohwy, M., Reif, B., McMahon, M. T., et al. (2002b). De Novo determination of Peptide Structure with Solid-State Magic-Angle Spinning NMR Spectroscopy. *Proc. Natl. Acad. Sci.* 99, 10260–10265. doi:10.1073/pnas.152346599

Sawaya, M. R., Hughes, M. P., Rodriguez, J. A., Riek, R., and Eisenberg, D. S. (2021). The Expanding Amyloid Family: Structure, Stability, Function, and Pathogenesis. *Cell* 184, 4857–4873. doi:10.1016/j.cell.2021.08.013

Schmitz, A., Pinheiro Marques, J., Oertig, I., Maharjan, N., and Saxena, S. (2021). Emerging Perspectives on Dipeptide Repeat Proteins in C9ORF72 ALS/FTD. *Front. Cel. Neurosci.* 15, 637548. doi:10.3389/fncel.2021.637548

Schneider, R., Schumacher, M. C., Mueller, H., Nand, D., Klaukien, V., Heise, H., et al. (2011). Structural Characterization of Polyglutamine Fibrils by Solid-State NMR Spectroscopy. *J. Mol. Biol.* 412, 121–136. doi:10.1016/j.jmb.2011.06.045

Schütz, A. K., Soragni, A., Hornemann, S., Aguzzi, A., Ernst, M., Böckmann, A., et al. (2011). The Amyloid-Congo Red Interface at Atomic Resolution. *Angew. Chem. Int. Ed.* 50, 5956–5960. doi:10.1002/anie.201008276

Sharma, D., Shinchuk, L. M., Inouye, H., Wetzel, R., and Kirschner, D. A. (2005). Polyglutamine Homopolymers Having 8-45 Residues Form Slablike β-crystallite Assemblies. *Proteins* 61, 398–411. doi:10.1002/prot.20602

Shen, Y., and Bax, A. (2007). Protein Backbone Chemical Shifts Predicted from Searching a Database for Torsion Angle and Sequence Homology. *J. Biomol. NMR* 38, 289–302. doi:10.1007/s10858-007-9166-6

Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009). TALOS+: a Hybrid Method for Predicting Protein Backbone Torsion Angles from NMR Chemical Shifts. *J. Biomol. NMR* 44, 213–223. doi:10.1007/s10858-009-9333-z

Sinha, N., and Hong, M. (2003). X-1H Rotational-echo Double-Resonance NMR for Torsion Angle Determination of Peptides. *Chem. Phys. Lett.* 380, 742–748. doi:10.1016/j.cplett.2003.09.088

Sivanandam, V. N., Jayaraman, M., Hoop, C. L., Kodali, R., Wetzel, R., and van der Wel, P. C. A. (2011). The Aggregation-Enhancing Huntingtin N-Terminus Is Helical in Amyloid Fibrils. *J. Am. Chem. Soc.* 133, 4558–4566. doi:10.1021/ja110715f

Smith, A. N., Märker, K., Piretra, T., Boatz, J. C., Matlahov, I., Kodali, R., et al. (2018). Structural Fingerprinting of Protein Aggregates by Dynamic Nuclear Polarization-Enhanced Solid-State NMR at Natural Isotopic Abundance. *J. Am. Chem. Soc.* 140, 14576–14580. doi:10.1021/jacs.8b09002

Torbeev, V. Y., and Hilvert, D. (2013). Both the Cis-Trans Equilibrium and Isomerization Dynamics of a Single Proline Amide Modulate 2-microglobulin Amyloid Assembly. *Proc. Natl. Acad. Sci.* 110, 20051–20056. doi:10.1073/pnas.1310414110

van Beek, J. D., and Meier, B. H. (2006). A DOQSY Approach for the Elucidation of Torsion Angle Distributions in Biopolymers: Application to Silk. *J. Magn. Reson.* 178, 106–120. doi:10.1016/j.jmr.2005.09.004

van der Wel, P. C. A., Eddy, M. T., Ramachandran, R., and Griffin, R. G. (2009). Targeted13C-13C Distance Measurements in a Microcrystalline Protein via J-Decoupled Rotational Resonance Width Measurements. *ChemPhysChem* 10, 1656–1663. doi:10.1002/cphc.200900102

van der Wel, P. C. A. (2017). Insights into Protein Misfolding and Aggregation Enabled by Solid-State NMR Spectroscopy. *Solid State. Nucl. Magn. Reson.* 88, 1–14. doi:10.1016/j.ssnmr.2017.10.001

van der Wel, P. C. A., Lewandowski, J. R., and Griffin, R. G. (2007). Solid-state NMR Study of Amyloid Nanocrystals and Fibrils Formed by the Peptide GNNQQNY from Yeast Prion Protein Sup35p. *J. Am. Chem. Soc.* 129, 5117–5130. doi:10.1021/ja068633m

van der Wel, P. C. A., Lewandowski, J. R., and Griffin, R. G. (2010). Structural Characterization of GNNQQNY Amyloid Fibrils by Magic Angle Spinning NMR. *Biochemistry* 49, 9457–9469. doi:10.1021/bi100077x

Van der Wel, P. C. A. (2018). New Applications of Solid-State NMR in Structural Biology. *Emerg. Top. Life Sci.* 2, 57–67. doi:10.1042/ETLS20170088

Venning, F. A., Wullkopf, L., and Erler, J. T. (2015). Targeting ECM Disrupts Cancer Progression. *Front. Oncol.* 5, 1377. doi:10.3389/fonc.2015.00224

Veshtort, M., and Griffin, R. G. (2006). SPINEVOLUTION: a Powerful Tool for the Simulation of Solid and Liquid State NMR Experiments. *J. Magn. Reson.* 178, 248–282. doi:10.1016/j.jmr.2005.07.018

Weliky, D. P., and Tycko, R. (1996). Determination of Peptide Conformations by Two-Dimensional Magic Angle Spinning NMR Exchange Spectroscopy with Rotor Synchronization. *J. Am. Chem. Soc.* 118, 8487–8488. doi:10.1021/ja960178e

Wetzel, R. (2012). Physical Chemistry of Polyglutamine: Intriguing Tales of a Monotonous Sequence. *J. Mol. Biol.* 421, 466–490. doi:10.1016/j.jmb.2012.01.030

Wi, S., and Spano, J. (2011). Site-specific ϕ- and ψ-torsion Angle Determination in a Uniformly/extensively 13C- and 15N-Labeled Peptide. *J. Magn. Reson.* 212, 431–439. doi:10.1016/j.jmr.2011.08.019

Wylie, B. J., Bhate, M. P., and McDermott, A. E. (2014). Transmembrane Allosteric Coupling of the gates in a Potassium Channel. *Proc. Natl. Acad. Sci. USA* 111, 185–190. doi:10.1073/pnas.1319577110

Wylie, B. J., Sperling, L. J., Nieuwkoop, A. J., Franks, W. T., Oldfield, E., and Rienstra, C. M. (2011). Ultrahigh Resolution Protein Structures Using NMR Chemical Shift Tensors. *Proc. Natl. Acad. Sci.* 108, 16974–16979. doi:10.1073/pnas.1103728108

Yakupova, E. I., Bobyleva, L. G., Vikhlyantsev, I. M., and Bobylev, A. G. (2019). Congo Red and Amyloids: History and Relationship. *Biosci. Rep.* 39, BSR20181415. doi:10.1042/BSR20181415

Zhang, R., Mroue, K. H., and Ramamoorthy, A. (2017). Proton-Based Ultrafast Magic Angle Spinning Solid-State NMR Spectroscopy. *Acc. Chem. Res.* 50, 1105–1113. doi:10.1021/acs.accounts.7b00082

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 09 December 2021
doi: 10.3389/fmolb.2021.791026

Dipolar Order Parameters in Large Systems With Fast Spinning

W. Trent Franks[1,2]*, *Ben P. Tatman*[1,2], *Jonah Trenouth*[2] *and Józef R. Lewandowski*[2]*

[1]Department of Physics, University of Warwick, Coventry, United Kingdom, [2]Department of Chemistry, University of Warwick, Coventry, United Kingdom

Order parameters are a useful tool for quantifying amplitudes of molecular motions. Here we measure dipolar order parameters by recoupling heteronuclear dipole-dipole couplings under fast spinning. We apply symmetry based recoupling methods to samples spinning under magic angle at 60 kHz by employing a variable flip angle compound inversion pulse. We validate the methods by measuring site-specific ^{15}N-^{1}H order parameters of a microcrystalline protein over a small temperature range and the same protein in a large, precipitated complex with antibody. The measurements of the order parameters in the complex are consistent with the observed protein undergoing overall motion within the assembly.

Keywords: symmetry, dynamics, fast MAS NMR, recoupling, order parameter, proton detection

OPEN ACCESS

Edited by:
Mark Pfuhl,
King's College London,
United Kingdom

Reviewed by:
Johanna Becker-Baldus,
Goethe University Frankfurt, Germany
Sergey Dvinskikh,
Royal Institute of Technology, Sweden

Correspondence:
W. Trent Franks
t.franks@warwick.ac.uk
Józef R. Lewandowski
j.r.lewandowski@warwick.ac.uk

Specialty section:
This article was submitted to Structural Biology, a section of the journal Frontiers in Molecular Biosciences

Received: *07 October 2021*
Accepted: *05 November 2021*
Published: *09 December 2021*

Citation:
Franks WT, Tatman BP, Trenouth J and Lewandowski JR (2021) Dipolar Order Parameters in Large Systems With Fast Spinning. Front. Mol. Biosci. 8:791026. doi: 10.3389/fmolb.2021.791026

INTRODUCTION

The astounding number of structures found in the protein databank speaks to the usefulness of structural data to provide insights into the structure-function relationship in biology and biochemistry (Berman et al., 2000; Burley et al., 2021). With advent of powerful computational structure prediction approaches such as AlphaFold (Jumper et al., 2021) and RoseTTAFold (Baek et al., 2021) there is an almost exponential increase of systems for which a reasonable quality structure or models become available. However, a structure is a snapshot that does not necessarily capture the choreography of the protein it needs to execute in order to perform its function (Koshland, 1958; Frauenfelder et al., 1991). The motion of a protein is often intrinsic to its activity. Understanding the dynamics and how structure changes in time is sometimes nearly as important as knowing a single, even high-resolution, snapshot. An ultimate example of this idea are intrinsically disordered proteins (IDP) and intrinsically disordered regions (IDR), which are involved in controlling countless processes in eukaryotic but also prokaryotic organisms, *e.g.* biosynthetic steps in production of bioactive natural products (Jenner et al., 2018; Kosol et al., 2019; Fage et al., 2021).

Nuclear magnetic resonance (NMR), can be used to find molecular motions under near physiological conditions at atomic resolution over a several orders of magnitude of the time scale, from as fast as picoseconds, to as slow as months, but will only report on local conditions and typically over small distance scales (Palmer, 2004; Kovermann et al., 2016; Sekhar and Kay, 2019; Alderson and Kay, 2021). However, to access such a vast range of dynamics a battery of different tools reporting on different parameters of motion in different regimes is required. For example, NMR relaxation is sensitive to both amplitudes and time scales of motions typically in the picoseconds-nanoseconds range in solution and picoseconds-milliseconds range in the solid state (Lewandowski, 2013), which provides some unique opportunities for characterizing protein motions (Castellani et al., 2002; Chevelkov et al., 2003; Chevelkov et al., 2006; Lewandowski et al., 2011; Asami and Reif, 2012; Lamley et al., 2014; Lamley et al., 2015a; Lamley et al., 2015b; Sternberg et al., 2018; Öster et al.,

2019). However, the extended range of time scales of motions, which influence relaxation in the solid state, comes also at a price: reliable quantification of motional amplitudes with relaxation rates alone is challenging and sometimes impossible. Often reliable quantification of dynamics using relaxation rates requires them to be combined with measurements of order parameters (typically dipolar order parameters), which constrain the overall amplitude of motions (Schanda and Ernst, 2016).

Order parameters can be obtained by recoupling specific terms of the NMR Hamiltonian in a separated local field (SLF) experiment (Hester et al., 1976). SLF techniques use pulses to create a Hamiltonian where one term of the full Hamiltonian is recoupled, and the other terms are averaged to zero. Examples of such methods include C7 (Hohwy et al., 1998), RFDR (Bennett et al., 1992), REDOR (Gullion and Schaefer, 1989a; Gullion and Schaefer, 1989b), TMREV (Hohwy et al., 2000; Franks et al., 2005; Franks et al., 2006a), and many others. The recoupling portion of SLF experiments have been summarized into a uniform theory using symmetry principles (Levitt, 2002). The development has mainly been focused on slowly spinning samples, as fast rotation was not available at the time. However, fast magic angle spinning (MAS) NMR experiments, introduced since the main formulation of this theory, have improved the process of assignment and structure calculation of large proteins and complexes that were very difficult to solve using solid-state NMR otherwise (Zhou et al., 2007a; Barbet-Massin et al., 2010; Knight et al., 2012; Barbet-Massin et al., 2014). For example, the membrane protein OMPG had been actively studied for almost 15 years using carbon detection but the assignment and structure were finally solved with the use of fast spinning and proton detection (Hiller et al., 2005; Retel et al., 2017). Thus, it is desirable to extend symmetry methods to this attractive new regime. Unfortunately, symmetry sequences require applied fields that scale linearly with the spinning rate, and thus the applied field requirements cannot be usually achieved under fast spinning conditions. The application of symmetry principles to heteronuclear dipole-dipole recoupling was previously demonstrated under 40 kHz spinning (Hou et al., 2011), but did not engender optimism for application at higher spinning rates. However, spinning rates of ~60 kHz are routine at the time of writing of this manuscript, with 100 kHz spinning becoming more common, and current cutting-edge probes reach rates on the order of 150 kHz (Penzel et al., 2019; Schledorn et al., 2020) and even 200 kHz. Consequently, symmetry methods have not been applied extensively to fast spinning samples (Brinkmann and Levitt, 2001; Levitt, 2002). SLF experiments undertaken with spinning frequencies of 60 kHz or greater have been cross-polarization-based (CP) (Chevelkov et al., 2009; Chevelkov et al., 2010; Paluch et al., 2013; Paluch et al., 2018) or use phase modulated rotary resonance pulses (Liang et al., 2021). Symmetry-based recoupling is comparable to CP based SLF experiments and has many of the same disadvantages, but symmetry be advantageous in a few ways. First, the symmetry sequences can be constructed to be very selective of the terms allowed, where only a few terms in the NMR Hamiltonian are still active. Second, there is only one channel that has high-power pulses applied which limits the power deposition, where the CP methods apply high fields on both channels simultaneously.

In this work, we introduce an approach to generate pulse sequences with optimized recoupling at fast spinning given probe performance requirements. Candidate symmetry sequences are generated, the scaling factor (κ) is optimized *in silico* using variable flip angle pulse sequence elements, and then the highest performing sequences are selected. The most promising sequences are tested against B_1 inhomogeneity and match condition mis-set. The experimental performance of several candidate sequences is evaluated on a favorable model sample, the micro-crystalline protein GB1 (β1 immunoglobulin binding domain of protein G) which is uniformly labelled with ^{2}H, ^{13}C, ^{15}N, and then back-exchanged with ^{1}H at all exchangeable sites, and on a more challenging >300 kDa precipitated complex of GB1 with immunoglobulin G (IgG). We have previously investigated differences in protein dynamics for protein GB1 in these two environments. The analysis of various relaxation and relaxation dispersion experiments has indicated that while ps-ns motions and some μs motions appear to be largely similar for GB1 in the two environments, there appears to be an additional overall motional mode present only in GB1 in the complex with IgG (Lamley et al., 2015a; Öster et al., 2019). Thus we decided to investigate whether we see the presence of this additional dynamic mode reflected in the measured order parameters. In addition, we have previously performed variable temperature dynamics measurements on crystalline GB1 and in the analysis assumed the observed trends are dominated by changes to the time scales of the motions rather than changes in amplitude and assumed a constant order parameter (Lewandowski et al., 2015; Busi et al., 2018). The initial variable temperature molecular dynamics simulations suggested that indeed the ^{15}N-^{1}H order parameter changes only very slightly with temperature in the explored 30 °C range but we thought this study to be a good opportunity to begin to explore validity of such approximation experimentally. Incidentally, crystalline GB1 and GB1 in the complex with IgG cover the range of a favorable model sample and a challenging "real" sample.

Symmetry Based Pulse Sequences

Symmetry-based sequences allow for the selection of portions of the full NMR Hamiltonian (Levitt, 2002). The performance of the pulse sequence with regards to the extent of the reintroduction for an interaction is indicated by a scaling factor (κ). The scaling factor is the magnitude of any coupling when compared to the static limit, which can vary between 0 and 1, where a larger κ indicates a more efficient recoupling/reintroduction. The recoupling performance can be altered in two ways: by using a different symmetry, or by using a different rotation element. This work demonstrates a strategy to find high performance heteronuclear recoupling pulse sequences by exploring the possible variations of symmetry derived pulse sequences.

The primary limitation for the application of symmetry at high spinning frequencies is the electronic performance of the probe. The nutation of the spins which are induced by the applied radio

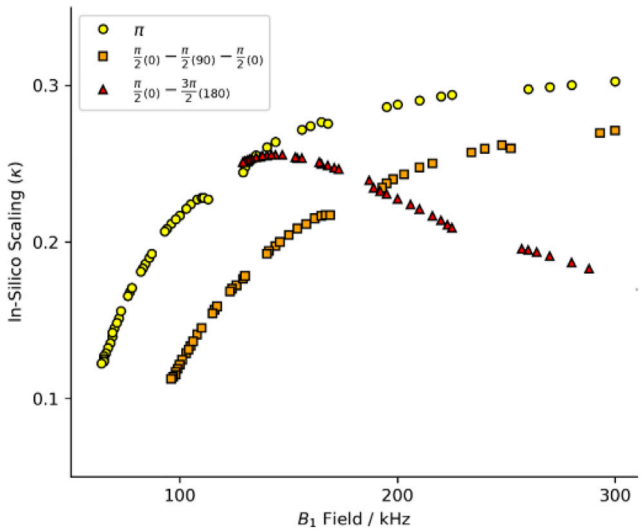


FIGURE 1 | Heteronuclear dipole-dipole recoupling scaling factor for symmetries found in Hou et al. (2011). The full range of symmetry match conditions for standard composite pulses. Yellow circles are for *R*-symmetries with a standard rectangular π pulse. Red triangles report the scaling with a compound *R*-element of a $\pi/2_{(0)}-\pi/2_{(90)}-\pi/2_{(0)}$ with the equivalent match of a 3π/2 pulse. Orange squares report the scaling with a compound *R*-element of a $\pi/2_{(0)}-3\pi/2_{(180)}$ with the equivalent match of a 2π pulse. See **Supplementary Figure S1** for full B_1 range.

frequency field in the probe must match the conditions specified by the symmetry sequence, which is dependent on the spinning rate. The same symmetry will require higher radio frequency fields with higher spinning rates. For example, in the well-known $R18_1^7$ sequence (Levitt, 2002; Wylie and Rienstra, 2008; Wylie et al., 2011) the match condition is 9 times the spinning rate, which means that ν_1 = 90 kHz at 10 kHz magic angle spinning (MAS) rate, but ν_1 = 540 kHz at 60 kHz spinning frequency. The electronics in the probe will likely break down due to the high voltages needed to generate such strong pulses, or alternatively, the protein sample will denature when the temperature gets too high from radiative heating. Most modern probes are specified to work with an applied ^{1}H field of ~100 kHz for long pulses (i.e., decoupling during acquisition), with fields on the order of 250 kHz available for short periods (and nonconductive samples). Since the prior art does not produce many appropriate choices, a method to identify and optimize symmetry-based recoupling to measure the heteronuclear dipole-dipole coupling was devised.

Initially, we intended to apply a series of symmetries identified for use at 40 kHz MAS (Hou et al., 2011). We simulated these schemes at 60 kHz MAS with a standard π pulse as the R-element (see **Figure 1**; **Supplementary Figure S1**, yellow circles). The curve of scaling factor against match condition matches the literature well but the scaling factor is smaller at 60 kHz spinning than at 40 kHz (Hou et al., 2011). That is, the scaling increases with increasing field, approaching some upper limit (here, κ ~ 0.3). We chose the best candidate symmetries to test experimentally where the criterion for selection was an applied field ν_1 < 130 kHz. After careful calibration of the applied field using long nutation experiments, the performance of the candidate sequences was found to be poor, with little to no recoupling apparent. The disappointing performance was attributed to poor ^{1}H channel B_1 homogeneity of the probe. The ratio of the NMR signal intensity at the 90°, 450° and 810° pulse is used as a proxy of an actual B_1 homogeneity measurement. The probes used in this study typically showed I_{450}/I_{90} ~70%, and I_{810}/I_{90} ~55% for the ^{1}H channel.

Composite pulses were implemented to compensate for the probe deficiencies since the standard symmetry recoupling experiments were unsuccessful. Two composite-π pulses are prominent in the literature, the $[90_{(0)}90_{(90)}90_{(0)}]$ and $[90_{(0)}270_{(180)}]$ where the flip angle of a portion of the pulse is denoted by the large number and the phase of that portion is denoted by the subscripted number in parentheses (Levitt, 2002). The comparison of the scaling factors determined *in silico* by simulations with the SIMPSON program (method described below) shows a dramatic change in the scaling factor when the symmetry element is changed. The performance of the $[90_{(0)}90_{(90)}90_{(0)}]$ composite pulse (**Figure 1**; **Supplementary Figure S1**, orange squares) follows the same general trend as the standard π pulse (yellow circles), but with worse efficiency and higher applied field. However, the $[90_{(0)}270_{(180)}]$ composite pulse (**Figure 1**; **Supplementary Figure S1**, red triangles) does not follow the same trend. The shape of the curve produced by the $[90_{(0)}270_{(180)}]$ composite pulse has a maximum in the curve, whereas the other pulses asymptotically rise. The maximum scaling factor is found at a relatively low field, with performance similar to the π pulse variant.

While it is made clear in the literature that the specifics of the *R*-elements contribute to the efficiency of the recoupling by altering the scaling factor (κ), the magnitude of this contribution was underappreciated. These preliminary simulations demonstrate that more symmetries than those found in the literature should be tried. Those with lower match conditions can be a viable option with composite pulse rotation elements. Also, a variety of symmetry elements will be useful to identify the best candidate sequences given the desired experimental conditions. To these ends, we present tools to generate appropriate symmetry lists, tools to test these symmetries, and experimentally test the best candidates.

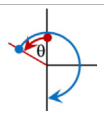
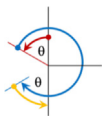
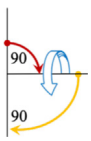
Generating Candidate Symmetry Sequences

An RN_n^ν or CN_n^ν multiple pulse sequence is applied such that "N" spin-space rotations are contained in "n" sample rotations and the phase (ϕ) of each element alternates as dictated by "ν" where ϕ = ±πν/N for R sequences. This averages some terms of the NMR Hamiltonian to zero, but not others. A brief discussion of the selection rules can be found in the symmetry selection rules section of the supporting information, and in depth in Levitt (2002) and references therein. Each element of the symmetry sequence is a specific rotation where an *R*-element is an inversion (π rotation) and a *C*-element has a 2π total rotation. Therefore, the amplitude for the radiofrequency (ν_1, B_1) match condition is proportional to the spinning rate and symmetry as

$$\upsilon_1 = k_p \frac{\omega_r}{2\pi} \frac{N}{2n}$$

where k_p is determined by the specifics of the *R* or *C*-element.

TABLE 1 | Selected variable flip angle *R*-elements, with field and timing dependencies. Red arrows represent the first portion of the composite pulse τ_1, blue is the second, τ_2, and yellow is the third portion, τ_3.

R-element	Diagram	Name	k_p ($B_1 = k_p\omega_r N/2n$)	$\tau_{1,2,3\ldots}$ ($2\tau_r n/N$)
$\theta_{(0)}$ $[180 + \theta]_{(180)}$		1a	(180 + 2θ)/180	τ_1 = θ/(180 + 2θ) τ_2= (180 + θ)/(180 + 2θ)
$\theta_{(0)}$ $[180 + 2\theta]_{(180)}$ $\theta_{(0)}$		2a	(180 + 4θ)/180	τ_1 = θ/(180 + 4θ) τ_2= (180 + 2θ)/(180 + 4θ) τ_3 = θ/(180 + 4θ)
$90_{(0)}$ $\theta_{(90)}$ $90_{(0)}$		4a	(180 + θ)/180	τ_1= (90)/(180 + θ) τ_2= (θ)/(180 + θ) τ_3= (90)/(180 + θ)

A custom program was written in Python 3 (see **Supplementary Material**: SymmetrySelector/SymSelect.py) to generate new symmetry sequences that are more relevant to faster spinning rates. This program reproduces the full list of sequences from Levitt (2002) (excepting 3 minor book-keeping errors, see symmetry selection rules section of the **Supplementary Information**; **Supplementary Tables S2–S4**). An arbitrarily large number of candidate symmetry sequences for application at 60 kHz spinning frequency is then generated. Symmetries for most terms in the Hamiltonian can be generated, with the possibility to limit the output based on experimental considerations. 203 symmetries fit the selection criteria (**Supplementary Table S5**); the sequence must allow heteronuclear dipole-dipole couplings, disallow homonuclear dipole-dipole coupling, and the base applied field is between 0.1 and 150 kHz. It should be noted that this program can be used to generate symmetry sequences for most spin ½ Hamiltonians, not just the heteronuclear dipole-dipole coupling.

Composite Rotation Element Pulses

Complex inversion pulses have been used for *R*-sequences, such as adiabatic inversions (Herbst et al., 2011; Herbst et al., 2015), numerically optimized optimal control pulses (Nielsen et al., 2009), and composite inversion pulses (Levitt, 1982). However, the *R*-element seems to have previously been chosen for some desirable property of the composite pulse and the expected shortcomings of the symmetry sequence, not explicitly for the performance of the sequence.

Since the timing and phase behavior of the multiple-pulse sequence must still fulfill the symmetry requirements the match field grows in proportion to the total arc swept out by the composite pulse. This term is the pulse contribution to the match condition "k_p," where:

$$k_p = \frac{\sum \text{Sweep Angle}^\circ}{180^\circ}$$

For example, a simple π inversion pulse sweeps an arc of 180°, so $k_p = 1$. For the composite pulses $[90_{(0)}180_{(90)}90_{(0)}]$ and $[90_{(0)}270_{(180)}]$ the RF sweeps out 360°, but the result is only an inversion of the polarization. These specific examples result in $k_p = 2$. The match field B_1 is, likewise, twice that of the symmetry alone.

It is unclear which composite pulses produce the highest scaling factor for the smallest applied field. The sweep angles for compound pulses have previously been constrained to only use the principal directions, i.e., they only include integer multiples of 90°, but such a constraint is not enforced in this study. For example, when we modify the $[90_{(0)}270_{(180)}]$ composite pulse we trade the initial $90_{(0)}$ portion for a $\theta_{(0)}$ pulse, and the $270_{(180)}$ becomes $(180 + \theta)_{(180)}$. The portions of the composite pulse are consecutively numbered $\tau_{1,2\ldots n}$.The field, timing dependencies, and a diagram of this element, named "1a," and two others are shown in **Table 1**. The "2a" element is a slight variation on 1a, it has an extra "wiggle" before finishing. The "4a" element is a variation of the $90_{(0)}180_{(90)}90_{(0)}$ element where the middle, out of plane, pulse is allowed to vary. It is possible to change θ continuously, and smoothly for these compound pulses and the end point (inversion) will not be changed. In total, 22 *R*-elements were constructed where the sweep angle (and thus the applied field) can be varied continuously, but which always produces a traceable inversion pulse. A simple inversion occurs when θ is zero for many, but not all the *R*-elements. All 22 of the *R*-elements with their pulse timings, field match dependence, and a visualization is found in **Supplementary Table S6**.

SIMPSON SIMULATIONS

Numerical simulations were conducted to determine the scaling factor with the SIMPSON NMR calculation software (Bak et al., 2000). The numerical scaling factor is comparable to the analytical solution presented in Levitt (2002), but differs in magnitude by up to 0.05, where the numerical method always over-estimates the scaling in comparison to the analytical

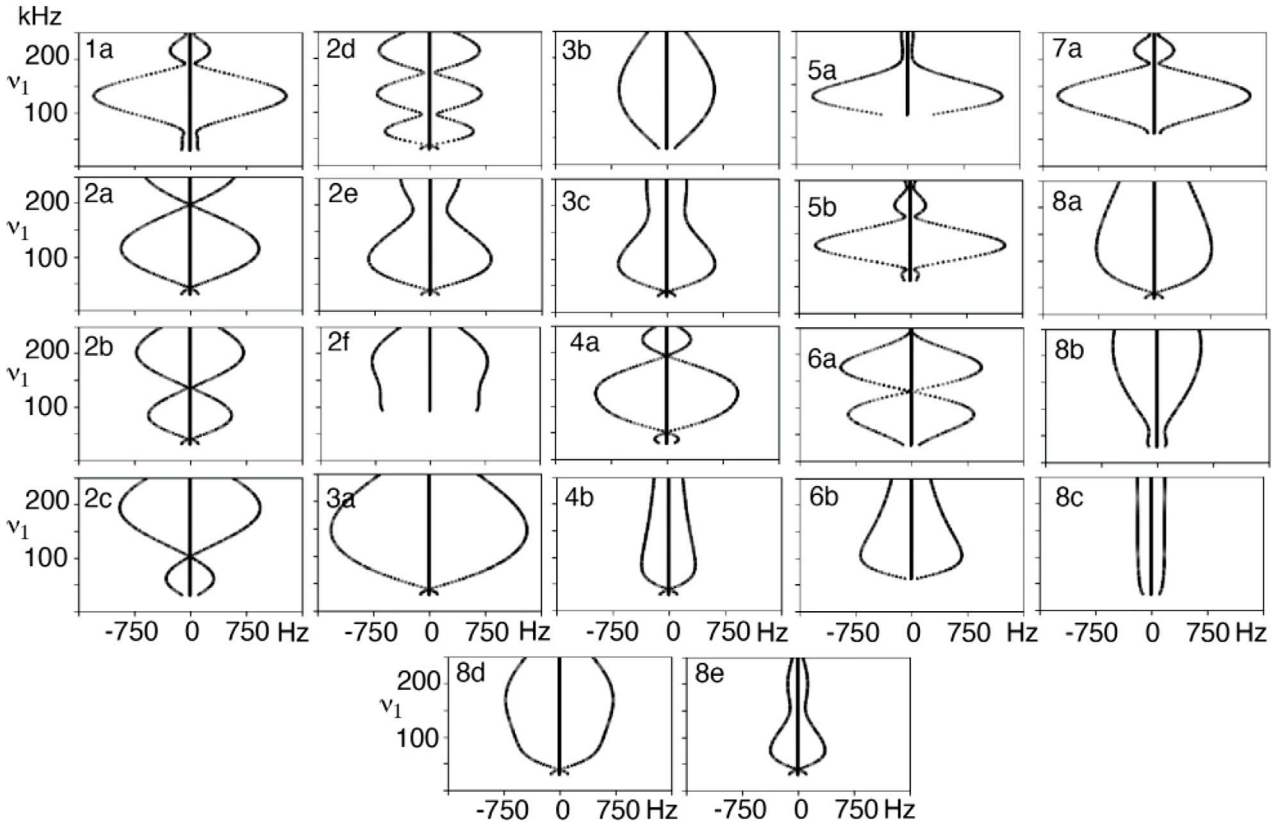


FIGURE 2 | Dipole-dipole recoupling spectra of the $R22_{21}^{9}$ symmetry as a function of ν_1 match condition for different variable flip angle pulse sequence elements.

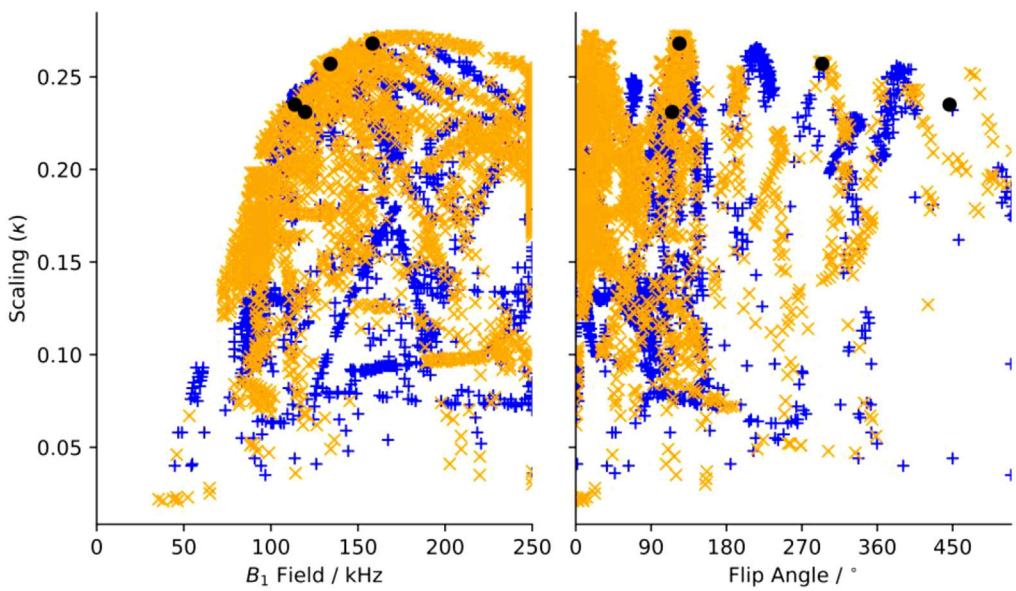


FIGURE 3 | Maximum scaling factor for heteronuclear dipole-dipole scaling factor determined *in silico* of all test symmetries and pulse sequence elements as a function of (a) applied (B_1) field and of additional tilt angle (θ). Blue"+"'s indicate sequences where "m" and "μ" are correlated, and orange "x"s are anti-correlated. Black circles indicate the position of sequences chosen for experimental verification.

solution. The spin system for testing the scaling factor is an isolated two spin (^{1}H-^{15}N) system with the dipole-dipole coupling defined to be 10 kHz which corresponds to an internuclear distance of 1.07 Å. The single crystallite crystal file "alpha0beta90" was used so that the maximum dipole-dipole coupling will be produced using the method of Hou et al. (2011) (via personal correspondence).

A series of simulations were performed where the applied field was varied from the lowest match condition up to 250 kHz for the composite *R*-elements to produce dipole recoupling efficiency curves. These curves are well-defined, as demonstrated by the simulations of the symmetry sequence $R22_{21}^{9}$ in **Figure 2**. It is convenient to vary the applied field and back calculate the variable flip angles and pulse durations to maintain a valid range for the applied field. The choice of the recoupling element results in decoupling in a sequence which is supposed to recouple, which may point to further selection rules for symmetry elements since there are several instances which result in a coupling of zero. These zero-coupling points might be useful for other applications such as decoupling. These zero points are also present with more crystallites, and thus do not appear to be an artifact of the single crystal simulations. The shape of the curve is an indication of the sensitivity to, for example, B_1 inhomogeneity or a mis-set match condition, where a flatter curve should mean less sensitive condition. There is a significant zero-frequency in the simulated spectra indicating that some fraction of the polarization is not recoupled, although this component is lost when more crystallites are included in the calculation when the calculations are repeated with crystal averaging, as seen in **Supplementary Figure S2**.

Dipole recoupling efficiency curves are generated for all 22 *R*-elements with 101 applied fields for each of the 203 symmetries. The highest scaling factor for each symmetry and element pair are plotted in **Figure 3**. Here, the recoupling for the "m" and "μ" (space and spin component) quantum numbers are either correlated (orange X), or anticorrelated (blue +), where this correlation indicates a second order dependency on frequency offset from the carrier (correlated) or on a field mis-set as in B_1 inhomogeneity (anticorrelated), where the frequency offset dependence is the more favorable deficiency.

In **Figure 3A**, there is a clear limit to the performance as a function of the applied field, where the maximum scaling rises quickly up to a maximum of about $\kappa = 0.27$ and $\omega_1 = 175$ kHz and then slowly reduces as the applied field increases. There are many candidates with $\kappa > 0.225$ and $\omega_1 < 150$ kHz, which are suitable for further testing. When the scaling is plotted against the flip angle, we find that there are certain flip angles that are favored. The first local maximum is at zero, indicating that at least some sequences do not improve using composite pulses. There are local maxima at about 15° and 35°, but the global maximum occurs when there is a flip angle of about 125°, and then approximately every 90° after.

The SIMPSON input files used to evaluate the pulse sequences and Python scripts to process the simulated datasets (i.e., find the maximum in a simple spectrum and report the scaling factor of a 2D simulation) are available online (see *Materials and Methods*).

The 1a element was found to produce the largest scaling factor (or is tied for the largest) for any given symmetry sequence. Additionally, the 1a element has the lowest requirements of any element that performs similarly. That is, the 1a element consistently produces the highest scaling factor for a given symmetry for the least applied field amongst all other competitive R-elements. Therefore, we only considered the 1a element further. The plot of the scaling factors for only the 1a elements as a function applied field and flip angle can be found in **Supplementary Figure S3**. The intensity of the applied field during the recoupling period is the largest concern, however the maximum duration is short so we felt that we could push the limits of the probe and chose an upper limit of about 150 kHz. Since compensating for poor B_1 homogeneity was the motivation to use compound pulses, we devised a fast test to show the B_1, or mis-set, dependence. The 2-spin simulation is run with the optimum flip angle, but the applied field is multiplied by 0.9 or 1.1. About half of the sequences respond strongly and move by hundreds of Hz, while the other half move less than ~50 Hz. The 0.9 mis-set spectra generally shift more than the 1.1. We found that the mis-set dependence depends on the specific components that are recoupled, specifically if the space component "m" and the spin component "μ" are correlated (both positive or both negative within the same allowed set), the mis-set dependence is typically small. This is due to a "pulse" term in the second order Hamiltonian which is allowed in the anticorrelated set of symmetries and disallowed in the correlated set. The "pulse" term is replaced by a second order frequency offset term in the correlated set of symmetries, which can be demonstrated in the simulations by observing the dependence of the scaling when introducing a frequency offset. This is not an absolute rule, though, since some anti-correlated symmetries are not greatly affected by mis-set, such as $R14_{8}^{5}$.

The criteria to choose a symmetry sequence are to find a sequence that works and will not damage either the sample or the hardware. Initially, the criteria were that the maximum applied field should be less than ~150 kHz, we should use the "1a" element for homonuclear recoupling, the scaling factor should be as large as possible, the spectral width should be large enough that the spectrum does not fold, and the echo time should be short. Ideally, the allowed "m" and "μ" components should be the same sign and the R-element should not stray too far from those previously devised. On close inspection, one notices a small gap at around 150 kHz, after which the scaling factors no longer greatly improve. The sequence immediately after this gap is $R20_{9}^{8}(124_{(0)}304_{(180)})$. Reducing the flip angle to 115° [making $R20_{9}^{8}(115_{(0)}295_{(180)})$] reduces the applied field to 152 kHz, but does not affect the scaling factor [the scaling is the same down to a 105° flip angle ($B_1 = 145$ kHz)]. $R20_{9}^{8}(115_{(0)}295_{(180)})$ thus fulfills all our desired traits. There are many candidate sequences with good scaling, and less demanding match conditions that meet the criteria, but were not tested experimentally. However, the sequences $R22_{21}^{9}$ $(300_{(0)},480_{(180)})$ and $R14_{22}^{5}$ $(460_{(0)},640_{(180)})$ were chosen to test the robustness of the simulations approach, since the unusual flip-angles in the composite pulses change the scaling factors from almost 0 to above 0.2.

This same *in silico* analysis can, of course, be made under 100 kHz spinning (**Supplementary Figure S4**). 310 gamma encoded candidate symmetry sequences were generated where $N = 10$ through 42, $n = 1$ through 37, and $\nu \leq N/2$ and the base match field is limited so that $\omega_1 < 200$ kHz. The recoupling is less efficient at 100 kHz spinning, as the curve equivalent to the one

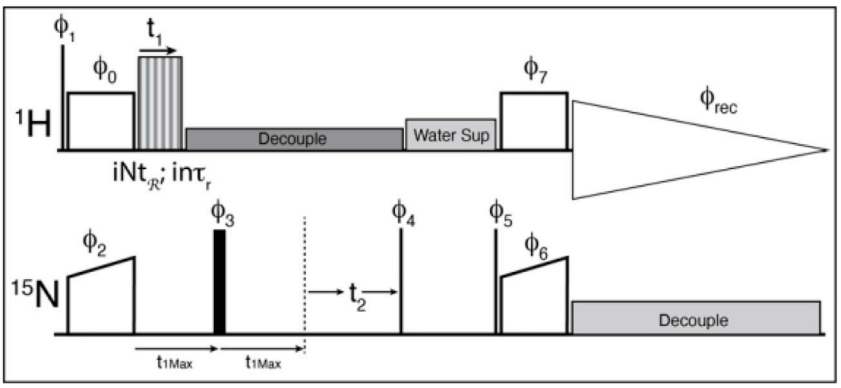


FIGURE 4 | The pulse sequence used for the experiments. Thin lines represent 90° pulses, and thick lines represent 180° pulses, boxes indicate cross polarization spinlock pulses. The sequence is modified from the usual solid-state NMR CP-HSQC by the addition of an echo period, and the application of the recoupling period. ϕ_0 = 1; ϕ_1 = 0 0 2 2; ϕ_2 = 0*4 2*4; ϕ_3 = 1*8 2*8; ϕ_4 = 3 1; ϕ_5 = 1; ϕ_6 = 0; ϕ_7 = 1; ϕ_{rec} = 1 3 3 1 3 1 1 3 3 1 1 3 1 3 3 1 where X*4 indicates that phase X is repeated 4 times. The phase shifts for the recoupling experiment are determined by the symmetry sequence and must be adjusted by hand.

shown in **Figure 3A** is shifted to higher match conditions at 100 kHz spinning (**Supplementary Figure S4A**). The dependence on the flip angle is the same as at 60 kHz spinning, where there are local maxima at ~15°, 35°, 125°, and then approximately every 90° afterwards (**Supplementary Figure S4B**). Still, there are several candidate sequences with scaling between κ = 0.15 and κ = 0.20 with relevant match conditions. If we limit the search to the "1a" element, with a match condition of less than ω_1 = 150 kHz we find 8 candidates, $R22_{16}^{1}$(θ = 104), $R26_{19}^{1}$(θ = 105), $R30_{22}^{1}$(θ = 106), $R32_{23}^{2}$(θ = 103), $R34_{25}^{1}$(θ = 104), $R38_{28}^{1}$(θ = 102), $R40_{29}^{2}$(θ = 104), and $R42_{31}^{1}$ (θ = 97).

EXPERIMENTAL VERIFICATION

A cross-polarization-based ^{1}H-detected ^{1}H-^{15}N correlation experiment was converted into a 3D experiment, where the third dimension consists of a constant time echo period on the low-gamma frequency (**Figure 4**). The recoupling time increases to create the third, separated local field dimension. The echo period (τ_{echo}) is calculated according to the chosen symmetry and data sampling by the following equation:

$$\tau_{\text{echo}} = 2 \times n\tau_r \times k \times (\text{points} - 1)$$

Where "*n*" is the space winding number from the symmetry sequence, "τ_r" is the rotor period, "*k*" is the number of repeated symmetry sequences (usually 1) to better fit the spectral width, and "points-1" is the zero-time-point inclusive number of points. We have chosen to apply the recoupling sequence only during the first half of the echo, and a standard decoupling pulse train for the rest of the echo and chemical shift dimensions. That is, the recoupling sequence increments up to the echo pulse, and the echo time is constant throughout the experiment. This is not the ideal case since the echo period should ideally be as short as possible so as little signal is lost as possible. However, the logic needed to loop the symmetry elements before, after and during the π pulse, to properly invert the phases of the symmetry pulses after the echo, and to maintain the proper timings in the Bruker scripting language was too cumbersome, so a simple echo was settled on since the Nitrogen T_2^* is quite long for these samples. It might be possible to use "compound pulse decoupling" pulses to simplify the logic, but we were unsuccessful in our attempts. Alternatively, the π pulse could be used for chemical shift evolution in a constant time evolution experiment, but this will reduce the sensitivity further.

Three recoupling sequences with good theoretical scaling factors and appropriate match conditions were chosen to validate our approach, and to determine which candidate scheme is the most promising. These sequences: $R20_9^8$ ($115_{(0)}$,$295_{(180)}$); $R22_{21}^{9}$ ($300_{(0)}$,$480_{(180)}$); and $R14_{22}^{5}$ ($460_{(0)}$,$640_{(180)}$) were tested with both their π-pulse version and the numerically optimized sequence elements (**Supplementary Figure S5**). The sequences were chosen partially because of the diversity of the flip angle, applied field, and the difference in scaling between standard and optimized sequences. The *in-silico* performance of these three sequences, along with $R14_8^5$ ($115_{(0)}$,$295_{(180)}$) are summarized in **Table 2**. The experimental performance closely follows the *in-silico*

TABLE 2 | Scaling factor and match conditions for Selected Symmetry sequences.

Symmetry	κ(π)	$\nu_1(\pi)$ kHz	$\nu_1(\theta)$ kHz	κ(θ)	θ in $\theta_{(0)}$–(θ + 180) $_{(180)}$
$R22_{21}^{9}$	0.0223	31.4	136	0.2567	300°
$R14_{22}^{5}$	0.0075	19.1	114	0.2354	460°
$R20_9^8$	0.1354	66.7	152	0.2682	115°
$R14_8^5$	0.0674	52.5	119.5	0.2314	115°

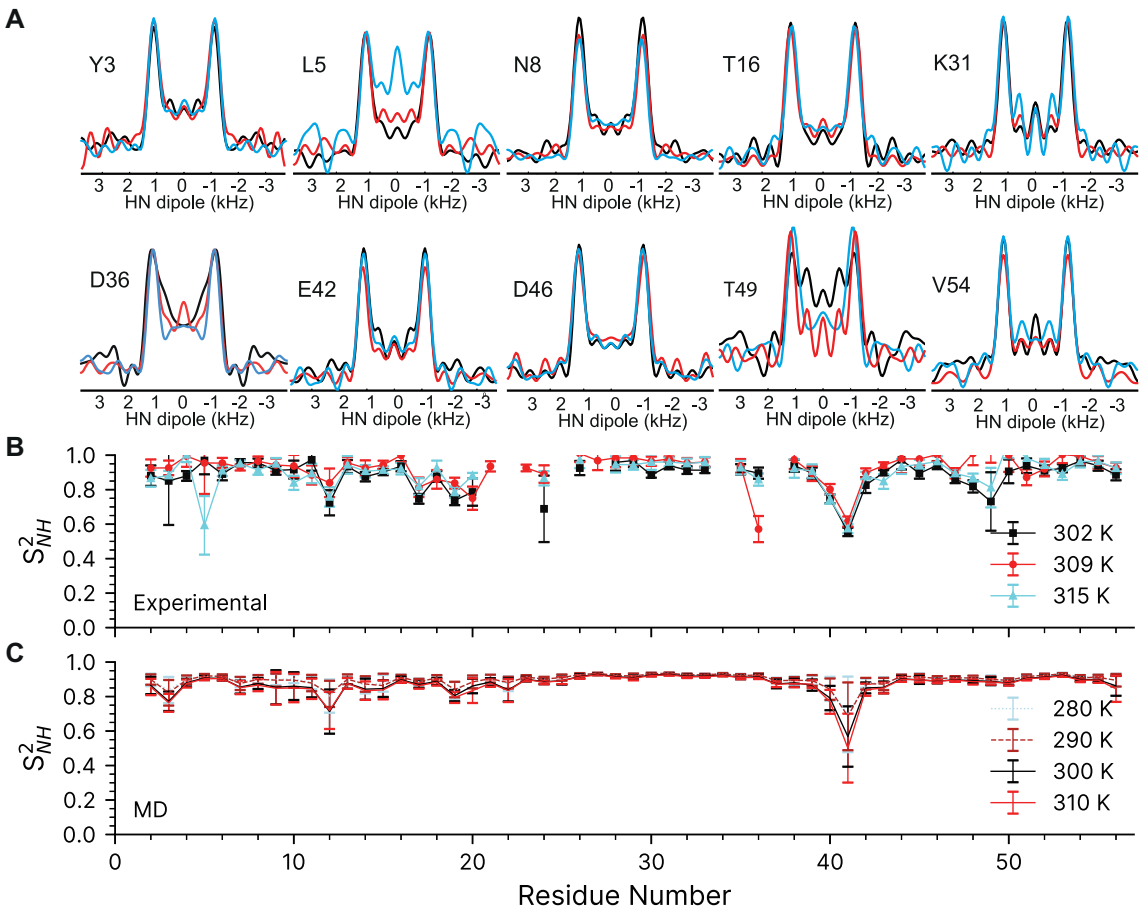


FIGURE 5 | ^{15}N-^{1}H dipolar couplings as a function of temperature in crystalline GB1. **(A)** Representative dipole-dipole recoupling spectra using the R14$_8^5$ (115$_{(0)}$,295$_{(180)}$) pulse sequence at ~302 K, ~309 and ~315 K at 60 kHz spinning. **(B)** S^2_{NH} obtained from the measured dipolar couplings, error bars are drawn at 1σ. **(C)** S^2_{NH} obtained from variable temperature molecular dynamics simulations. MD derived order parameters were extracted from 400 ns molecular dynamics simulations of a 3 × 3 × 3 supercell of GB1 containing 108 monomers.

performance as demonstrated by strong recoupling with the variable flip angle pulses which lends credence to the approach, and the robustness of symmetry theory in general.

While the R20$_9^8$ (115$_{(0)}$,295$_{(180)}$) seems to be a great option, its match condition is considerable at ν_1 = 152 kHz. During the power calibration with long nutation experiments, the signal disappeared in a few ms at the fields needed for R20$_9^8$ (115$_{(0)}$,295$_{(180)}$). The probe detuned indicating that such high match conditions would likely damage the sample or the probe, however fields up to about 130 kHz were long lived. The R14$_{22}^5$ (460$_{(0)}$,640$_{(180)}$) was also tested, but did not produce any recoupling, so the R22$_{21}^9$ (300$_{(0)}$,480$_{(180)}$) sequence was also thrown out since both have quite long echo times, and quite large flip angles. Other candidate sequences were identified from the scaling curve that have more conventional *R*-elements and lower match conditions. Amongst a handful of candidate sequences, the R14$_8^5$ (115$_{(0)}$,295$_{(180)}$) sequence was the first one that worked (it was also the first one tested). The R14$_8^5$ (115$_{(0)}$,295$_{(180)}$) has a lower match condition than the R20$_9^8$(115$_{(0)}$,295$_{(180)}$), the R-element is not far from a canonical R-element, and it seems to have a high tolerance for field mis-set in numerical simulations. The match condition (ν_1 = 119.5 kHz) for R14$_8^5$ (115$_{(0)}$,295$_{(180)}$) is very near the $\nu_1/2p = 2\omega_r$ rotary resonance condition, although this does not seem detrimental to the quality of the data in the microcrystalline sample. If the match is a greater concern, the *R*-element could be adjusted for a lower match condition as it was for the R20$_9^8$ (115$_{(0)}$,295$_{(180)}$). The scaling factor is the same for all "1a" *R*-elements between θ = 100° and θ = 115°, where ν_1 ranges from 111.8 to 119.5 kHz. The ability to turn down the power was not appreciated at the time the experiment was conducted.

Temperature Dependent Order Parameters

To validate the designed experiments, we have first applied them to a favorable model sample of crystalline 100% H_2O [U-^{2}H,^{13}C,^{15}N]GB1, which means that the protein is uniformly ^{13}C and ^{15}N labelled and perdeuterated with only exchangeable protons being reintroduced at 100%. We used R14$_8^5$ (115$_{(0)}$,295$_{(180)}$) sequence to measure ^{15}N-^{1}H order parameters at three different temperatures, nominally 263.2,

273.2, and 283.2 K, which correspond to sample temperatures of ~302 K, ~309 and ~315 K (larger temperature differences are difficult to obtain on our 1.3 mm probes at 60 kHz spinning; see *Materials and Methods*). A representative ^{1}H-^{15}N correlation spectrum taken from the first plane of the 3D can found in the **Supplementary Figure S9**.

The quality of the recoupled line-shapes for all temperatures is excellent as evidenced in **Figure 5A**. There is surprisingly little intensity at the zero-frequency, and the dipole line-shape is clear and strong. As expected, there is little difference in the experimentally observed (fitted) coupling over the explored temperature range and the obtained ^{15}N-^{1}H order parameters are generally very similar for all three measurements (see **Supplementary Table S7**). We only observe a noticeable change of ^{15}N-^{1}H dipole-dipole coupling (from 11.1 to 10.2 kHz) as a function of temperature for L5, D36, and T49 which are low intensity peaks in the 2D fingerprint spectrum. In the ~309 K data, the order parameter for D36 is spuriously low for reasons which are not clear (See **Supplementary Figure S8**). The precision of the experiment is very good: the 1σ standard deviation in the dipole couplings is around ±220 Hz, with some down to ±100 Hz, and outliers ranging from ±500 to 1500 Hz. These errors are comparable to previous symmetry-based methods on the same protein (Franks et al., 2005; Franks et al., 2006b). However, the precision may be overstated due to poor noise estimation in the Monte Carlo error analysis. Monte Carlo error analysis first finds the best fit for the experimental data to a simulation, here the time-domain dipole-dipole recoupling curve. Noise is then added to each point of the experimental data and the best fit is found again, and the values saved. Noise is added to the original experimental data several times (here, 5,000 times) to estimate the amount of spread in the simulation values found in the experimental data. The method we have used to add the noise could be improved. First, most peaks have very similar initial intensities, so a constant noise value was used for all peaks in the Monte Carlo analysis, which was ±7.5% of the total. All trajectories were normalized to 1 during the integrations, which results in an undesirable loss of information regarding the intensity. The result is that a resonance that is 100 intensity units high will have noise ranging from −7.5 to +7.5 added, while a peak that is only 10 Intensity units high will only have noise added that ranges from −0.75 to +0.75 units. The noise estimate works well for peaks with a typical intensity, and may even be larger than necessary, but the error analysis fails for peaks with poor intensity. A typical Monte Carlo fit is shown for residue K28 in **Supplementary Figure S6**. Similar figures for all fits can be found in their corresponding datasets in the online materials. Those residues with worse sensitivity will have spuriously good fits, such as found for the K13 peak in the GB1+IgG complex (**Supplementary Figure S8**).

The ^{15}N-^{1}H dipolar order parameters shown in **Figure 5B** fit well with those reported previously (Franks et al., 2005; Franks et al., 2006b). The ^{15}N-^{1}H dipolar order parameters are generally around 0.9 (order parameters, S^2, take values between 0 and 1, which mean unrestricted motion and no motion respectively) with an occasional dip to around 0.8 near the loops. The largest amplitudes of motions are observed for residue G41, which is in a loop between the alpha helix and beta strand 3.

We have also compared the experimentally determined ^{15}N-^{1}H order parameters to those obtained from molecular dynamics simulations performed in the 280–310 K temperature range (**Figure 5C**). The simulated rates are in general good agreement with the experimental ones (see **Supplementary Figure S7**). As before (Busi et al., 2018), the MD simulations predict that there should be little change in the order parameters of the expected temperature range.

ORDER PARAMETERS OF THE GB1-IGG COMPLEX

The complex of 100% H_2O [U-^{2}H,^{13}C,^{15}N]GB1 with IgG is a much more challenging sample compared to crystalline GB1 to apply the described methods. As a precipitate it is more heterogenous, >90% of the sample volume is taken by the antibody resulting in lower sensitivity and GB1 in the complex exhibits much more pronounced slow motions in the microsecond range. (Lamley et al., 2014; Lamley et al., 2015a; Öster et al., 2019). A paramagnetic doping agent, 2 mM Gd (DTPA-BMA), was added to speed up the measurements by reducing the required relaxation delays (Linser et al., 2007), although in this particular application sample heating is a large concern so the recovery delay remains quite long at 1s. The ^{1}H-^{15}N correlation spectrum taken from the first plane of the 3D can found in the **Supplementary Figure S9**.

Despite the more challenging nature of the sample, the GB1 in the complex still produces high-quality dipole-recoupling spectra, as seen in **Figure 6A**. There is a significant zero-frequency component in most of these spectra, and the sensitivity is generally worse (especially for T16). The origin of the zero-frequency component is possibly due to the larger amount of ^{1}H atoms in the sample, less efficient ^{1}H-^{1}H homo-decoupling, increased dynamics, sample heating, and/or probe detuning. The determined ^{15}N-^{1}H order parameters are generally lower in the GB1 complex (average S^2_{NH} ~0.7) than in crystalline GB1 (average S^2_{NH} ~0.9). If one uses expression for order parameter in diffusion in a cone model this difference corresponds on average to ~23° additional motional amplitude for most residues. This provides further support for presence of a microsecond range overall motion of GB1 in the complex with IgG as proposed previously (Lamley et al., 2015a).

CONCLUSION

We have presented a method to apply symmetry-based recoupling theory to fast MAS experiments using variable flip angle compound pulses. The method generates many candidate sequences that have a reasonably high scaling factor and applicable match conditions. Being able to apply symmetry principles under fast-MAS makes SLF methods applicable to faster spinning and the other benefits that usually comes with

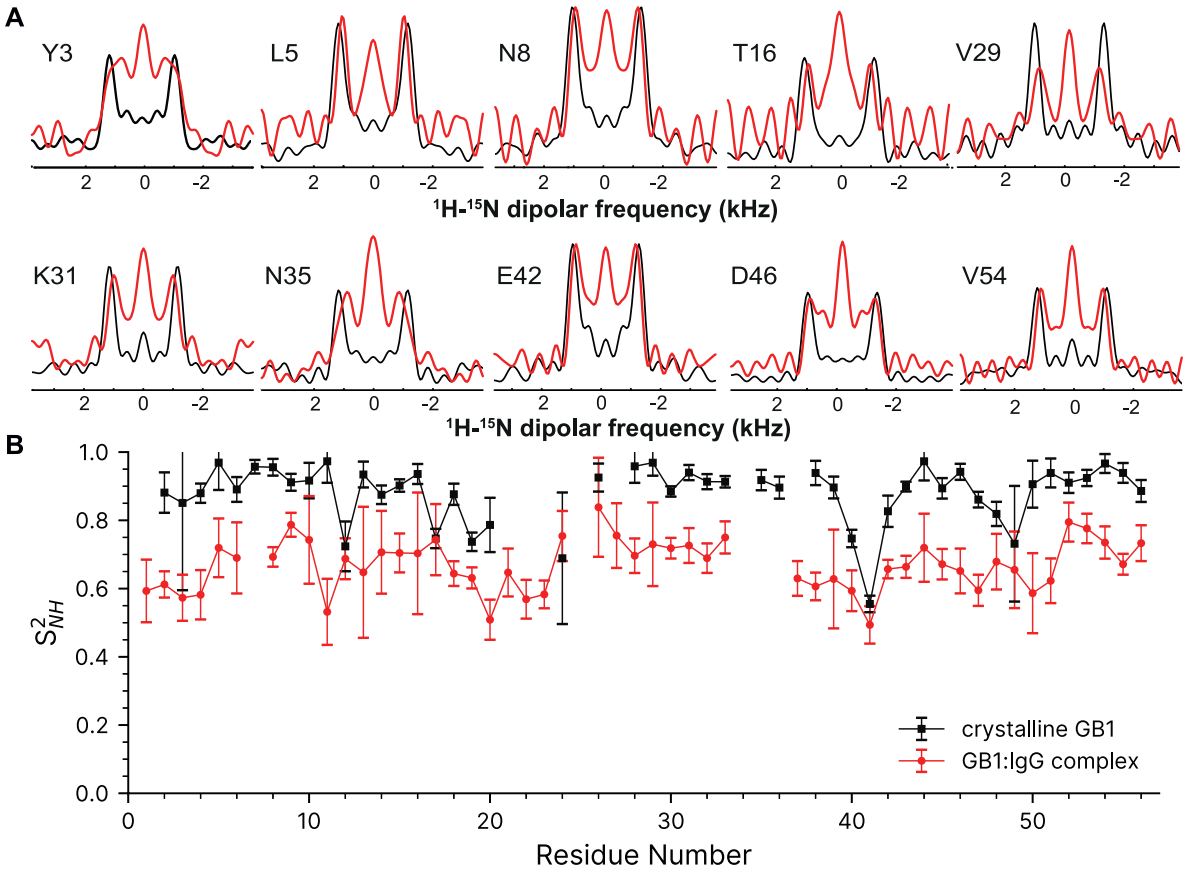


FIGURE 6 | ^{15}N-^{1}H dipolar order parameter data for GB1 in a crystal (black line) and in a precipitated complex with IgG (red line). **(A)** Overlay of representative dipole recoupling spectra. **(B)** ^{15}N-^{1}H order parameters as a function of residue, error bars are drawn at 1σ. Both data sets were recorded at the nominal temperature of 263.2 K corresponding to sample temperature of ~302 K at 60 kHz spinning frequency.

it. The heteronuclear dipole-dipole coupling can be measured site-specifically in microcrystalline GB1 and in the GB1-IgG complex. The order parameters found using the variable flip angle method are consistent with previous datasets and molecular dynamics in the microcrystalline sample. The GB1 in the complex shows both localized differences in dynamics and global increase of cumulative motional amplitudes.

MATERIALS AND METHODS

Sample Preparation

Uniformly (^{2}H, ^{13}C, ^{15}N) GB1 was produced as described previously (Franks et al., 2005; Zhou et al., 2007b). After production in ^{2}H buffer with ^{13}C glucose and ^{15}N NH_4Cl, the protein is placed in ^{1}H containing buffer, and heated so that the exchangeable ^{1}H sites are ^{1}H labelled. The protein is then either crystallized or incubated with natural abundance IgG in an equimolar ratio (Lamley et al., 2014). All back-bone amide sites on the GB1 molecules are thus labeled, but the strong ^{1}H dipole coupling network is disrupted locally by the ^{2}H labelling of the sidechain carbons. The buffer of the IgG-GB1 complex contained 2 mM of paramagnetic Gd (DTPA-BMA).

NMR Spectroscopy

All experiments were performed on either a Bruker Avance III spectrometer at 700.13 MHz ^{1}H Larmor frequency or a Bruker Avance II spectrometer or at a 599.4 MHz ^{1}H Larmor frequency. A Bruker 1.3 mm HCN Probe operating in HCN triple resonance mode with a sample spinning rate of 60 kHz ± 3 Hz was used with both instruments. 1,200 L/h of cooling gas was used at the nominal temperatures of 263.2, 273.2, and 283.2 K. The nutation frequencies for the 90° pulses were calibrated so that the hard pulses for ^{1}H were 2.1 μs (ν_1 = 120 kHz); ^{13}C, 2.5 μs (ν_1 = 100 kHz); and ^{15}N, 3.25 μs (ν_1 = 77 kHz). The ^{1}H carrier radiofrequency (RF) was centred on the H_2O signal (~4.7 ppm), ^{15}N at 120 ppm, and ^{13}C at 100 ppm. Heteronuclear ^{1}H decoupling (~10 kHz SPINAL-16) (Fung et al., 2000) was used during the indirect chemical shift dimension, and during the echo period when the recoupling sequence was not being applied, and approximately 10 kHz WALTZ-64 (Zhou et al., 2007c) ^{15}N decoupling was used during ^{1}H acquisition. The States-TPPI method was employed

for quadrature detection in the indirect chemical shift dimension (Marion et al., 1989) and only the real portion of the dipole-coupling dimension is acquired. The MISSISSIPPI (Zhou and Rienstra, 2008) solvent suppression scheme was applied with a spinlock field of ~30 kHz for four 15 ms intervals before detection. The symmetry match condition was calibrated to the theoretical value by varying the applied power in a 2D, nitrogen edited, ^{1}H nutation experiment until the nutation experiment is within 1 Hz.

The total experiment time for each temperature point of the crystalline GB1 was 16 h each ^{1}H free induction decay was acquired for 40 ms with a spectral width of 30 ppm with 32 coadded transients. The ^{15}N dimension for the microcrystalline protein were acquired with 80 rows with a dwell of 333.33 μs for a total of 13.3 ms in the indirect dimensions. The $R14_8^5$ ($115_{(0)}$,$295_{(180)}$) dimension was acquired for 15 real points with an increment of $8^*\tau_r$ = 133.33 μs for a total of 1.87 ms (3.73 ms total echo time). The recovery delay was 1.5 s.

The spectrum of the GB1-IgG complex was collected in four blocks of 34.1 h each which were later summed together, for a total of 5 days 16.5 h. There was 1,200 L/h of variable temperature gas flow at the nominal temperature of 263.2 K. Each ^{1}H free induction decay was acquired for 30 ms with a spectral width of 30 ppm in four blocks of 128 coadded transients (512 total), the ^{15}N dimension for the microcrystalline protein were acquired with 64 rows with a dwell of 333.33 μs, with a spectral width of 42 ppm (^{15}N) for a total of 10.7 ms in the indirect dimensions, the $R14_8^5$ ($115_{(0)}$,$295_{(180)}$) dimension was acquired for 15 real points with an increment of $8^*\tau_r$ = 133.33 μs for a total of 1.87 ms (3.73 ms total echo time) with a relaxation delay of 1 s All 3D data was processed using NMRPipe and the four blocks were added using NMRPipe (Delaglio et al., 1995). The dipole-recoupling dimension was not Fourier-transformed in NMRPipe so that peak volumes could be extracted. The Fourier-transform of the dipole recoupling dimension was performed on the peak volumes extracted by NMRPipe by a the fast fourier transform routine found in python's numpy package. The imaginary components of the trajectories were filled with zeroes, and the trajectory was zero filled to 512 points. The 2D datasets with a dipole recoupling dimension (**Supplementary Figure S5**) were processed in Topspin with by performing a Hilbert-transform to fill the imaginary portion of the dipole trajectory before Fourier transforming the dipole recoupled dimension. (See **Supplementary Material**: BrukerMacros/2DHN).

External KBr (Thurber and Tycko, 2009) and neat methanol (Ammann et al., 1982) were used as external standards to calibrate the temperature. The samples did not have adequate resolution to unambiguously identify the bulk water signal from the isopropanol and methyl-pentane-diol OH signals, which precluded temperature calibration by the chemical shift difference between water and DSS (Hoogen et al., 1988; Wishart et al., 1995). Temperatures derived using either the chemical shift or the T_1 of KBr (Thurber and Tycko, 2009) were not self-consistent on the 700 MHz instrument at either 10 or 60 kHz spinning. There is approximately 10°C difference between the two methods. The T_1 method indicates a 20°C difference across the nominal temperatures, where the chemical shift method indicates a difference of 13.6°C. Calibrating the temperature by the chemical shift difference in the ^{1}H spectrum of methanol (Ammann et al., 1982) indicates temperatures of 301.9 (263.2), 309.0 (273.2), and 314.9 K (283.2) or 28.7, 35.8, and 41.7°C under the experimental conditions, for a range of 12.9°C.

Data Processing, Fitting and Simulations

The GB1 resonances were identified in the 2D finger-print spectra from the literature and previous work (Franks et al., 2005; Zhou et al., 2007a). Peaks were integrated using NMRPipe (Delaglio et al., 1995), where the peak volumes were converted into dipole recoupling trajectories. The volume of the peak on the first plane is used to normalize the rest of the curve, so all curves range between ±1, and start with an intensity of 1.00. Overlapping resonances were fit, but these resonances were not included in the figures or analysis since the resonance could not be unambiguously identified.

A library of numerical simulations was created in SIMPSON 4.2.1 on an Apple MacBook Pro for use in the Monte-Carlo fitting routine. The simulation library was created using 251 dipole-dipole couplings ranging from 7,500 Hz to 12,500 Hz in steps of 20 Hz, the spin rate was 60 kHz, the calculation method was "direct", the crystal file was "zcw376," and 16 gamma angles were used. The time-domain trajectories and frequency domain spectra were saved as a 2D SIMPSON file.

5000 Monte Carlo steps were used for error analysis for all datasets only using the time domain SIMPSON library. The library was expanded using simple operations on the time-domain. The DC parameter is used to add a constant is to all data points (varied between −0.2 and +0.5). The scaling factor multiplies each point by a constant (varied from 0.9 to 1.1). Finally, relaxation is simulated by applying line-broadening for each trajectory, that is, the simulation is multiplied by a time dependent exponential function (from 0 Hz to 2,500 Hz). To Fourier transform the dipole trajectories, the imaginary time portion was filled with zeroes, and the trajectory was zero-filled to 128 points. For the crystalline GB1, most line shapes fit with a small, negative DC offset (−0.05), a scaling multiplier of 1.00, and less than 300 Hz of line-broadening. For the IgG-GB1 complex, there is a small positive DC offset (+0.14), a scale of 1.00, and approximately 600 Hz of line-broadening on average. The rigid limit for the N-H dipolar coupling is taken as 11,477.3 Hz to determine the Order Parameters, which corresponds to an N-H bond length of 1.02 Å.

Molecular Dynamics Simulations

A molecular dynamics trajectory for a 3 × 3 × 3 supercell of GB1 containing 108 monomers was computed using AMBER MD (Case et al., 2005; Doshi and Hamelberg, 2009; Maier et al., 2015; Tian et al., 2019). The coordinates of the X-ray structure of GB1 (PDB: 2gi9 Franks et al., 2006b) were taken as a starting conformation. To the supercell, 108 PO_4^{3-} counter ions were added. 12,852 explicit water molecules were added, followed by charge balancing with sodium ions giving an overall box size of 75.591 Å × 107.152 Å × 150.822 Å. The ff19SB (Tian et al., 2019) forcefield was used for the GB1 proteins, with OPC water (Izadi et al., 2014) and GAFF cocrystallites (Wang et al., 2004). After minimization, the system was replicated and heated to the temperatures indicated in the

figures (280, 290, 300, 310 K). The systems were then simulated for full 400 ns runs. For each, a 2 fs timestep was used with a cut-off of 11 Å for non-bonded interactions. Temperatures were maintained using a Langevin thermostat, and the SHAKE algorithm (Ryckaert et al., 1977) was applied to all bond lengths involving a hydrogen atom. Anisotropic pressure scaling was used with periodic boundary conditions.

Prior to processing, the Cα carbons between timesteps were aligned using cpptraj (Roe and Cheatham, 2013). Then, correlation functions for each N-H vector were calculated according to the iRED framework using cpptraj (Prompers and Brüschweiler, 2002). The median was calculated for each residue over all GB1s in the supercell for which order parameters could be extracted, and the error taken as twice the median absolute difference.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

WF designed, optimized, and performed the experiment, analyzed the simulated and experimental NMR data, and prepared the manuscript. BT performed and analyzed the molecular dynamics simulations. JT performed and analyzed NMR simulations. JL designed the experiment, analyzed the data, and prepared the manuscript.

FUNDING

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement 639907, which supported JL and WF. JL also acknowledges funding from EPSRC EP/L025906/1 and BBSRC Grant BB/R010218/1.

ACKNOWLEDGMENTS

BT thanks Bruker and the University of Warwick for PhD funding as part of the Warwick Centre for Doctoral Training in Analytical Science.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.791026/full#supplementary-material

REFERENCES

Alderson, T. R., and Kay, L. E. (2021). NMR Spectroscopy Captures the Essential Role of Dynamics in Regulating Biomolecular Function. *Cell* 184, 577–595. doi:10.1016/j.cell.2020.12.034

Ammann, C., Meier, P., and Merbach, A. (1982). A Simple Multinuclear NMR Thermometer. *J. Magn. Reson. (1969)* 46, 319. doi:10.1016/0022-2364(82)90147-0

Asami, S., and Reif, B. (2012). Assignment Strategies for Aliphatic Protons in the Solid-State in Randomly Protonated Proteins. *J. Biomol. NMR* 52, 31–39. doi:10.1007/s10858-011-9591-4

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., et al. (2021). Accurate Prediction of Protein Structures and Interactions Using a Three-Track Neural Network. *Science* 373, 871–876. doi:10.1126/science.abj8754

Bak, M., Rasmussen, J. T., and Nielsen, N. C. (2000). SIMPSON: A General Simulation Program for Solid-State NMR Spectroscopy. *J. Magn. Reson.* 147, 296–330. doi:10.1006/jmre.2000.2179

Barbet-Massin, E., Pell, A. J., Retel, J. S., Andreas, L. B., Jaudzems, K., Franks, W. T., et al. (2014). Rapid Proton-Detected NMR Assignment for Proteins with Fast Magic Angle Spinning. *J. Am. Chem. Soc.* 136, 12489–12497. doi:10.1021/ja507382j

Barbet-Massin, E., Ricagno, S., Lewandowski, J. R., Giorgetti, S., Bellotti, V., Bolognesi, M., et al. (2010). Fibrillar vs Crystalline Full-Length β-2-Microglobulin Studied by High-Resolution Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 132, 5556–5557. doi:10.1021/ja1002839

Bennett, A. E., Griffin, R. G., Ok, J. H., and Vega, S. (1992). Chemical Shift Correlation Spectroscopy in Rotating Solids: Radio Frequency-driven Dipolar Recoupling and Longitudinal Exchange. *J. Chem. Phys.* 96, 8624–8627. doi:10.1063/1.462267

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., et al. (2000). The Protein Data Bank. *Nucleic Acids Res.* 28, 235–242. doi:10.1093/nar/28.1.235

Brinkmann, A., and Levitt, M. H. (2001). Symmetry Principles in the Nuclear Magnetic Resonance of Spinning Solids: Heteronuclear Recoupling by Generalized Hartmann-Hahn Sequences. *J. Chem. Phys.* 115, 357–384. doi:10.1063/1.1377031

Burley, S. K., Bhikadiya, C., Bi, C., Bittrich, S., Chen, L., Crichlow, G. V., et al. (2021). RCSB Protein Data Bank: Powerful New Tools for Exploring 3D Structures of Biological Macromolecules for Basic and Applied Research and Education in Fundamental Biology, Biomedicine, Biotechnology, Bioengineering and Energy Sciences. *Nucleic Acids Res.* 49, D437. doi:10.1093/nar/gkaa1038

Busi, B., Yarava, J. R., Hofstetter, A., Salvi, N., Cala-De Paepe, D., Lewandowski, J. R., et al. (2018). Probing Protein Dynamics Using Multifield Variable Temperature NMR Relaxation and Molecular Dynamics Simulation. *J. Phys. Chem. B* 122, 9697–9702. doi:10.1021/acs.jpcb.8b08578

Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., et al. (2005). The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* 26, 1668–1688. doi:10.1002/jcc.20290

Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K., and Oschkinat, H. (2002). Structure of a Protein Determined by Solid-State Magic-Angle-Spinning NMR Spectroscopy. *Nature* 420, 98–102. doi:10.1038/nature01070

Chevelkov, V., Fink, U., and Reif, B. (2009). Accurate Determination of Order Parameters from $^{1}H,^{15}N$ Dipolar Couplings in MAS Solid-State NMR Experiments. *J. Am. Chem. Soc.* 131, 14018–14022. doi:10.1021/ja902649u

Chevelkov, V., Rehbein, K., Diehl, A., and Reif, B. (2006). Ultrahigh Resolution in Proton Solid-State NMR Spectroscopy at High Levels of Deuteration. *Angew. Chem. Int. Ed.* 45, 3878–3881. doi:10.1002/anie.200600328

Chevelkov, V., van Rossum, B. J., Castellani, F., Rehbein, K., Diehl, A., Hohwy, M., et al. (2003). H Detection in MAS Solid-State NMR Spectroscopy of

Biomacromolecules Employing Pulsed Field Gradients for Residual Solvent Suppression. *J. Am. Chem. Soc.* 125, 7788–7789. doi:10.1021/ja029354b

Chevelkov, V., Xue, Y., Linser, R., Skrynnikov, N. R., and Reif, B. (2010). Comparison of Solid-State Dipolar Couplings and Solution Relaxation Data Provides Insight into Protein Backbone Dynamics. *J. Am. Chem. Soc.* 132, 5015–5017. doi:10.1021/ja100645k

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *J. Biomol. NMR* 6, 277–293. doi:10.1007/BF00197809

Doshi, U., and Hamelberg, D. (2009). Reoptimization of the AMBER Force Field Parameters for Peptide Bond (Omega) Torsions Using Accelerated Molecular Dynamics. *J. Phys. Chem. B* 113, 16590–16595. doi:10.1021/jp907388m

Fage, C. D., Kosol, S., Jenner, M., Öster, C., Gallo, A., Kaniusaite, M., et al. (2021). Communication Breakdown: Dissecting the COM Interfaces between the Subunits of Nonribosomal Peptide Synthetases. *ACS Catal.* 11, 10802–10813. doi:10.1021/acscatal.1c02113

Franks, W. T., Wylie, B. J., Stellfox, S. A., and Rienstra, C. M. (2006). Backbone Conformational Constraints in a Microcrystalline U-^{15}N-Labeled Protein by 3D Dipolar-Shift Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 128, 3154–3155. doi:10.1021/ja058292x

Franks, W. T., Wylie, B. J., Stellfox, S. A., and Rienstra, C. M. (2006). Backbone Conformational Constraints in a Microcrystalline U-^{15}N-Labeled Protein by 3D Dipolar-Shift Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 128, 3154–3155. doi:10.1021/ja058292x

Franks, W. T., Zhou, D. H., Wylie, B. J., Money, B. G., Graesser, D. T., Frericks, H. L., et al. (2005). Magic-Angle Spinning Solid-State NMR Spectroscopy of the β1 Immunoglobulin Binding Domain of Protein G (GB1): ^{15}N and ^{13}C Chemical Shift Assignments and Conformational Analysis. *J. Am. Chem. Soc.* 127, 12291–12305. doi:10.1021/ja044497e

Frauenfelder, H., Sligar, S. G., and Wolynes, P. G. (1991). The Energy Landscapes and Motions of Proteins. *Science* 254, 1598–1603. doi:10.1126/science.1749933

Fung, B. M., Khitrin, A. K., and Ermolaev, K. (2000). An Improved Broadband Decoupling Sequence for Liquid Crystals and Solids. *J. Magn. Reson.* 142, 97–101. doi:10.1006/jmre.1999.1896

Gullion, T., and Schaefer, J. (1989). Detection of Weak Heteronuclear Dipolar Coupling by Rotational-Echo Double-Resonance Nuclear Magnetic Resonance. *Adv. Magn. Opt. Reson.* 13, 57–83. doi:10.1016/B978-0-12-025513-9.50009-4

Gullion, T., and Schaefer, J. (1989). Rotational-echo Double-Resonance NMR. *J. Magn. Reson. (1969)* 81, 196–200. doi:10.1016/0022-2364(89)90280-1

Herbst, C., Bellstedt, P., Görlach, M., and Ramachandran, R. (2015). MAS Solid State NMR of Proteins: Simultaneous ^{15}N-^{13}CA and ^{15}N-^{13}CO Dipolar Recoupling via Low-Power Symmetry-Based RF Pulse Schemes. *J. Biomol. NMR* 62, 7–15. doi:10.1007/s10858-015-9910-2

Herbst, C., Herbst, J., Leppert, J., Ohlenschläger, O., Görlach, M., and Ramachandran, R. (2011). Chemical Shift Correlation at High MAS Frequencies Employing Low-Power Symmetry-Based Mixing Schemes. *J. Biomol. NMR* 50, 277–284. doi:10.1007/s10858-011-9516-2

Hester, R. K., Ackerman, J. L., Neff, B. L., and Waugh, J. S. (1976). Separated Local Field Spectra in NMR: Determination of Structure of Solids. *Phys. Rev. Lett.* 36, 1081–1083. doi:10.1103/physrevlett.36.1081

Hiller, M., Krabben, L., Vinothkumar, K. R., Castellani, F., van Rossum, B.-J., Kühlbrandt, W., et al. (2005). Solid-state Magic-Angle Spinning NMR of Outer-Membrane Protein G from *Escherichia coli*. *Chembiochem* 6, 1679–1684. doi:10.1002/cbic.200500132

Hohwy, M., Jakobsen, H. J., Edén, M., Levitt, M. H., and Nielsen, N. C. (1998). Broadband Dipolar Recoupling in the Nuclear Magnetic Resonance of Rotating Solids: A Compensated C7 Pulse Sequence. *J. Chem. Phys.* 108, 2686–2694. doi:10.1063/1.475661

Hohwy, M., Jaroniec, C. P., Reif, B., Rienstra, C. M., and Griffin, R. G. (2000). Local Structure and Relaxation in Solid-State NMR: Accurate Measurement of Amide N–H Bond Lengths and H–N–H Bond Angles. *J. Am. Chem. Soc.* 122, 3218–3219. doi:10.1021/ja9913737

Hoogen, Y. T., Treurniet, S. J., Roelen, H. C. P. F., Vroom, E., Marel, G. A., Boom, J. H., et al. (1988). Conformational Analysis of the Tetranucleotides m62A-m62A-U-m62A (m62A = N6-Dimethyladenosine) and U-m62A-U-m62A and of the Hybrid dA-r(U-A). A One- and Two-Dimensional NMR Study. *Eur. J. Biochem.* 171, 155–162. doi:10.1111/j.1432-1033.1988.tb13771.x

Hou, G., Byeon, I.-J. L., Ahn, J., Gronenborn, A. M., and Polenova, T. (2011). ^{1}H-^{13}C/^{1}H-^{15}N Heteronuclear Dipolar Recoupling by R-Symmetry Sequences under Fast Magic Angle Spinning for Dynamics Analysis of Biological and Organic Solids. *J. Am. Chem. Soc.* 133, 18646–18655. doi:10.1021/ja203771a

Izadi, S., Anandakrishnan, R., and Onufriev, A. V. (2014). Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* 5, 3863–3871. doi:10.1021/jz501780a

Jenner, M., Kosol, S., Griffiths, D., Prasongpholchai, P., Manzi, L., Barrow, A. S., et al. (2018). Mechanism of Intersubunit Ketosynthase-Dehydratase Interaction in Polyketide Synthases. *Nat. Chem. Biol.* 14, 270–275. doi:10.1038/nchembio.2549

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* 596, 583–589. doi:10.1038/s41586-021-03819-2

Knight, M. J., Pell, A. J., Bertini, I., Felli, I. C., Gonnelli, L., Pierattelli, R., et al. (2012). Structure and Backbone Dynamics of a Microcrystalline Metalloprotein by Solid-State NMR. *Proc. Natl. Acad. Sci.* 109, 11095–11100. doi:10.1073/pnas.1204515109

Koshland, D. E. (1958). Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc. Natl. Acad. Sci.* 44, 98–104. doi:10.1073/pnas.44.2.98

Kosol, S., Gallo, A., Griffiths, D., Valentic, T. R., Masschelein, J., Jenner, M., et al. (2019). Structural Basis for Chain Release from the Enacyloxin Polyketide Synthase. *Nat. Chem.* 11, 913–923. doi:10.1038/s41557-019-0335-5

Kovermann, M., Rogne, P., and Wolf-Watz, M. (2016). Protein Dynamics and Function from Solution State NMR Spectroscopy. *Quart. Rev. Biophys.* 49, e6. doi:10.1017/s0033583516000019

Lamley, J. M., Iuga, D., Öster, C., Sass, H.-J., Rogowski, M., Oss, A., et al. (2014). Solid-state NMR of a Protein in a Precipitated Complex with a Full-Length Antibody. *J. Am. Chem. Soc.* 136, 16800–16806. doi:10.1021/ja5069992

Lamley, J. M., Lougher, M. J., Sass, H. J., Rogowski, M., Grzesiek, S., and Lewandowski, J. R. (2015). Unraveling the Complexity of Protein Backbone Dynamics with Combined ^{13}C and ^{15}N Solid-State NMR Relaxation Measurements. *Phys. Chem. Chem. Phys.* 17, 21997–22008. doi:10.1039/c5cp03484a

Lamley, J. M., Öster, C., Stevens, R. A., and Lewandowski, J. R. (2015). Intermolecular Interactions and Protein Dynamics by Solid-State NMR Spectroscopy. *Angew. Chem. Int. Ed.* 54, 15374–15378. doi:10.1002/anie.201509168

Levitt, M. H. (1982). Symmetrical Composite Pulse Sequences for NMR Population Inversion. II. Compensation of Resonance Offset. *J. Magn. Reson. (1969)* 50, 95–110. doi:10.1016/0022-2364(82)90035-x

Levitt, M. H. (2002). Symmetry-Based Pulse Sequences in Magic-Angle Spinning Solid-State NMR. *Encyclopedia Magn. Reson.* 9, 165–196.

Lewandowski, J. R. (2013). Advances in Solid-State Relaxation Methodology for Probing Site-specific Protein Dynamics. *Acc. Chem. Res.* 46, 2018–2027. doi:10.1021/ar300334g

Lewandowski, J. R., Halse, M. E., Blackledge, M., and Emsley, L. (2015). Direct Observation of Hierarchical Protein Dynamics. *Science* 348, 578–581. doi:10.1126/science.aaa6111

Lewandowski, J. R., Sass, H. J., Grzesiek, S., Blackledge, M., and Emsley, L. (2011). Site-specific Measurement of Slow Motions in Proteins. *J. Am. Chem. Soc.* 133, 16762–16765. doi:10.1021/ja206815h

Liang, L., Ji, Y., Zhao, Z., Quinn, C. M., Han, X., Bao, X., et al. (2021). Accurate Heteronuclear Distance Measurements at All Magic-Angle Spinning Frequencies in Solid-State NMR Spectroscopy. *Chem. Sci.* 12, 11554–11564. doi:10.1039/d1sc03194e

Linser, R., Chevelkov, V., Diehl, A., and Reif, B. (2007). Sensitivity Enhancement Using Paramagnetic Relaxation in MAS Solid-State NMR of Perdeuterated Proteins. *J. Magn. Reson.* 189, 209–216. doi:10.1016/j.jmr.2007.09.007

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theor. Comput.* 11, 3696–3713. doi:10.1021/acs.jctc.5b00255

Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989). Rapid Recording of 2D NMR Spectra without Phase Cycling. Application to the Study of Hydrogen Exchange in Proteins. *J. Magn. Reson. (1969)* 85, 393–399. doi:10.1016/0022-2364(89)90152-2

Nielsen, A. B., Bjerring, M., Nielsen, J. T., and Nielsen, N. C. (2009). Symmetry-based Dipolar Recoupling by Optimal Control: Band-Selective Experiments for

Assignment of Solid-State NMR Spectra of Proteins. *J. Chem. Phys.* 131, 025101. doi:10.1063/1.3157737

Öster, C., Kosol, S., and Lewandowski, J. R. (2019). Quantifying Microsecond Exchange in Large Protein Complexes with Accelerated Relaxation Dispersion Experiments in the Solid State. *Sci. Rep.* 9, 11082. doi:10.1038/s41598-019-47507-8

Palmer, A. G. (2004). NMR Characterization of the Dynamics of Biomacromolecules. *Chem. Rev.* 104, 3623–3640. doi:10.1021/cr030413t

Paluch, P., Pawlak, T., Amoureux, J.-P., and Potrzebowski, M. J. (2013). Simple and Accurate Determination of X-H Distances under Ultra-fast MAS NMR. *J. Magn. Reson.* 233, 56–63. doi:10.1016/j.jmr.2013.05.005

Paluch, P., Pawlak, T., Ławniczak, K., Trébosc, J., Lafon, O., Amoureux, J.-P., et al. (2018). Simple and Robust Study of Backbone Dynamics of Crystalline Proteins Employing 1H-15N Dipolar Coupling Dispersion. *J. Phys. Chem. B* 122, 8146–8156. doi:10.1021/acs.jpcb.8b04557

Penzel, S., Oss, A., Org, M.-L., Samoson, A., Böckmann, A., Ernst, M., et al. (2019). Spinning Faster: Protein NMR at MAS Frequencies up to 126 kHz. *J. Biomol. NMR* 73, 19–29. doi:10.1007/s10858-018-0219-9

Prompers, J. J., and Brüschweiler, R. (2002). General Framework for Studying the Dynamics of Folded and Nonfolded Proteins by NMR Relaxation Spectroscopy and MD Simulation. *J. Am. Chem. Soc.* 124, 4522–4534. doi:10.1021/ja012750u

Retel, J. S., Nieuwkoop, A. J., Hiller, M., Higman, V. A., Barbet-Massin, E., Stanek, J., et al. (2017). Structure of Outer Membrane Protein G in Lipid Bilayers. *Nat. Commun.* 8, 2073. doi:10.1038/s41467-017-02228-2

Roe, D. R., and Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theor. Comput.* 9, 3084–3095. doi:10.1021/ct400341p

Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977). Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* 23, 327–341. doi:10.1016/0021-9991(77)90098-5

Schanda, P., and Ernst, M. (2016). Studying Dynamics by Magic-Angle Spinning Solid-State NMR Spectroscopy: Principles and Applications to Biomolecules. *Prog. Nucl. Magn. Reson. Spectrosc.* 96, 1–46. doi:10.1016/j.pnmrs.2016.02.001

Schledorn, M., Malär, A. A., Torosyan, A., Penzel, S., Klose, D., Oss, A., et al. (2020). Protein NMR Spectroscopy at 150 kHz Magic-Angle Spinning Continues to Improve Resolution and Mass Sensitivity. *ChemBioChem* 21, 2540–2548. doi:10.1002/cbic.202000341

Sekhar, A., and Kay, L. E. (2019). An NMR View of Protein Dynamics in Health and Disease. *Annu. Rev. Biophys.* 48, 297–319. doi:10.1146/annurev-biophys-052118-115647

Sternberg, U., Witter, R., Kuprov, I., Lamley, J. M., Oss, A., Lewandowski, J. R., et al. (2018). 1H Line Width Dependence on MAS Speed in Solid State NMR - Comparison of experiment and Simulation. *J. Magn. Reson.* 291, 32–39. doi:10.1016/j.jmr.2018.04.003

Thurber, K. R., and Tycko, R. (2009). Measurement of Sample Temperatures under Magic-Angle Spinning from the Chemical Shift and Spin-Lattice Relaxation Rate of 79Br in KBr Powder. *J. Magn. Reson.* 196, 84–87. doi:10.1016/j.jmr.2008.09.019

Tian, C., Kasavajhala, K., Belfon, K. A. A., Raguette, L., Huang, H., Migues, A. N., et al. (2019). ff19SB: Amino-acid-specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theor. Comput.* 16, 528–552. doi:10.1021/acs.jctc.9b00591

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004). Development and Testing of a General Amber Force Field. *J. Comput. Chem.* 25, 1157–1174. doi:10.1002/jcc.20035

Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., et al. (1995). ^{1}H, ^{13}C and ^{15}N Chemical Shift Referencing in Biomolecular NMR. *J. Biomol. NMR* 6, 135–140. doi:10.1007/bf00211777

Wylie, B. J., and Rienstra, C. M. (2008). Multidimensional Solid State NMR of Anisotropic Interactions in Peptides and Proteins. *J. Chem. Phys.* 128, 052207. doi:10.1063/1.2834735

Wylie, B. J., Sperling, L. J., Nieuwkoop, A. J., Franks, W. T., Oldfield, E., and Rienstra, C. M. (2011). Ultrahigh Resolution Protein Structures Using NMR Chemical Shift Tensors. *Proc. Natl. Acad. Sci. U. S. A.* 108, 16974–16979. doi:10.1073/pnas.1103728108

Zhou, D. H., Shea, J. J., Nieuwkoop, A. J., Franks, W. T., Wylie, B. J., Mullen, C., et al. (2007). Solid-state Protein-Structure Determination with Proton-Detected Triple-Resonance 3D Magic-Angle-Spinning NMR Spectroscopy. *Angew. Chem. Int. Ed. Engl.* 46, 8380–8383. doi:10.1002/anie.200702905

Zhou, D. H., and Rienstra, C. M. (2008). High-performance Solvent Suppression for Proton Detected Solid-State NMR. *J. Magn. Reson.* 192, 167–172. doi:10.1016/j.jmr.2008.01.012

Zhou, D. H., Shea, J. J., Nieuwkoop, A. J., Franks, W. T., Wylie, B. J., Mullen, C., et al. (2007). Solid-state Protein-Structure Determination with Proton-Detected Triple-Resonance 3D Magic-Angle-Spinning NMR Spectroscopy. *Angew. Chem. Int. Ed.* 46, 8380–8383. doi:10.1002/anie.200702905

Zhou, Z., Kümmerle, R., Qiu, X., Redwine, D., Cong, R., Taha, A., et al. (2007). A New Decoupling Method for Accurate Quantification of Polyethylene Copolymer Composition and Triad Sequence Distribution with 13C NMR. *J. Magn. Reson.* 187, 225–233. doi:10.1016/j.jmr.2007.05.005

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 10 December 2021
doi: 10.3389/fmolb.2021.767040

Determination of Histidine Protonation States in Proteins by Fast Magic Angle Spinning NMR

Roman Zadorozhnyi[1,2], Sucharita Sarkar[1,2], Caitlin M. Quinn[1], Kaneil K. Zadrozny[3], Barbie K. Ganser-Pornillos[3], Owen Pornillos[3], Angela M. Gronenborn[2,4] and Tatyana Polenova[1,2]**

[1]Department of Chemistry and Biochemistry, University of Delaware, Newark, DE, United States, [2]Pittsburgh Center for HIV Protein Interactions, University of Pittsburgh School of Medicine, Pittsburgh, PA, United States, [3]Department of Molecular Physiology and Biological Physics, University of Virginia School of Medicine, Charlottesville, VA, United States, [4]Department of Structural Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA, United States

Histidine residues play important structural and functional roles in proteins, such as serving as metal-binding ligands, mediating enzyme catalysis, and modulating proton channel activity. Many of these activities are modulated by the ionization state of the imidazole ring. Here we present a fast MAS NMR approach for the determination of protonation and tautomeric states of His at frequencies of 40–62 kHz. The experiments combine ^{1}H detection with selective magnetization inversion techniques and transferred echo double resonance (TEDOR)–based filters, in 2D heteronuclear correlation experiments. We illustrate this approach using microcrystalline assemblies of HIV-1 CA_{CTD}-SP1 protein.

Keywords: Magic angle spinning (MAS), nuclear magnetic resonance (NMR) spectroscopy, histidine protonation state, transferred echo double resonance (TEDOR), Fast MAS NMR, solid-state NMR

OPEN ACCESS

Edited by:
Amir Goldbourt,
Tel Aviv University, Israel

Reviewed by:
Loren B. Andreas,
Max Planck Institute for Biophysical Chemistry, Germany
Hans-Heinrich Limbach,
Freie Universität Berlin, Germany

****Correspondence:***
Tatyana Polenova
tpolenov@udel.edu
Angela M. Gronenborn
amg100@pitt.edu

Specialty section:
This article was submitted to Structural Biology, a section of the journal Frontiers in Molecular Biosciences

Received: *30 August 2021*
Accepted: *15 October 2021*
Published: *10 December 2021*

Citation:
Zadorozhnyi R, Sarkar S, Quinn CM, Zadrozny KK, Ganser-Pornillos BK, Pornillos O, Gronenborn AM and Polenova T (2021) Determination of Histidine Protonation States in Proteins by Fast Magic Angle Spinning NMR. Front. Mol. Biosci. 8:767040. doi: 10.3389/fmolb.2021.767040

INTRODUCTION

Histidines (His) play important structural and functional roles in proteins such as metal binding (Stryer et al., 1964; Perutz and Mathews, 1966; Adams et al., 1969; Liljas et al., 1972), proton transfer (Hoffee et al., 1967; Blow et al., 1969; Campbell et al., 1974), and stability (Perutz et al., 1969; Lewis et al., 1976; Loewenthal et al., 1992). These functions are often correlated with the ionization state of the histidine sidechain (**Figure 1A**) (Bachovchin and Roberts, 1978; Kossiakoff and Spencer, 1981; Lewis et al., 1981). While the pK_a of the imidazole ring for free histidine is 6.5 (Blomberg et al., 1977), in proteins the pK_a values vary widely, from 3 to 9, depending on the interactions with neighboring residues and degree of burial (Zhou et al., 1993; Plesniak et al., 1996). At pH values above the pK_a, anionic τ and π tautomers with hydrogens at either $N^{\varepsilon 2}$ or $N^{\delta 1}$ are present, while below the pK_a the protonated imidazole ring possesses hydrogens at both $N^{\varepsilon 2}$ and $N^{\delta 1}$. For a protein at intermediate pH values, it is possible that a fraction of His residues is protonated and the remaining fraction unprotonated (French and Hammes, 1965; Edwards and Sykes, 1980; Hass et al., 2008).

Methods to determine His ionization states in proteins are solution NMR (Kilmartin et al., 1973; Markley, 1975; Bachovchin and Roberts, 1978; Perutz et al., 1985; Pelton et al., 1993; Shimba et al., 1998; Hass et al., 2008; Hansen and Kay, 2014) or neutron diffraction (Kossiakoff and Spencer, 1980; Maeda et al., 2004; Kovalevsky et al., 2010), with the latter limited to very large single crystals and requiring a neutron source, both difficult conditions to meet routinely. Therefore, solid-state magic angle spinning (MAS) NMR constitutes a viable alternative (Wei et al., 1999). Similar to solution NMR, the tautomeric state of histidines can be unambiguously determined from a unique combination of ^{15}N sidechain chemical shifts (Munowitz et al., 1982;

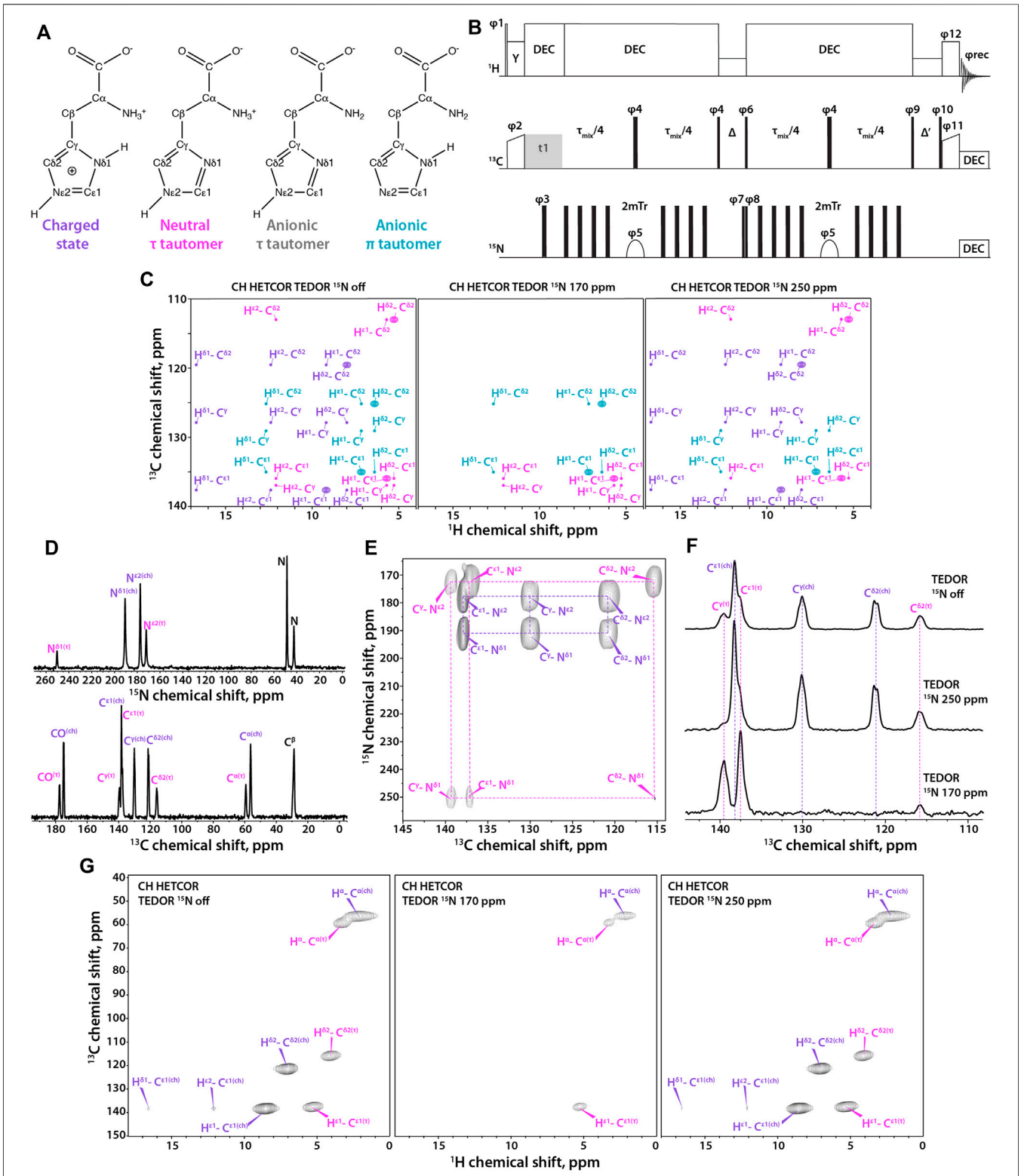


FIGURE 1 | (A) Four states of histidine: left to right, charged state, neutral τ tautomer, anionic τ tautomer, and anionic π tautomer. **(B)** Pulse sequence for the ^{1}H-detected TEDOR-based ^{15}N selective filtered experiment. T_r is the MAS rotor period, τ_{mix} is the total TEDOR mixing time. The phase on the individual pulses are: φ1 = 16 × (0) 16 × (2), φ2 = 1, φ3 = 0,φ4 = 0, φ5 = 0213 2031, φ6 = 2, φ7 = 0, φ8 = 02, φ9 = 1133, φ10 = 4 × (0) 4 × (1) 4 × (2) 4 × (3), φ11 = 4 × (1) 4 × (0), φ12 = 4 × (1) 4 × (0) 4 × (1) 4 × (0) 4 × (3) 4 × (2) 4 × (3) 4 × (2), φrec = 3113 0220 1331 2002, where 0 = *x*, 1 = *y*, 2 = –*x*, and 3 = –*y*. Δ is set to one rotor period during which ^{1}H rf field of ω_r amplitude is applied for effective *Z*-filtering. MISSISSIPPI water suppression sequence is applied during Δ′ time period. **(C)** Synthetic ^{1}H-detected TEDOR-based ^{15}N selective

(*Continued*)

FIGURE 1 | filtered CH HETCOR spectra showing cross peaks expected for each tautomer. Left to right, soft pulse turned off, soft pulse at 170 ppm, soft pulse at 250 ppm. The filtering patterns for neutral and anionic τ tautomers are identical. **(D)** ^{15}N **(top)** and ^{13}C **(bottom)** CPMAS NMR spectra of crystalline histidine. **(E)** Aromatic region expansion of 2D NCA spectrum of crystalline histidine. **(F)** 1D ^{13}C spectra using TEDOR-based ^{15}N selective filtering in the aromatic region. Top to bottom, soft pulse turned off; soft pulse at 250 ppm; soft pulse at 170 ppm. **(G)** Three complementary ^{1}H-detected TEDOR-based ^{15}N selective filtered CH HETCOR spectra. Left to right, soft pulse turned off, soft pulse at 170 ppm, soft pulse at 250 ppm. The MAS frequency was 60 kHz in all experiments. Signals of charged state are shown in purple, neutral τ tautomer – in magenta, anionic τ tautomer - in grey, and anionic π tautomer – in teal.

Wei et al., 1999; Miao et al., 2014) and the corresponding N-H distances can be estimated, allowing for hydrogen bonding studies (Shenderovich et al., 2015). Protonation states for the crystalline histidine amino acid have been determined by MAS NMR for different pH values (Li and Hong, 2011) and crystalline short peptides (Platzer et al., 2014). Using ^{15}N selective filtered, ^{13}C-detected experiments with the inversion pulses at frequencies of the different tautomers (Miao et al., 2014) permits their identification. For proteins containing several histidine residues, the above experiments are challenging due to low sensitivity and spectral overlap. Therefore, only a handful of such studies have been reported to date (Hu et al., 2006; Hu et al., 2010; Miao et al., 2015; Kwon et al., 2019; Maciejko et al., 2019; Vasa et al., 2019; Movellan et al., 2020). In order to increase resolution, the original pulse sequence can be reconfigured as a 2D experiment by introducing a ^{13}C-^{13}C mixing period based on proton-driven spin diffusion (PDSD) (Bloembergen, 1949) and extending the second Z-filter (Miao et al., 2014). 2D and 3D proton-based experiments were also introduced with ^{1}H chemical shifts either recorded in the indirect dimension (Miao et al., 2015) or detected directly (Shenderovich et al., 2015; Vasa et al., 2019; Movellan et al., 2020).

Herein, we present an alternative MAS experiment that uses ^{1}H detected transferred-echo double resonance (TEDOR)-based ^{15}N selectively filtered 2D correlations at fast MAS frequencies of 40–60 kHz. The advantages of the ^{1}H-detected fast-MAS experiments presented here are: i) improved sensitivity due to ^{1}H detection, and ii) improved resolution *via* the second dimension and selective recoupling of aromatic resonances directly attached to ^{15}N atoms. Microcrystalline assemblies of U-^{13}C,^{15}N- and fractionally deuterated (FD) (Mance et al., 2015) ^{13}C,^{15}N-HIV-1 CA_{CTD}-SP1 protein samples, possessing solely a single His residue, His-226, are ideally suited for pulse sequence optimization and therefore were selected for illustrating our current approach. Extension to ultrafast MAS frequencies (up to 110 kHz), should yield even higher sensitivity and resolution for proteins with multiple histidines.

MATERIALS AND METHODS

Sample Preparation

U-^{13}C,^{15}N-L-histidine was purchased from Cambridge Isotope Laboratories, recrystallized from an aqueous solution at pH 6.0, adjusted by mixing HCl and NaOH. The sample was packed into a 1.3 mm MAS rotor. Microcrystalline assemblies of U-^{13}C,^{15}N- and FD-^{13}C,^{15}N-HIV-1 CA_{CTD}-SP1 were prepared in the presence of the assembly cofactor inositol hexakisphosphate (IP6) as described previously (Wagner et al., 2016) except for growing *Escherichia coli* in M9 medium containing ^{13}C glucose, ^{15}N NH_4Cl, isotopically labeled precursors, and (for the deuterated sample) D_2O. Proteins were assembled with 1.6 mM IP6 (Sigma-Aldrich), for a final reaction volume of 1 ml at pH 8.0. Assemblies were incubated overnight at 20°C and packed into 3.2 mm (U-^{13}C,^{15}N), 1.9 mm (FD-^{13}C,^{15}N), or 1.3 mm MAS rotors (U-^{13}C,^{15}N).

MAS NMR Spectroscopy

MAS NMR experiments on U-^{13}C,^{15}N-CA_{CTD}-SP1 and FD-^{13}C,^{15}N-CA_{CTD}-SP1 microcrystalline assemblies were performed on a 20.0 T Bruker AVIII spectrometer outfitted with 3.2 mm E-Free HCN and 1.9 HCN probes, respectively. The MAS frequency was 14 and 40 kHz, respectively, controlled to within ± 10 Hz by a Bruker MAS controller. The actual sample temperature was maintained at 4 ± 1°C throughout the experiments using the Bruker temperature controller.

The Larmor frequencies were 850.4 MHz (^{1}H), 213.9 MHz (^{13}C) and 86.2 MHz (^{15}N). The typical 90° pulse lengths were 2.6–3.0 μs for ^{1}H, 4.3–4.5 μs for ^{13}C, and 4.2–4.7 μs for ^{15}N. The ^{1}H-^{13}C and ^{1}H-^{15}N cross-polarization employed a linear amplitude ramp of 90–110% on ^{1}H, and the center of the ramp was matched to a Hartmann–Hahn condition at the first spinning sideband; contact times of 0.7–1.5 ms and 1.0–1.7 ms were used, respectively. 50 ms CORD (Hou et al., 2013) mixing time was applied to facilitate ^{13}C-^{13}C mixing.

MAS NMR experiments on U-^{13}C,^{15}N-L-histidine and FD-^{13}C,^{15}N-CA_{CTD}-SP1 microcrystalline assemblies were performed on a 14.1 T Bruker AVIII spectrometer outfitted with 1.3 mm HCN probe. Larmor frequencies were 599.8 MHz (^{1}H), 150.8 MHz (^{13}C), and 60.7 MHz (^{15}N). The MAS frequency was 60 kHz, controlled to within ± 10 Hz by a Bruker MAS controller. The actual sample temperature was maintained at 40 ± 1°C throughout the experiments using the Bruker temperature controller. The typical 90° pulse lengths were 1.4–1.6 μs for ^{1}H, 2.7–3.0 μs for ^{13}C, and 3.3–3.6 μs for ^{15}N. The ^{1}H-^{13}C and ^{1}H-^{15}N cross-polarization employed a linear amplitude ramp of 90–110% on ^{1}H, center of the ramp was matched to a Hartmann–Hahn condition at the first spinning sideband, with contact times of 1.0–5.0 ms and 1.3–5.0 ms, respectively. Band-selective ^{15}N-^{13}C SPECIFIC-CP contact time was 5.0–6.0 ms. SWFTPPM (Vinod Chandran et al., 2008) decoupling (15 kHz) was used during the TEDOR block and acquisition periods. The selective ^{15}N 180° r-SNOB (Kupce et al., 1995) pulse length in the Z-filtered TEDOR experiments was 500 μs and the bandwidth — 2 kHz; the rf power was 4 kHz. During the Z-filter time period Δ, 60 kHz CW decoupling was applied for τ_r on ^{1}H channel, while during the time period Δ', MISSISSIPPI (Zhou and Rienstra, 2008) water suppression was applied. The TEDOR block duration was 1–3 ms.

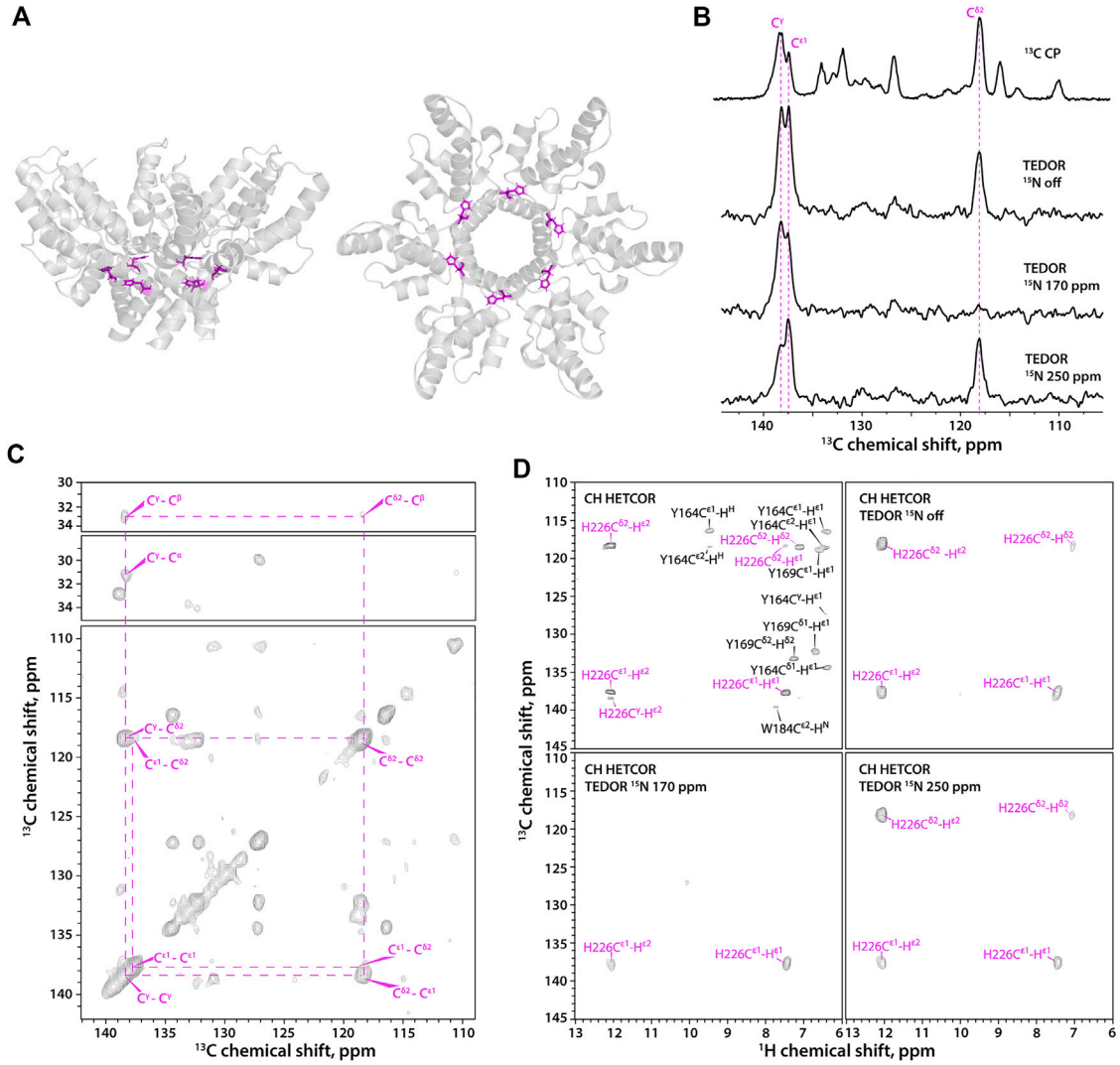


FIGURE 2 | (A) A hexameric unit of HIV-1 CA_{CTD}-SP1 in the microcrystalline assembly (PDB 5I4T) shown as side view **(left)** and top view **(right)**. **(B)** 1D ^{13}C MAS NMR spectra of FD-^{13}C,^{15}N-CA_{CTD}-SP1 with TEDOR-based ^{15}N selective filtering in the aromatic region. Top to bottom: CPMAS spectrum; TEDOR-based ^{15}N selectively filtered spectra with soft pulse turned off, soft pulse at 170 ppm, and soft pulse at 250 ppm. **(C)** 2D CORD spectrum of FD-^{13}C,^{15}N-CA_{CTD}-SP1 (MAS frequency 14 kHz). **(D)** Aromatic regions of ^{1}H-detected TEDOR-based ^{15}N selective filtered CH HETCOR spectra in FD-^{13}C,^{15}N-CA_{CTD}-SP1: TEDOR filter and soft pulse turned off **(top left)**, soft pulse turned off **(top right)**, soft pulse at 170 ppm **(bottom left)**, soft pulse at 250 ppm **(bottom right)**. The MAS frequency was 40 kHz in all experiments, unless indicated otherwise. Signals of τ tautomer are shown in magenta.

Data Processing

All MAS NMR data were processed using NMRPipe (Delaglio et al., 1995). The ^{13}C and ^{15}N chemical shifts were referenced with respect to the external standards adamantane (Morcombe and Zilm, 2003) and ammonium chloride (Bertani et al., 2014), respectively. The 2D and 3D data sets were processed by applying 30, 45, 60, and 90° shifted sine bell apodization followed by a Lorentzian-to-Gaussian transformation in both dimensions. Forward linear prediction to twice the number of the original data points was used in the indirect dimension followed by zero filling. The processed spectra were analyzed in NMRFAM-Sparky (Goddard and Kneller, 2004; Lee et al., 2015) and CCPN (Stevens et al., 2011).

RESULTS

Here, we report on a 2D ^{1}H-detected TEDOR-based Z-filtered experiment, which incorporates ^{15}N selective filters for the determination of histidine tautomeric states. The pulse sequence is shown in **Figure 1B**. The experiment is well suited for fast MAS frequencies of 40 kHz and above. The tautomeric states of His residues are unambiguously determined using a combination of three CH HETCOR experiments comprising: i) ^{15}N selective TEDOR filter, containing ^{13}C resonances of all protonation and tautomeric states present; ii) ^{15}N selective TEDOR filter with a soft pulse at 170 ppm, removing resonances of the protonated state while $C^{\varepsilon1}$ and $C^{\delta2}$ atoms of π tautomer and $C^{\varepsilon1}$ and C^{γ} atoms of τ tautomer

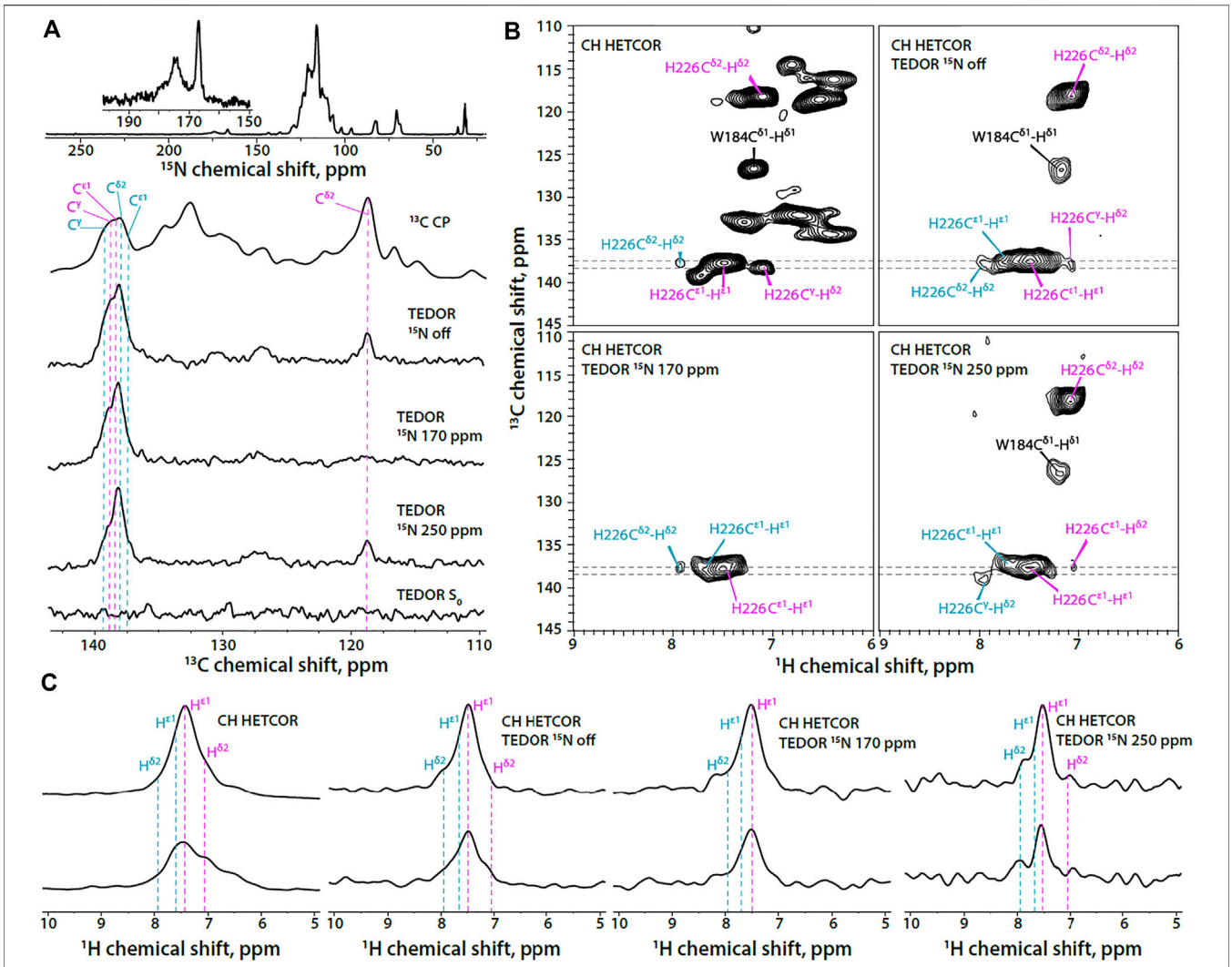


FIGURE 3 | (A) 1D ^{15}N CPMAS and ^{13}C MAS NMR spectra of U-^{13}C,^{15}N-CA$_{CTD}$-SP1 with TEDOR-based ^{15}N selective filtering in the aromatic region (MAS frequency 14 kHz). Top to bottom: ^{15}N CPMAS spectrum; ^{13}C CPMAS spectrum; TEDOR-based ^{15}N selectively filtered spectra with soft pulse turned off, soft pulse at 170 ppm, soft pulse at 250 ppm, and a reference (S_0) experiment. **(B)** Aromatic regions of ^{1}H-detected TEDOR-based ^{15}N selective filtered CH HETCOR spectra in U-^{13}C,^{15}N-CA$_{CTD}$-SP1: TEDOR filter and soft pulse turned off **(top left)**, soft pulse turned off **(top right)**, soft pulse at 170 ppm **(bottom left)**, soft pulse at 250 ppm **(bottom right)**. The first contour was set at 5× the noise rmsd. **(C)** 1D ^{1}H slices of ^{1}H-detected TEDOR-based ^{15}N selective filtered CH HETCOR spectra in U-^{13}C,^{15}N-CA$_{CTD}$-SP1, extracted at ^{13}C shifts shown as gray dashed lines in panel **(B)**. Left to right: TEDOR filter and soft pulse turned off, soft pulse turned off, soft pulse at 170 ppm, soft pulse at 250 ppm. The MAS frequency was 60 kHz. Signals of τ tautomer and π tautomer are shown in magenta and teal, respectively.

remain; and iii) ^{15}N selective TEDOR filter with a soft pulse at 250 ppm, retaining all signals of the charged state, C$^{\varepsilon1}$ and C$^{\gamma}$ of the π tautomer as well as C$^{\varepsilon1}$ and C$^{\delta2}$ atoms of the τ tautomer. C$^{\varepsilon1}$ of anionic tautomers is always present in TEDOR filtered spectra, but has reduced peak intensity when ^{15}N selective pulse is applied as C-N dipolar interaction with the non-selectively irradiated nitrogen atom is recoupled. The sequence was first tested on a crystalline L-histidine sample prepared at pH 6.0. The ^{13}C and ^{15}N 1D CPMAS and 2D NCA spectra are shown in **Figures 1D, E**, respectively. The spectra clearly indicate the presence of two forms of L-histidine, the charged monohydrate and the τ tautomer, in approximately 2:1 ratio. As shown in **Figure 1F**, conventional ^{13}C-detected TEDOR-based experiments are well suited for the determination of protonation states in this sample. To test the ^{1}H-detected sequences proposed herein, three complementary experiments were performed. As shown in **Figure 1G**, ^{15}N selective TEDOR-filtered CH HETCOR without or with a soft pulse at 250 ppm (left and right panels, respectively) yield the sidechain signals of both protonation states, while ^{15}N selective TEDOR-filtered CH HETCOR with soft pulse at 170 ppm retains only C$^{\varepsilon1}$ resonance of the τ tautomer (chemical shifts provided in **Supplementary Table S1**). Water suppression was incorporated into the second Z-filter, allowing to record spectra on hydrated samples.

HIV-1 CA$_{CTD}$-SP1 (**Figure 2A**) contains a single His residue, His-226. The outstanding high spectral resolution in the

microcrystalline FD-^{13}C,^{15}N-CA$_{CTD}$-SP1 sample allows for the determination of histidine protonation and tautomeric states even in the ^{13}C-detected mode (**Figure 2B**). The C$^{\varepsilon1}$ and C$^{\gamma}$ resonances are present in 1D experiments, while the C$^{\delta2}$ resonance is absent in the ^{15}N selective TEDOR-filtered ^{13}C CPMAS experiment with the soft pulse at 170 ppm since its magnetization does not build up during the TEDOR block due to the very weak dipolar coupling to N$^{\delta1}$ (chemical shifts provided in **Supplementary Table S2**). The 2D ^{13}C-^{13}C CORD spectrum clearly shows a single set of resonances, indicating the presence of only one histidine species (**Figure 2C**), although the protonation and tautomeric state cannot be determined without additional experiments. The three complementary ^{1}H-detected TEDOR-based ^{15}N selective CH HETCOR spectra (**Figure 2D**) also indicate the presence of a single species, which is unambiguously assigned as τ tautomer. These ^{1}H-detected 2D spectra contain no resonances of aromatic residues other than His (shown in black in the CH HETCOR spectrum) and Trp (these are weak or absent in the spectra of the deuterated sample), as only carbons attached to nitrogens are selected, making assignment of histidine resonances straightforward.

In contrast to the FD-^{13}C,^{15}N-CA$_{CTD}$-SP1, the His-226 protonation state in U-^{13}C,^{15}N-CA$_{CTD}$-SP1 assemblies cannot be easily determined using the 1D ^{13}C-detected version of TEDOR-based ^{15}N selective filtered experiments due to low resolution and spectral overlap **(Figure 3A)**. In contrast, the 2D ^{1}H-detected TEDOR-based ^{15}N selective filtered spectra (**Figure 3B**) suggest the presence of a small fraction of π tautomer along with the predominant τ tautomer in this sample (chemical shifts provided in **Supplementary Table S2**).

In addition to the His signals, the indole ring signals of the Trp184 residue are also present in the ^{1}H-detected TEDOR-based experiments when the soft pulse is either turned off or centered at 250 ppm. This is expected due to the nitrogen atom N$^{\varepsilon1}$ in the indole ring, which allows for magnetization build up on adjacent carbon atoms (C$^{\delta1}$ and C$^{\varepsilon2}$) during TEDOR transfer. Tryptophan sidechain resonances appear much stronger in non-deuterated protein assemblies compared to the FD-^{13}C,^{15}N -CA$_{CTD}$-SP1 and can be distinguished from those corresponding to the histidine based on chemical shift.

CONCLUSION

We demonstrated that ^{1}H-detected 2D Z-filtered TEDOR experiments incorporating ^{15}N selective filters permit unambiguous assignment of histidine protonation and tautomeric states in microcrystalline proteins and protein assemblies. This approach combines all the advantages of fast MAS and proton detection. Extending the experiments to MAS frequencies of 110 kHz and above can further improve the quality of data sets and allow unambiguous assignment of His protonation and tautomeric states in larger proteins and protein assemblies.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

TP and AMG conceived the project and guided the work. RZ performed NMR experiments and analyzed the experimental data. CMQ assisted with the NMR experiments and data analysis. SS assigned the CA$_{CTD}$-SP1 chemical shifts. KKZ, BKG-P, and OP prepared samples of microcrystalline CA$_{CTD}$-SP1 assemblies. RZ and TP took the lead in writing the manuscript. All authors discussed the results and contributed to the manuscript preparation.

FUNDING

This work was supported by the National Institutes of Health (NIH Grants P50AI1504817 and R01AI129678). We acknowledge the support of the National Science Foundation (NSF Grant CHE0959496) for the acquisition of the 850 MHz NMR spectrometer and of the National Institutes of Health (NIH Grant P30GM110758) for the support of core instrumentation infrastructure at the University of Delaware; and of the National Institutes of Health (NIH Grant S10OD012213) for the acquisition of the 750 MHz NMR spectrometer at the University of Pittsburgh.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.767040/full#supplementary-material

REFERENCES

Adams, M. J., Blundell, T. L., Dodson, E. J., Dodson, G. G., Vijayan, M., Baker, E. N., et al. (1969). Structure of Rhombohedral 2 Zinc Insulin Crystals. *Nature* 224, 491–495. doi:10.1038/224491a0

Bachovchin, W. W., and Roberts, J. D. (1978). Nitrogen-15 Nuclear Magnetic Resonance Spectroscopy. The State of Histidine in the Catalytic Triad of Alpha-Lytic Protease. Implications for the Charge-Relay Mechanism of Peptide-Bond Cleavage by Serine Proteases. *J. Am. Chem. Soc.* 100, 8041–8047. doi:10.1021/ja00494a001

Bertani, P., Raya, J., and Bechinger, B. (2014). ^{15}N Chemical Shift Referencing in Solid State NMR. *Solid State. Nucl. Magn. Reson.* 61-62, 15–18. doi:10.1016/j.ssnmr.2014.03.003

Bloembergen, N. (1949). On the Interaction of Nuclear Spins in a Crystalline Lattice. *Physica* 15, 386–426. doi:10.1016/0031-8914(49)90114-7

Blomberg, F., Maurer, W., and Rueterjans, H. (1977). Nuclear Magnetic Resonance Investigation of Nitrogen-15-Labeled Histidine in Aqueous Solution. *J. Am. Chem. Soc.* 99, 8149–8159. doi:10.1021/ja00467a005

Blow, D. M., Birktoft, J. J., and Hartley, B. S. (1969). Role of a Buried Acid Group in the Mechanism of Action of Chymotrypsin. *Nature* 221, 337–340. doi:10.1038/221337a0

Campbell, I. D., Lindskog, S., and White, A. I. (1974). A Study of the Histidine Residues of Human Carbonic Anhydrase B Using 270 MHz Proton Magnetic Resonance. *J. Mol. Biol.* 90, 469–489. doi:10.1016/0022-2836(74)90229-0

Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a Multidimensional Spectral Processing System Based on UNIX Pipes. *J. Biomol. NMR* 6, 277–293. doi:10.1007/bf00197809

Edwards, B. F. P., and Sykes, B. D. (1980). Nuclear Magnetic Resonance Evidence for the Coexistence of Several Conformational States of Rabbit Cardiac and Skeletal Tropomyosins. *Biochemistry* 19, 2577–2583. doi:10.1021/bi00553a007

French, T. C., and Hammes, G. G. (1965). Relaxation Spectra of Ribonuclease. II. Isomerization of Ribonuclease at Neutral pH Values. *J. Am. Chem. Soc.* 87, 4669–4673. doi:10.1021/ja00949a002

Goddard, T. D., and Kneller, D. G. (2004). *SPARKY 3*. Univ. California, San Fransisco.

Hansen, A. L., and Kay, L. E. (2014). Measurement of Histidine pKa Values and Tautomer Populations in Invisible Protein States. *Proc. Natl. Acad. Sci.* 111, E1705–E1712. doi:10.1073/pnas.1400577111

Hass, M. A. S., Hansen, D. F., Christensen, H. E. M., Led, J. J., and Kay, L. E. (2008). Characterization of Conformational Exchange of a Histidine Side Chain: Protonation, Rotamerization, and Tautomerization of His61 in Plastocyanin from Anabaena Variabilis. *J. Am. Chem. Soc.* 130, 8460–8470. doi:10.1021/ja801330h

Hoffee, P., Lai, C. Y., Pugh, E. L., and Horecker, B. L. (1967). The Function of Histidine Residues in Rabbit Muscle Aldolase. *Proc. Natl. Acad. Sci.* 57, 107–113. doi:10.1073/pnas.57.1.107

Hou, G., Yan, S., Trébosc, J., Amoureux, J.-P., and Polenova, T. (2013). Broadband Homonuclear Correlation Spectroscopy Driven by Combined $R2_n^v$ Sequences Under Fast Magic Angle Spinning for NMR Structural Analysis of Organic and Biological Solids. *J. Magn. Reson.* 232, 18–30. doi:10.1016/j.jmr.2013.04.009

Hu, J., Fu, R., Nishimura, K., Zhang, L., Zhou, H.-X., Busath, D. D., et al. (2006). Histidines, Heart of the Hydrogen Ion Channel from Influenza A Virus: toward an Understanding of Conductance and Proton Selectivity. *Proc. Natl. Acad. Sci.* 103, 6865–6870. doi:10.1073/pnas.0601944103

Hu, F., Luo, W., and Hong, M. (2010). Mechanisms of Proton Conduction and Gating in Influenza M2 Proton Channels from Solid-State NMR. *Science* 330, 505–508. doi:10.1126/science.1191714

Kilmartin, J. V., Breen, J. J., Roberts, G. C. K., and Ho, C. (1973). Direct Measurement of the pK Values of an Alkaline Bohr Group in Human Hemoglobin. *Proc. Natl. Acad. Sci.* 70, 1246–1249. doi:10.1073/pnas.70.4.1246

Kossiakoff, A. A., and Spencer, S. A. (1980). Neutron Diffraction Identifies His 57 as the Catalytic Base in Trypsin. *Nature* 288, 414–416. doi:10.1038/288414a0

Kossiakoff, A. A., and Spencer, S. A. (1981). Direct Determination of the Protonation States of Aspartic Acid-102 and Histidine-57 in the Tetrahedral Intermediate of the Serine Proteases: Neutron Structure of Trypsin. *Biochemistry* 20, 6462–6474. doi:10.1021/bi00525a027

Kovalevsky, A. Y., Chatake, T., Shibayama, N., Park, S.-Y., Ishikawa, T., Mustyakimov, M., et al. (2010). Direct Determination of Protonation States of Histidine Residues in a 2 Å Neutron Structure of Deoxy-Human Normal Adult Hemoglobin and Implications for the Bohr Effect. *J. Mol. Biol.* 398, 276–291. doi:10.1016/j.jmb.2010.03.016

Kupce, E., Boyd, J., and Campbell, I. D. (1995). Short Selective Pulses for Biochemical Applications. *J. Magn. Reson. Ser. B* 106, 300–303. doi:10.1006/jmrb.1995.1049

Kwon, B., Roos, M., Mandala, V. S., Shcherbakov, A. A., and Hong, M. (2019). Elucidating Relayed Proton Transfer through a His-Trp-His Triad of a Transmembrane Proton Channel by Solid-State NMR. *J. Mol. Biol.* 431, 2554–2566. doi:10.1016/j.jmb.2019.05.009

Lee, W., Tonelli, M., and Markley, J. L. (2015). NMRFAM-SPARKY: Enhanced Software for Biomolecular NMR Spectroscopy. *Bioinformatics* 31, 1325–1327. doi:10.1093/bioinformatics/btu830

Lewis, S. D., Johnson, F. A., and Shafer, J. A. (1976). Potentiometric Determination of Ionizations at the Active Site of Papain. *Biochemistry* 15, 5009–5017. doi:10.1021/bi00668a010

Lewis, S. D., Johnson, F. A., and Shafer, J. A. (1981). Effect of Cysteine-25 on the Ionization of Histidine-159 in Papain as Determined by Proton Nuclear Magnetic Resonance Spectroscopy. Evidence for a Histidine-159-Cysteine-25 Ion Pair and its Possible Role in Catalysis. *Biochemistry* 20, 48–51. doi:10.1021/bi00504a009

Li, S., and Hong, M. (2011). Protonation, Tautomerization, and Rotameric Structure of Histidine: a Comprehensive Study by Magic-Angle-Spinning Solid-State NMR. *J. Am. Chem. Soc.* 133, 1534–1544. doi:10.1021/ja108943n

Liljas, A., Kannan, K. K., Bergstén, P.-C., Waara, I., Fridborg, K., Strandberg, B., et al. (1972). Crystal Structure of Human Carbonic Anhydrase C. *Nat. New Biol.* 235, 131–137. doi:10.1038/newbio235131a0

Loewenthal, R., Sancho, J., and Fersht, A. R. (1992). Histidine-aromatic Interactions in Barnase. *J. Mol. Biol.* 224, 759–770. doi:10.1016/0022-2836(92)90560-7

Maciejko, J., Kaur, J., Becker-Baldus, J., and Glaubitz, C. (2019). Photocycle-dependent Conformational Changes in the Proteorhodopsin Cross-Protomer Asp-His-Trp Triad Revealed by DNP-Enhanced MAS-NMR. *Proc. Natl. Acad. Sci. USA* 116, 8342–8349. doi:10.1073/pnas.1817665116

Maeda, M., Chatake, T., Tanaka, I., Ostermann, A., and Niimura, N. (2004). Crystallization of a Large Single crystal of Cubic Insulin for Neutron Protein Crystallography. *J. Synchrotron Radiat.* 11, 41–44. doi:10.1107/s0909049503023859

Mance, D., Sinnige, T., Kaplan, M., Narasimhan, S., Daniëls, M., Houben, K., et al. (2015). An Efficient Labelling Approach to Harness Backbone and Side-Chain Protons in ^{1}H-Detected Solid-State NMR Spectroscopy. *Angew. Chem. Int. Ed.* 54, 15799–15803. doi:10.1002/anie.201509170

Markley, J. L. (1975). Observation of Histidine Residues in Proteins by Nuclear Magnetic Resonance Spectroscopy. *Acc. Chem. Res.* 8, 70–80. doi:10.1021/ar50086a004

Miao, Y., Cross, T. A., and Fu, R. (2014). Differentiation of Histidine Tautomeric States Using ^{15}N Selectively Filtered ^{13}C Solid-State NMR Spectroscopy. *J. Magn. Reson.* 245, 105–109. doi:10.1016/j.jmr.2014.06.005

Miao, Y., Fu, R., Zhou, H.-X., and Cross, T. A. (2015). Dynamic Short Hydrogen Bonds in Histidine Tetrad of Full-Length M2 Proton Channel Reveal Tetrameric Structural Heterogeneity and Functional Mechanism. *Structure* 23, 2300–2308. doi:10.1016/j.str.2015.09.011

Morcombe, C. R., and Zilm, K. W. (2003). Chemical Shift Referencing in MAS Solid State NMR. *J. Magn. Reson.* 162, 479–486. doi:10.1016/S1090-7807(03)00082-X

Movellan, K. T., Wegstroth, M., Overkamp, K., Leonov, A., Becker, S., and Andreas, L. B. (2020). Imidazole-Imidazole Hydrogen Bonding in the pH-Sensing Histidine Side Chains of Influenza A M2. *J. Am. Chem. Soc.* 142, 2704–2708. doi:10.1021/jacs.9b10984

Munowitz, M., Bachovchin, W. W., Herzfeld, J., Dobson, C. M., and Griffin, R. G. (1982). Acid-base and Tautomeric Equilibriums in the Solid State: Nitrogen-15 NMR Spectroscopy of Histidine and Imidazole. *J. Am. Chem. Soc.* 104, 1192–1196. doi:10.1021/ja00369a007

Pelton, J. G., Torchia, D. A., Meadow, N. D., and Roseman, S. (1993). Tautomeric States of the Active-Site Histidines of Phosphorylated and Unphosphorylated IIIGlc, a Signal-Transducing Protein from *Escherichia coli*, Using Two-Dimensional Heteronuclear NMR Techniques. *Protein Sci.* 2, 543–558. doi:10.1002/pro.5560020406

Perutz, M. F., and Mathews, F. S. (1966). An X-ray Study of Azide Methaemoglobin. *J. Mol. Biol.* 21, 199–202. doi:10.1016/0022-2836(66)90088-x

Perutz, M. F., Muirhead, H., Mazzarella, L., Crowther, R. A., Greer, J., and Kilmartin, J. V. (1969). Identification of Residues Responsible for the Alkaline Bohr Effect in Haemoglobin. *Nature* 222, 1240–1243. doi:10.1038/2221240a0

Perutz, M. F., Gronenborn, A. M., Clore, G. M., Fogg, J. H., and Shih, D. T.-b. (1985). The pKa Values of Two Histidine Residues in Human Haemoglobin, the Bohr Effect, and the Dipole Moments of α-helices. *J. Mol. Biol.* 183, 491–498. doi:10.1016/0022-2836(85)90016-6

Platzer, G., Okon, M., and Mcintosh, L. P. (2014). pH-Dependent Random Coil ^{1}H, ^{13}C, and ^{15}N Chemical Shifts of the Ionizable Amino Acids: a Guide for Protein pK a Measurements. *J. Biomol. NMR* 60, 109–129. doi:10.1007/s10858-014-9862-y

Plesniak, L. A., Connelly, G. P., Mcintosh, L. P., and Wakarchuk, W. W. (1996). Characterization of a Buried Neutral Histidine Residue in Bacillus Circulansxylanase: NMR Assignments, pH Titration, and Hydrogen Exchange. *Protein Sci.* 5, 2319–2328. doi:10.1002/pro.5560051118

Shenderovich, I. G., Lesnichin, S. B., Tu, C., Silverman, D. N., Tolstoy, P. M., Denisov, G. S., et al. (2015). NMR Studies of Active-Site Properties of Human Carbonic Anhydrase II by Using ^{15}N-Labeled 4-Methylimidazole as a Local Probe and Histidine Hydrogen-Bond Correlations. *Chem. Eur. J.* 21, 2915–2929. doi:10.1002/chem.201404083

Shimba, N., Takahashi, H., Sakakura, M., Fujii, I., and Shimada, I. (1998). Determination of Protonation and Deprotonation Forms and Tautomeric States of Histidine Residues in Large Proteins Using Nitrogen–Carbon J Couplings in Imidazole Ring. *J. Am. Chem. Soc.* 120, 10988–10989. doi:10.1021/ja982153g

Stevens, T. J., Fogh, R. H., Boucher, W., Higman, V. A., Eisenmenger, F., Bardiaux, B., et al. (2011). A Software Framework for Analysing Solid-State MAS NMR Data. *J. Biomol. NMR* 51, 437–447. doi:10.1007/s10858-011-9569-2

Stryer, L., Kendrew, J. C., and Watson, H. C. (1964). The Mode of Attachment of the Azide Ion to Sperm Whale Metmyoglobin. *J. Mol. Biol.* 8, 96–IN10. doi:10.1016/s0022-2836(64)80152-2

Vasa, S. K., Singh, H., Grohe, K., and Linser, R. (2019). Assessment of a Large Enzyme-Drug Complex by Proton-Detected Solid-State NMR Spectroscopy without Deuteration. *Angew. Chem. Int. Ed.* 58, 5758–5762. doi:10.1002/anie.201811714

Vinod Chandran, C., Madhu, P. K., Kurur, N. D., and Bräuniger, T. (2008). Swept-frequency Two-Pulse Phase Modulation (SWf-TPPM) Sequences with Linear Sweep Profile for Heteronuclear Decoupling in Solid-State NMR. *Magn. Reson. Chem.* 46, 943–947. doi:10.1002/mrc.2285

Wagner, J. M., Zadrozny, K. K., Chrustowicz, J., Purdy, M. D., Yeager, M., Ganser-Pornillos, B. K., et al. (2016). Crystal Structure of an HIV Assembly and Maturation Switch. *eLife* 5, e17063. doi:10.7554/eLife.17063

Wei, Y., De Dios, A. C., and McDermott, A. E. (1999). Solid-State ^{15}N NMR Chemical Shift Anisotropy of Histidines: Experimental and Theoretical Studies of Hydrogen Bonding. *J. Am. Chem. Soc.* 121, 10389–10394. doi:10.1021/ja9919074

Zhou, D. H., and Rienstra, C. M. (2008). High-performance Solvent Suppression for Proton Detected Solid-State NMR. *J. Magn. Reson.* 192, 167–172. doi:10.1016/j.jmr.2008.01.012

Zhou, M. M., Davis, J. P., and Van Etten, R. L. (1993). Identification and pKa Determination of the Histidine Residues of Human Low-Molecular-Weight Phosphotyrosyl Protein Phosphatases: A Convenient Approach Using MLEV-17 Spectral Editing Scheme. *Biochemistry* 32, 8479–8486. doi:10.1021/bi00084a012

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 25 January 2022
doi: 10.3389/fmolb.2021.789478

In-Cell NMR of Intact Mammalian Cells Preserved with the Cryoprotectants DMSO and Glycerol Have Similar DNP Performance

Yiling Xiao[1†‡], *Rupam Ghosh*[1†‡] *and Kendra K. Frederick*[1,2*†]

[1]*Department of Biophysics, UT Southwestern Medical Center, Dallas, TX, United States,* [2]*Center for Alzheimer's and Neurodegenerative Disease, UT Southwestern Medical Center, Dallas, TX, United States*

OPEN ACCESS

Edited by:
Loren B Andreas,
Max Planck Institute for Biophysical Chemistry, Germany

Reviewed by:
Moreno Lelli,
University of Florence, Italy
Henrike Heise,
Heinrich Heine University of Düsseldorf, Germany

***Correspondence:**
Kendra K. Frederick
kendra.frederick@utsouthwestern.edu

†ORCID:
Yiling Xiao
orcid.org/0000-0002-4548-4833
Rupam Ghosh
orcid.org/0000-0002-3095-621X
Kendra K. Frederick
orcid.org/0000-0002-1656-5167

‡*These authors have contributed equally to this work and share first authorship*

Specialty section:
This article was submitted to Structural Biology, a section of the journal Frontiers in Molecular Biosciences

Received: *05 October 2021*
Accepted: *31 December 2021*
Published: *25 January 2022*

Citation:
Xiao Y, Ghosh R and Frederick KK (2022) In-Cell NMR of Intact Mammalian Cells Preserved with the Cryoprotectants DMSO and Glycerol Have Similar DNP Performance. Front. Mol. Biosci. 8:789478. doi: 10.3389/fmolb.2021.789478

NMR has the resolution and specificity to determine atomic-level protein structures of isotopically-labeled proteins in complex environments and, with the sensitivity gains conferred by dynamic nuclear polarization (DNP), NMR has the sensitivity to detect proteins at their endogenous concentrations. Prior work established that DNP MAS NMR is compatible with cellular viability. However, in that work, 15% glycerol, rather than the more commonly used 10% DMSO, was used as the cellular cryoprotectant. Moreover, incubation of cells cryoprotected 15% glycerol with the polarization agent, AMUPol, resulted in an inhomogeneous distribution of AMUPol through the cellular biomass, which resulted in a spatial bias of the NMR peak intensities. Because 10% DMSO is not only the most used cryoprotectant for mammalian cells, but also because DMSO is often used to improve delivery of molecules to cells, we sought to characterize the DNP performance of cells that were incubated with AMUPol and cryoprotected with 10% DMSO. We found that, like cells preserved with 15% glycerol, cells preserved with 10% DMSO retain high viability during DNP MAS NMR experiments if they are frozen at a controlled rate. However, DMSO did not improve the dispersion of AMUPol throughout the cellular biomass. Cells preserved with 15% glycerol and with 10% DMSO had similar DNP performance for both the maximal DNP enhancements as well as the inhomogeneous dispersion of AMUPol throughout the cellular biomass. Therefore, 10% DMSO and 15% glycerol are both appropriate cryoprotectant systems for DNP-assisted MAS NMR of intact viable mammalian cells.

Keywords: Dynamic nuclear polarization (DNP), AMUPol, cryopreservation, DMSO (dimethyl sulphoxide), glycerol, in-cell NMR, HEK293

INTRODUCTION

In-cell structural biology enables the study of protein conformation in environments that maintain the identity, stoichiometry, concentrations and organization of the myriad of biomolecules that can interact with a protein of interest. (Frederick et al., 2015; Theillet et al., 2016; Burmann et al., 2020; Luchinat et al., 2020) Capturing the effect of these complicated environments on biomolecular conformation is of particular importance for proteins that have more than one stable conformation, interact with cellular components or contain regions of intrinsic disorder. Nuclear Magnetic Resonance (NMR) is uniquely suited to study proteins in these complicated contexts with

atomic level resolution. NMR spectroscopy detects only NMR-active nuclei. These nuclei are non-perturbative probes that can be specifically incorporated into a protein of interest that is either delivered to or expressed inside the cell. (Selenko et al., 2006; Inomata et al., 2009; Theillet et al., 2013; Majumder et al., 2015; Burmann et al., 2020) NMR has the resolution and specificity to determine atomic-level protein structures of isotopically-labeled proteins in complex environments (Sakakibara et al., 2009) and, with the sensitivity gains conferred by dynamic nuclear polarization (DNP), NMR has the sensitivity to detect proteins at their endogenous concentrations (Renault et al., 2012; Frederick et al., 2015; Albert et al., 2018; Costello et al., 2019; Narasimhan et al., 2019; Schlagnitweit et al., 2019).

We recently established that sample conditions that favor efficient DNP enhancements are compatible with cellular viability. In that work, we established methods that maintained cellular viability throughout the DNP NMR experiments and found that the magnitude of the sensitivity enhancements for such samples were high enough to enable detection of a protein at micromolar concentrations inside intact cells in experimentally tractable experimental times. (Ghosh et al., 2021) Briefly, cells were cryoprotected, transferred to rotors and frozen at the controlled rate of 1°C per minute before cryogenic transfer to the pre-cooled NMR spectrometer for analysis. (Ghosh et al., 2020; Ghosh et al., 2021) After structural characterization via DNP MAS NMR, these cells can be cultured or imaged and their phenotype can be determined and compared with cells before structural characterization. (Ghosh et al., 2021) However, that work only examined cells that were cryopreserved using 15% glycerol as the cryoprotectant. While the overall approach to sample preparation is likely to be generalizable to freezing media with different compositions, this has not been explicitly demonstrated. The most common cryoprotectant for cultured mammalian cells is dimethylsulfoxide (DMSO) at a concentration of 10% (v/v). Indeed, the handful of studies that examine preparations of mammalian cells using DNP NMR use DMSO (Albert et al., 2018; Narasimhan et al., 2019; Schlagnitweit et al., 2019; Overall et al., 2020), although the sample composition—including the choice of cryoprotectant—and post-experiment cellular viability, have only very recently been considered (Ghosh et al., 2020; Ghosh et al., 2021; Overall and Barnes, 2021). Given the widespread preference for 10% DMSO over 15% glycerol as the cryoprotectant for cellular cryopreservation, we sought to determine if cryoprotection using 10% DMSO could also support cellular viability throughout the DNP NMR experiments.

DNP increases the sensitivity of NMR spectroscopy through the transfer of the large spin polarization of an unpaired electron to nearby nuclei (Ni et al., 2013) which are typically introduced into a sample by doping with millimolar concentrations of stable biological radicals (Sauvée et al., 2013; Lund et al., 2020; Stevanato et al., 2020). The sensitivity enhancements from DNP rely upon proximity to a polarization agent. Thus, DNP-enhanced MAS NMR experiments are biased towards observation of molecules that are accessible to polarization agents. Despite how critical the dispersion of polarization agents in a sample is to both achieve high sensitivity and interpret the results, the dispersion of polarization agents in intact cells has only very recently been considered (Ghosh et al., 2021). In our recent work that described methods for DNP MAS NMR on viable cells we described two of many potential approaches to deliver polarization agents to intact cells. In that work, we introduced AMUPol to cells by incubation of intact cells with AMUPol and by electroporation of intact cells in the presence of AMUPol to transiently permeabilize the membrane. (Ghosh et al., 2021) We compared the distribution of AMUPol throughout the cellular biomass for cells prepared in these two different ways to the distribution of AMUPol throughout the cellular biomass for cellular lysates where the cellular membrane does not present a barrier to distribution. We found that while AMUPol was homogenously distributed in cellular lysates and cells where AMUPol had been introduced by electroporation. AMUPol was inhomogeneously distributed in cells where AMUPol was delivered by incubation. In samples of cells incubated with AMUPol, the signal intensity from DNA in the nucleus was lower than the signal intensity from proteins and RNA in the cytoplasm. Thus, data from experiments on such samples will report qualitatively, and not quantitatively, on the structural ensemble; any observed conformation in such samples certainly exists, but the relative population of that conformation to any other cannot be inferred from peak intensities. The method used to introduce the polarization agent affects the experimental result and therefore must be chosen to address the structural question under consideration. Interestingly, DMSO is not only often used as a cryoprotectant (Lovelock and Bishop, 1959) but is often also used as a cellular penetration enhancer (Williams and Barry, 2004). Here we assessed the performance of 10% DMSO to determine not only if it is able to support cellular viability throughout DNP MAS NMR but also to determine if it can improve delivery of the polarization agent, AMUPol, to the cell.

MATERIALS AND METHODS

Sample Preparation

To reduce experimental acquisition times, we uniformly isotopically labeled HEK293 cells by culturing them in isotopically-enriched media. Human embryonic kidney 293 (HEK293) cells were cultured in ^{13}C, ^{15}N labelled media (BioExpress 6000 Mammalian U-^{13}C, 98%; U-^{15}N, 98%, Cambridge Isotope Laboratories, MA, United States) with 10% (v/v) fetal bovine serum (FBS, qualified, Gibco) and 1% (v/v) PenStrep (Gibco) at 37 °C and 5% CO_2. Confluent plates were harvested using Tryp-LE Express (Gibco) and BioExpress 6000 media, transferred to 15 ml conical tube and centrifuged at 233 x *g* for 5 min at 22 °C using a swinging bucket rotor (Beckman Coulter). Pelleted cells were resuspended and washed once with 1x PBS (-$CaCl_2$, -$MgCl_2$, pH 7.4, Gibco). AMUPol was delivered to cells by incubation, to do so, a 50 μL cell pellet was mixed with 50 μL perdeuterated 1x PBS (85% D_2O + 15% H_2O, pH 7.4) containing AMUPol (Cortecnet, NY, United States) and 11 μL of d_6-DMSO. The 111 μL cell suspension had a final composition of 10% (v/v) d_6-DMSO, 76.5% (v/v) D_2O and 13.5% (v/v) H_2O. After delivery of AMUPol, cells were transferred into 3.2 mm sapphire rotor

(Bruker) by centrifugation in a swinging bucket rotor at 100 x *g* for 2 min at 22 °C. The supernatant was removed, and rotors were frozen at a controlled rate (1 °C/min) in "Cool Cell LX" (Corning) in the −80 °C freezer for 12–16 h. Finally, frozen rotors were transferred to liquid nitrogen storage until measurement by DNP NMR.

Trypan Blue Exclusion Assay

Pelleted cells (10 μL) were diluted into 100 μL unlabeled DMEM and 10 μL of this cell suspension were mixed with 10 μL of Trypan Blue (0.4% solution). 10 μL of the Trypan Blue cell suspension was loaded onto Countess Chamber. Trypan blue membrane permeability was assessed using Countess automated cell counter (Life technologies) using the manufacturer's protocol.

Growth Assay

Equal number of cells (1 million cells) were plated in 10 cm dish containing complete media (DMEM) and grown for 9–14 days (as indicated before). After cells have settled down (post 8–10 h), media was removed to get rid of floating dead cells. 10–12 ml of DMEM is added to the 10 cm culture dish and cell growth is monitored using inverted light microscope till 100% confluency. Fitting of sigmoidal curves was performed with an equation of $y(t) = \frac{a}{1+e^{-k(t-t_0)}}$, where y($t$) denotes the cell culture time t, a and k are fitting parameters, and t_0 defines a lag time of t_L as $t_L = t_0 - 2/k$. (Nielsen et al., 2001) The error range for the fitting was estimated at the 95% confidence level.

DNP NMR Spectroscopy

Rotors were transferred in liquid nitrogen directly into the NMR probe that was pre-equilibrated at 100 K. All dynamic nuclear polarization magic angle spinning nuclear magnetic resonance (DNP MAS NMR) experiments were performed on a 600 MHz Bruker Ascend DNP NMR spectrometer/7.2 T Cryogen-free gyrotron magnet (Bruker), equipped with a ^{1}H, ^{13}C, ^{15}N triple-resonance, 3.2 mm low temperature (LT) DNP MAS NMR Bruker probe (600 MHz). The sample temperature was 104 K and the MAS frequency was 12 kHz. The DNP enhancement for the instrumentation set-up for a standard sample of 1.5 mg of uniformly ^{13}C, ^{15}N labeled proline (Isotech) suspended in 25 mg of 60:30:10 d_8-glycerol:D_2O:H_2O containing 10 mM AMUPol was between 130 and 140 and a $T_{B,on}$ of 4.6 s. For ^{13}C cross-polarization (CP) MAS experiments, the ^{13}C radio frequency (RF) amplitude was linearly swept from 75 to 37.5 kHz with an average of 56.25 kHz. ^{1}H RF amplitude was 68–72 kHz for CP, 83 kHz for 90 degree pulse, and 85 kHz for ^{1}H TPPM decoupling with phase alternation of ± 15° during acquisition of ^{13}C signal. The DNP enhancements were determined by comparing 1D ^{13}C CP spectra collected with and without microwaves irradiation. For $T_{B,on}$ measurements, recycle delays ranged from 0.1 to 300 s. To determine the $T_{B,on}$, the dependence of the recycle delay using saturation recovery on both ^{13}C peak intensity or volume was fit to the mono-exponential equation $I_t = I_0(1 - e^{\frac{-t}{T_{B,on}}})$ and the stretched-exponential equation $I_t = I_0 \times [1 - e^{-(\frac{t}{T_{B,on}})^{\beta}}]$, respectively.

^{13}C–^{13}C 2D correlations were measured using 20 ms DARR mixing with the ^{1}H amplitude at the MAS frequency. A total of 280 complex t_1 points with increment of 25 μs were recorded. For ^{13}C-^{15}N 1D and 2D correlations, a 24 rotor periods (2 ms) TEDOR sequence was applied with ^{13}C and ^{15}N pulse trains at 55.5 and 41.7 kHz, respectively. A total of 64 complex t_1 points with an increment of 80 μs were recorded. The recycle delay was 3.9 s and the same ^{1}H decoupling was applied. The experimental time required to collect a 2D TEDOR spectra with 32 scans was 2 h and to collect a 2D DARR of 16 scans was 5 h.

DNP NMR Data Analysis

For 1D experiments, the data were processed using NMRPipe (Delaglio et al., 1995). The real part of the processed spectrum was exported using pipe2txt.tcl command. Peaks were integrated, and the time constants were obtained by least-squares fitting with a single-exponential function. DNP enhancements were determined by peak intensity. For 2D experiments, the TEDOR and DARR data were both apodized with a Lorenz-to-Gauss window function with IEN of 15 Hz and GB of 75 Hz in the t_1 and t_2 time domains. The noise level and peak height from the 2D NMR spectrum was detected by the NMRDraw software for S/N estimation.

RESULTS AND DISCUSSION

HEK293 Cells Cryopreserved With 10% DMSO Remain Viable During DNP MAS NMR

The polarization agent, AMUPol, is not toxic to HEK293 cells in the presence of 10% DMSO

To determine whether AMUPol in the presence of the cryoprotectant 10% d_6-DMSO compromised cellular viability, we used a trypan blue dye exclusion test to determine the percentage of cells with intact membranes present in a sample. HEK293 viability was not compromised by replacement of media components with PBS, per-deuteration and addition 10% d_6-DMSO (**Supplementary Figure S1A**). Moreover, HEK293 viability was not compromised by addition of the polarization agent AMUPol at concentrations up to 50 mM (**Supplementary Figure S1B**).

Cells Cryopreserved With 10% DMSO Retain High Viability After DNP MAS NMR

To determine whether any of the manipulations required for DNP MAS NMR sample preparation compromise cellular viability when 10% DMSO is used as a cryoprotectant, we assessed trypan blue membrane permeability at several steps of our sample preparation workflow (**Figure 1A**, arrows). After harvesting adherent cells from tissue culture plates, the cells were rinsed with PBS and pelleted. At this point, cellular membrane integrity as assessed by trypan blue dye exclusion tests was high (95 ± 3%, **Figure 1**, dark red). Addition of 10% DMSO and AMUPol followed by transfer into 3.2 mm NMR rotors did not significantly decrease membrane integrity (91 ±

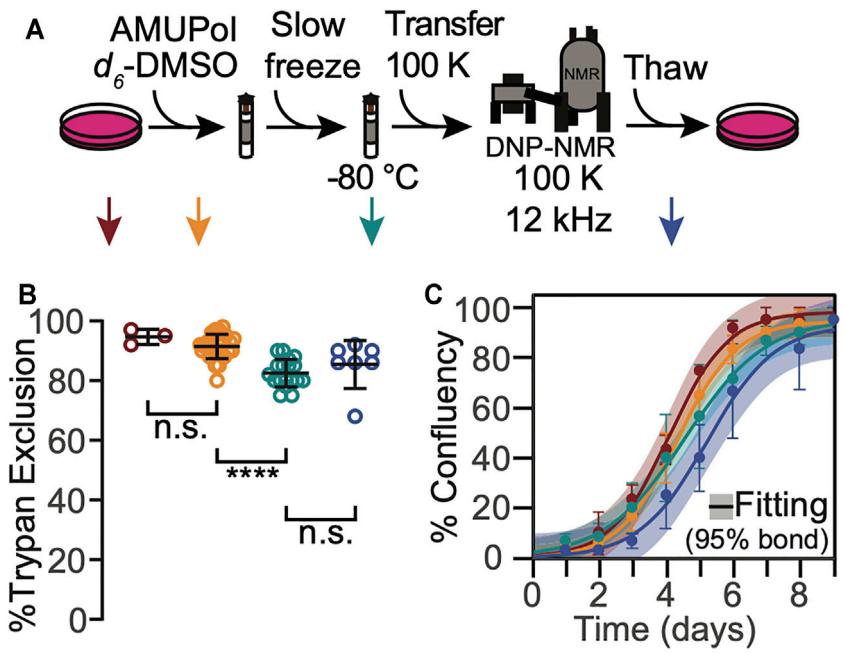


FIGURE 1 | HEK293 cells that are cryopreserved with 10% DMSO are viable throughout the DNP NMR process. **(A)**. Experimental scheme of the DNP NMR sample preparation procedure. Colored arrows indicate points at which sample viability was assessed. Viability was assessed for cells after trypsinization and washing (dark red), after suspension in AMUPol and cryoprotectants (orange), after being frozen at 1 °C per min (green), and after the entire DNP MAS NMR experiment (blue). **(B)** Percentage of cells with trypan impermeable membranes at each sample assessment point, colored as in A. Each point represents an independent sample. Black bars indicate average and standard deviation. Brackets indicate results of two-tailed homoscedastic student's *t*-tests. (n.s. $p > 0.05$, ****$p < 0.0001$). **(C)** Growth kinetics as assessed by confluency, colored as in A. The averages and standard deviations of three independent experiments are indicated by circles and error bars, respectively. The best fit of sigmoid is indicated in solid lines and the 95% confidence interval by the shaded area.

4%, **Figure 1**, orange; $p = 0.19$). Freezing cryoprotected cells at the controlled rate of 1 °C/min slightly compromised membrane integrity as assessed by trypan blue dye exclusion test (decrease of 10 ± 6% to 82 ± 5%, **Figure 1B**, green, p = 1e-5). Post-NMR, trypan blue membrane integrity was indistinguishable from that of slow frozen samples (83 ± 10%, **Figure 1**, blue, $p = 1$). Cryopreservation of cells with 10% DMSO is therefore compatible with high membrane integrity post-DNP NMR MAS.

The membrane integrity throughout the DNP MAS NMR sample preparation protocol for cells cryoprotected with 10% DMSO was similar, though not identical, to that for cells cryoprotected with 15% glycerol (Ghosh et al., 2021). The membrane integrity of frozen cells was lower than that of fresh cells for cells cryoprotected with either 10% DMSO or 15% glycerol, however the decrease in membrane integrity occurred at different points in the sample preparation. The membrane integrity of cells cryoprotected with 10% DMSO did not decrease upon addition of the cryoprotectant and slightly decreased (10 ± 6%) upon freezing. In contrast, the membrane integrity of cells cryoprotected with 15% glycerol slightly decreased upon addition of the cryoprotectant (5 ± 10%) and was unchanged by freezing. This difference likely reflects the difference in the mechanisms of interaction of the cryoprotectants with cellular membranes. In both cases, the viability of cryoprotected frozen sample, the state that is most representative of the state of the sample during NMR data collection, was the same. Interestingly, the membrane integrity of these sample after DNP MAS NMR was different. It was higher by 14 ± 14% for cells cryopreserved with 10% DMSO than it was for cells cryopreserved with 15% glycerol ($p = 0.03$). However, the loss in membrane integrity after DNP MAS NMR experimentation for cells that were cryopreserved with 15% glycerol is a result of the manipulations required to remove the cells from the rotor, and not the DNP MAS NMR experiment itself. (Ghosh et al., 2021) This indicated that membranes of cells cryopreserved with 10% DMSO were less sensitive to the manipulations required to unpack the rotor than cells cryopreserved with 15% glycerol. This may reflect differences in intracellular ice content, which can recrystallize under slow thawing conditions and damage cells, and/or in changes in diffusion and osmosis across the cellular membrane, which may result in membrane rupture if they exceed the tolerance of the cellular membrane (Pegg, 2007). Overall, this indicates cellular membrane integrity is maintained for cells cryopreserved with 10% DMSO before, during and after DNP MAS NMR. The maintenance of cellular membrane integrity for cells cryopreserved with 10% DMSO and 15% glycerol is similar before and during DNP MAS NMR experimentation and is better for cells that are cryopreserved with 10% DMSO than for cells cryopreserved with 15% glycerol after the DNP MAS NMR experiment.

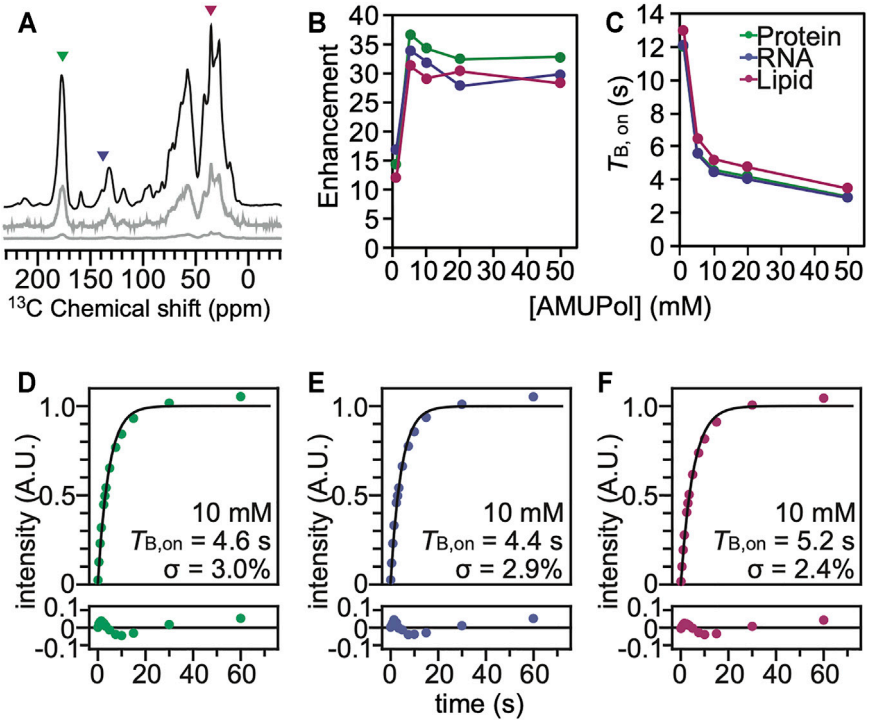


FIGURE 2 | The polarization agent, AMUPol, effectively polarizes all the components of HEK293 cells cryoprotected with 10% DMSO. **(A)** ^{13}C cross-polarization spectra of cryopreserved HEK293 cells grown on isotopically enriched media with 10 mM AMUPol at 100 K taken at 600 MHz with 12 kHz magic angle spinning and a recycle delay of 10 s. Displayed spectra are taken with (black) and without (grey) microwave irradiation. The microwave off spectrum is plotted on the same scale as the microwave on spectrum (bottom) and with the intensity multiplied by 10 (middle). Colored arrowheads indicate peaks that are representative of proteins (green), nucleotides (blue) and lipids (pink). **(B)** DNP enhancement and **(C)** $T_{B,on}$ values from saturation recovery experiments are dependent upon the AMUPol concentration. Fits of the $T_{B,on}$ data to a mono-exponential equation (black line) for different biomass components for cells incubated with 10 mM AMUPol with 10% DMSO as a cryoprotectant. **(D)** The protein component had a $T_{B,on}$ value of 4.6 s with a regression error (lower plot) of 3.0%. **(E)** The nucleotide component had a $T_{B,on}$ value of 4.4 s with a regression error (lower plot) of 2.9%. **(F)** The lipid component had a $T_{B,on}$ value of 5.2 s with a regression error (lower plot) of 2.4%.

To determine whether any of the manipulations required for DNP MAS NMR sample preparation compromised cellular propagative ability, we next assessed cellular growth kinetics at each step in our workflow. We found none of the manipulations significantly altered growth kinetics for cells that have been cryoprotected with 10% DMSO (**Figure 1C**). Cell growth curves were well-fit by a sigmoidal function with lag phase (R^2 = 0.99 ± 0.01). The lag phases and cell growth rates were indistinguishable across all the tested conditions (p > 0.26) and all plates reached 100% confluency. Similar to the membrane integrity results, the growth kinetic results for cells cryoprotected with 10% DMSO were similar, though not identical, to those for cells cryoprotected with 15% glycerol (Ghosh et al., 2021). The most notable difference was that exposure to glycerol prolongs the lag phase by 1.5 ± 0.5 days (Ghosh et al., 2021) while exposure to 10% DMSO does not. Otherwise, as for cells cryopreserved with 10% DMSO, no other perturbations significantly altered the growth kinetics. The maintenance of cellular propagative ability for cells cryopreserved with 10% DMSO and 15% glycerol is similar throughout DNP MAS NMR experimentation, although cells cryopreserved with 10% DMSO do not experience a lag phase. This indicates that while both 10% DMSO and 15% glycerol are effective cryoprotectants, 10% DMSO may be a better choice of cryoprotectants for experiments that will benefit from post-NMR cellular growth-based phenotyping.

Addition of AMUPol to HEK293 Cells Results in DNP Enhancement of all Biomass Components

Cells cryopreserved with 10% DMSO and 15% glycerol have similar DNP performance

Using characteristic peaks in the NMR spectra as reporters of the different cellular biomass components (Ghosh et al., 2021), we assessed DNP performance for cells that had been incubated with AMUPol and cryopreserved using 10% DMSO. We collected ^{13}C cross-polarization (CP) spectra with and without microwave irradiation to determine the DNP enhancements for HEK293 cells incubated with a range of AMUPol concentrations. We determined DNP enhancements for peaks in the ^{13}C CP spectra that are representative of the major biomass components of HEK293 cells; proteins, nucleotides and lipids (**Figure 2**). While some chemical moieties, like carbonyls, are found in more than one major biomass component—the contribution of lipid head groups could be up to 7% of the "protein" peak

and the contribution of aliphatic side chains could be up to 22% of the "lipid" peak—these peaks serve as quantitative proxies for the different biomass components (Ghosh et al., 2021). We found that the DNP enhancements for intact HEK293 cells that were cryopreserved with 10% DMSO reached a maximum value of 39 for the protein component and addition of 5 mM AMUPol sufficed to attain this enhancement. Addition of higher concentrations of AMUPol to the sample did not significantly alter the DNP enhancement across all biomass components (**Figure 2B**). The DNP enhancements for cells cryoprotected with 10% DMSO and then incubated with AMUPol are very similar to those for cells cryoprotected with 15% glycerol and then incubated with AMUPol (Ghosh et al., 2021). The dependence of the DNP enhancements on the concentration of AMUPol for cells cryoprotected with 10% DMSO and 15% glycerol were indistinguishable (p = 0.44, n = 5). We next assessed the DNP build-up times ($T_{B,on}$) (Pinon et al., 2017) for cells that had been incubated with AMUPol and cryopreserved using 10% DMSO. As expected, we found that the value of $T_{B,on}$ decreased with increasing AMUPol concentrations. The $T_{B,on}$ for the protein component of cells incubated with 5 mM AMUPol was 5.6 s and decreased to 3.0 s for cells that were incubated with 50 mM AMUPol. The dependence of $T_{B,on}$ values on the concentration of AMUPol for cells cryoprotected with 10% DMSO are very similar to those for cells cryoprotected with 15% glycerol. The dependence of $T_{B,on}$ values on concentration of AMUPol for cells cryoprotected with 10% DMSO and 15% glycerol were indistinguishable (p = 0.16, n = 5). Interestingly, the maximal enhancement for proteins inside intact cells, regardless of the cryoprotectant, is ~40 which is half of the maximal enhancement for proteins in cellular lysates, where the plasma membrane of the cell doesn't present an accessibility barrier. The higher maximal enhancements and the much steeper dependance of $T_{B,on}$ on AMUPol concentration for cellular lysates than for intact cells that were incubated with AMUPol suggests that, as was previously observed for cells cryoprotected with 15% glycerol, the AMUPol concentration inside of cells cryoprotected with 10% DMSO is lower than the concentration of AMUPol that was added to the sample. This indicates that AMUPol is heterogeneously distributed in samples of intact cells cryopreserved with 10% DMSO.

AMUPol is Heterogeneously Distributed in Cells Cryopreserved in Both 10% DMSO and 15% Glycerol

To assess the homogeneity of the AMUPol concentration throughout each biomass component, we used the regression error of the fit of the $T_{B,on}$ data to a mono-exponential equation (Ghosh et al., 2021) as well as a stretched exponential function where β describes the degree of deviation from an exponential fit (Pinon et al., 2017; Rankin et al., 2019). The regression error is a modestly more sensitive measure for deviation from a monoexponential and the regression error and the β factor are strongly anti-correlated. Both the regression error and β are reported in **Supplementary Table S2**. If the concentration distribution of AMUPol is heterogenous, there will be a mixture of underlying $T_{B,on}$ values which will increase the regression error. For reference, the regression error of the fit of the $T_{B,on}$ data to a mono-exponential function of the amino acid proline suspended in a matrix of 60:30:10 (v/v) glycerol:D_2O:H_2O with 10 mM AMUPol was 0.5% and represents the error expected from experimental noise (Ghosh et al., 2021). For intact cells cryopreserved with 10% DMSO, the regression error for protein was 2.8 ± 0.8% and for nucleotide the regression error was 2.6 ± 0.6% (indistinguishable from protein, p = 0.19, n = 5), while the regression error for lipid was 2.2 ± 0.6% (lower than protein and nucleotide, p < 0.003, n = 5) (**Supplementary Table S2**). These regression errors were indistinguishable from those for intact cells cryopreserved with 15% glycerol (p = 0.74) and were significantly larger than the regression error for lysed cells, where the plasma membrane does not present a barrier to accessibility (p = 0.005) as well as for intact cells where AMUPol was introduced inside the cell by electroporation (p = 0.004) across all biomass components. When AMUPol is dispersed homogenously throughout the sample, the regression errors are small. The larger regression errors for cell incubated with AMUPol and cryopreserved with 10% DMSO indicates the concentration distribution of AMUPol is more heterogenous in these samples than in samples of lysed cells or cells where AMUPol is delivered by electroporation. DMSO is sometimes used to improve cellular permeability of small molecules. If DMSO improves delivery of AMUPol to cells, the regression error for cells incubated with AMUPol and cryopreserved with 10% DMSO should be smaller than those for cells incubated with AMUPol and cryopreserved with 15% glycerol. However, the regression errors are indistinguishable. This indicates that the choice of cryoprotectant does not alter the delivery of the polarization agent to the cell. Finally, it is possible that inhomogeneities in the dispersion of the radical could result from the formation of ice crystals, rather than from larger scale inhomogeneities that result from physical exclusion of the radical from cell interiors by the plasma membrane. However, this is unlikely. Ice crystal formation inside cells kills cells and these samples had high post thaw viability. Moreover, when AMUPol is introduced into cells by electroporation, which circumvents the physical exclusion of the radical from the cell interior, cryoprotected cells also have high post-thaw viabilities, along with homogenous radical dispersions and high DNP enhancements (Ghosh et al., 2021). Thus, incubation of cells with AMUPol results in inhomogenous distribution of the AMUPol throughout the sample, regardless of the choice of cryoprotectant.

To further explore the distribution of AMUPol in samples of intact cells cryopreserved with 10% DMSO, we collected DNP-enhanced 2D ^{13}C-^{15}N correlation spectra (TEDOR) (Jaroniec et al., 2002) and assessed the signal to noise ratios for biomass components with different cellular distributions. We compared the normalized peak intensities of cytosolic and nucleic components for cells incubated with AMUPol and cryopreserved with DMSO to those of lysed samples and intact samples that were either electroporated or incubated with AMUPol and cryopreserved with 15% glycerol (Ghosh

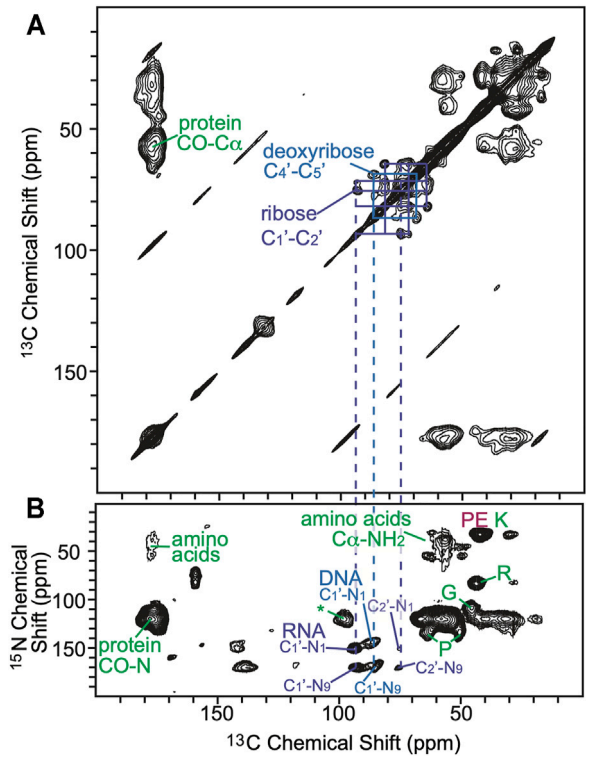


FIGURE 3 | (A) 2D homonuclear correlation spectra (DARR) of cells cryoprotected with DMSO. Selected ^{13}C–^{13}C correlations from carbons in the ribose (purple) and deoxyribose (blue) rings of RNA and DNA are annotated. **(B)** 2D heteronuclear correlation spectra (zTEDOR) of cells cryoprotected with DMSO. Selected ^{13}C-^{15}N correlations from the protein back bone and sidechains (green), from RNA (purple), from DNA (blue) and from lipid (pink) are annotated. The signal to noise ratios of selected peaks are reported in **Supplementary Table S1**.

et al., 2021). The TEDOR peak intensities were normalized to either the ribose-purine peak of RNA or DNA for each sample to control for differences in DNP-enhancements and cross-polarization efficiencies. TEDOR spectra have distinct peaks for DNA, RNA, protein backbone sites, protein side chain moieties, and free amino acids (**Figure 3**). DNA is located only in the nucleus, while RNA, proteins and free amino acids are entirely or largely cytoplasmically localized (i.e. more than 80% of the protein content of a cell is non-nuclearly localized). (Shaiken and Opekun, 2014) In addition to the RNA and DNA ribose purine and pyrimidine peaks, we determined peak intensities for the amide-carbonyl and amide-C_α sites for both proteins and free amino acids as well as the carbon-nitrogen bonds in the protein side chains of arginine and glycine (**Supplementary Table S1**). When the intensity of the amino acids peaks are compared to the ribose-purine peak of RNA, the ratio of the cross-peak intensities for the sample cryopreserved with DMSO were similar to those for the cellular lysate and intact cellular samples that had been prepared with 15% glycerol as the cryoprotectant, regardless of the method of AMUPol delivery (incubation, electroporation or addition to lysed cells) ($p > 0.06$, student's paired *t*-test) (**Supplementary Table S1**) (Ghosh et al., 2021). As an example, the glycine C_α-N cross peak was 1.23, 1.17, 1.43, and 1.16 times more intense than the ribose-purine cross peak of RNA for lysed cell, intact cells electroporated with AMUPol, intact cells incubated with AMUPol in 15% glycerol, and intact cells incubated AMUPol in DMSO, respectively. The similarity of the relative cross-peak intensities for the cytoplasmically-located biomass components across different approaches to sample preparation indicated that the cytoplasmic distribution of AMUPol is similar for all these samples. When the intensity of the amino acid and RNA peaks are compared to the deoxyribose-purine peak of DNA, we found that the sample incubated with AMUPol and cryopreserved with DMSO had indistinguishable peak intensity ratios to those for the sample incubated with AMUPol and cryopreserved with 15% glycerol ($p = 0.24$, Student's t-test, paired) and very different peak intensity ratios from the lysed and electroporated cells ($p < 0.02$, Student's t-test, paired). For example, the glycine C_α-N cross peak is 1.75, 1.80, 2.55 and 2.67 times more intense than the deoxyribose-purine cross peak of DNA for lysed cell, intact cells electroporated with AMUPol, intact cells incubated with AMUPol in 15% glycerol and intact cells incubated AMUPol in DMSO, respectively. The DNA peaks for intact cells incubated with AMUPol and then cryoprotected with either DMSO or glycerol were less intense than expected; the ratios of peak intensities for cytoplasmic to nuclear components were larger by 53% ± 19% ($p < 0.05$) (**Supplementary Table S1** and reference 14). Because the nuclear envelope is known to be permeable to AMUPol in intact cells (Ghosh et al., 2021), this suggested that the AMUPol concentration in the nucleus of cells incubated with AMUPol, regardless of choice of cryoprotectant, is lower than the concentration of AMUPol in the cytoplasm. Overall, AMUPol is heterogeneously distribution when intact cells are incubated with AMUPol. While AMUPol can polarize all the biomass components, including DNA, the relatively lower intensity of the DNA peaks combined with larger regression errors indicate that there is an AMUPol concentration gradient inside these cells. Although DMSO can improve delivery of small molecules to cells, there is no indication that DMSO improves delivery of AMUPol. The choice of cryoprotectant does not alter the delivery of the polarization agent to the cell.

CONCLUSION

Prior work established that sample conditions that favor efficient DNP enhancements are compatible with cellular viability and that the magnitude of the sensitivity enhancements is high enough to enable detection of a protein at micromolar concentrations in experimentally tractable experimental times. However, in that work, 15% glycerol, rather than the more commonly used 10% DMSO, was used as the cellular cryoprotectant. Moreover, incubation of cells cryoprotected 15% glycerol with AMUPol resulted in an inhomogeneous distribution of the polarization agent, AMUPol, through the cellular biomass, which will result in a spatial bias of the NMR peak intensities. Because 10% DMSO is not only the most used

cryoprotectant for mammalian cells, but also because DMSO is often used to improve delivery of molecules to cells, we sought to characterize the DNP performance of cells that were incubated with AMUPol and cryoprotected with 10% DMSO. We found that, like cells preserved with 15% glycerol, cells preserved with 10% DMSO retain high viability during DNP MAS NMR experiments. Moreover, cells preserved with 10% DMSO were less sensitive to the manipulations required to unpack cells from the NMR rotor, suggesting that it may be a better cryoprotectant for experiments that require post-NMR growth-based phenotyping. However, DMSO did not improve the dispersion of AMUPol throughout the cellular biomass. Cells preserved with 15% glycerol and with 10% DMSO had similar DNP performance for both the maximal DNP enhancements as well as the inhomogeneous dispersion of AMUPol throughout the cellular biomass. Therefore, we establish that 10% DMSO and 15% glycerol can be used interchangeably for DNP-assisted MAS NMR of intact viable mammalian cells.

Here we examined the cryopreservation and DNP properties for cells that were cryopreserved using concentrations of cryoprotectants at their established working concentrations of 10% for DMSO and 15% glycerol. At these working concentrations, the cryoprotective properties and DNP performance were indistinguishable. However, addition of different percentages of the same cryoprotectants can dramatically affect viability. Prior work established that suspension of cells in 60% glycerol, a percentage commonly used in biological DNP samples, or 60% DMSO both resulted in significant losses of membrane integrity and a complete loss of propagative ability (Ghosh et al., 2021). Thus, both cryoprotectants are compatible with high DNP enhancements on intact viable mammalian cells, but only at concentrations appropriate for cellular cryoprotection.

Because DMSO and glycerol have indistinguishable cryoprotective properties and DNP performance, both cryoprotectant systems are well-suited for in cell DNP MAS NMR of mammalian cells. Because both cryoprotectants fulfill the major requirements of viability maintenance and DNP efficiency, the choice of cryoprotectant depends upon question under investigation. For example, long term exposure to even low concentrations of DMSO is toxic. Although this is unlikely to be a major concern since the exposure to high concentrations of DMSO is transient, glycerol does not have the same toxicity profile and may be a more appropriate choice for sensitive cellular systems. However, we observed that cells preserved with 10% DMSO were less sensitive to the manipulations required to unpack cells from the NMR rotor, suggesting that DMSO may be a better cryoprotectant for experiments that require post-NMR growth-based phenotyping. Interestingly, DMSO and glycerol interact differently with the cell membrane. In general, DMSO de-solvates lipid membranes, increases the chain melting temperature (Yu and Quinn, 1995), induces water pores and increases floppiness in lipid membranes (Notman et al., 2006) while glycerol affects lipid hydration only to the same degree as it does of bulk water (Schrader et al., 2016) and does not alter the chain melting temperature of lipid membranes (McDaniel et al., 1983). Thus, while macroscopically DMSO may protect cellular membranes from rupturing, microscopically, glycerol may better preserve the local character of the membrane which could be particularly important for investigations of membrane-associated biomolecules and may be a more appropriate cryoprotectant for questions that require maintenance of the local structural integrity of lipid membranes. Finally, the work presented here was done on cells that were grown on isotopically enriched media. Therefore, the spectral contribution of the cryoprotectants to the signal were negligible because the ^{13}C content from natural abundance carbon in the cryoprotectants accounts for ~0.1% of the volume of the sample. However, for samples where the target molecule is at concentrations that are low enough that signals from natural abundance components make a significant contribution to the spectra (Costello et al., 2019), the contribution of the cryoprotectant peak to the spectra becomes a consideration. DMSO has one ^{13}C peak at 40 ppm and glycerol has two ^{13}C peaks at 65 and 75 ppm (**Supplementary Figure S2**). DMSO overlaps with protein sidechains while glycerol overlaps with the alpha carbons of some amino acids and the ribose ring of nucleic acids. While there is currently no source for ^{13}C-depleted d_6-DMSO, ^{13}C-depleted d_8-glycerol is commercially available and reduces the ^{13}C content of the cryoprotectant by an order of magnitude, which may make glycerol a more attractive choice for sensitivity-limited samples. Because both DMSO and glycerol are both well-suited for in-cell DNP MAS NMR, the choice of cryoprotectant system can be tailored to the system under investigation.

Finally, although DMSO is often used to improve delivery of molecules to cells, it did not improve the delivery of the polarization agent, AMUPol, to cells. The DNP enhancements, $T_{B,on}$ values and residual errors for samples preserved with 10% DMSO were indistinguishable from those for sample preserved with 15% glycerol. We considered the possibility that the delivery of AMUPol was improved in the presence of DMSO, but the reductive environment of the cell inactivated the AMUPol inside the cell (Jagtap et al., 2015; Karthikeyan et al., 2018), resulting in similar DNP performance. For cells that were cryopreserved with 15% glycerol, the DNP performance for cells incubated with AMUPol was relatively constant for room temperature incubation times of up to 2 h because the plasma membrane is semi-permeable to AMUPol and as the small amount that enters the cell is reduced, it is replenished by the large concentration of AMUPol in the interstitial space (Ghosh et al., 2021). More generally, the reduction of AMUPol by mammalian cells is slow relative to the sample preparation time; the half-life of AMUPol in intact cells is about an hour (Ghosh et al., 2021). Additionally, the $T_{B,on}$ values for samples prepared with both 10% DMSO and 15% glycerol are indistinguishable. Because the bi-nitroxide radicals in AMUPol, are reduced independent of each other. The monoradical form of AMUPol is DNP-silent, but still contributes to paramagnetic relaxation (McCoy et al., 2019). The accumulation of monoradical forms of AMUPol explains the observation that maximal enhancement for intact mammalian cells, where the AMUPol is reduced during the sample preparation time, is about half of the maximal enhancement for lysed cells, which can be flash frozen which prevents the build-up of monoradical forms, yet have similar $T_{B,on}$ values (Ghosh

et al., 2021). The monoradical forms of AMUPol shorten the $T_{B,on}$ without contributing to the enhancement. Thus, if more AMUPol is delivered to cells in the 10% DMSO condition but then is also reduced by cells, the enhancements could be similar but the $T_{B,on}$ values for those preparations should be shorter. However, this is not the case. Therefore, the presence of 10% DMSO did not improve delivery of a polarzation agent AMUPol to HEK293 cell.

Because the sensitivity enhancements from DNP rely upon proximity to a polarization agent, DNP-enhanced MAS NMR experiments are biased towards observation of molecules that are accessible to polarization agents. Here we found that for cells incubated with AMUPol and cryoprotected with 10% DMSO, a minority of the AMUPol enters the cell; the peak intensities for DNA are lower than expected and the $T_{B,on}$ fits indicate that the AMUPol concentration is heterogenous. Thus, while the identity, stoichiometry, concentrations and organization of the cellular components for cells incubated with AMUPol are all maintained, the experimental read-out from such samples are spatially biased, just like they are for cells incubated with AMUPol and cryoprotected with 15% glycerol. Data from experiments performed on intact cells incubated with AMUPol are qualitative rather than quantitative. While any observed conformation inside cells incubated with AMUPol exists, the relative populations of different conformations cannot be inferred from peak intensities. For in cell work where such quantitative information is required, alternative approaches that result in homogenous dispersion of the polarization agent—like electroporation—are more appropriate (Ghosh et al., 2021). Investigation of protein conformations inside viable cells using DNP MAS NMR creates an experimental system with the ability to tightly couple genotypes, phenotypes and environments (e.g., presence/absence of a drug) to specific structures or structural ensembles. Cryoprotection of cells using the commonly used cryoprotectant, DMSO, is compatible with in cell DNP MAS NMR.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

YX, RG and KKF contributed to conception and design of the study. RG and YX prepared samples and collected data. KKF wrote the first draft of the manuscript. RG, YX, and KKF wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

RG is supported by a fellowship from the O'Donnell Brain Institute Neural Science Training Program. This work was supported by grants from the National Institute of Health (NS-111236), the Welch Foundation (I-1923-20200401, the Lupe Murchison Foundation and the Kinship Foundation (Searle Scholars Program) to KKF.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.789478/full#supplementary-material

REFERENCES

Albert, B. J., Gao, C., Sesti, E. L., Saliba, E. P., Alaniva, N., Scott, F. J., et al. (2018). Dynamic Nuclear Polarization Nuclear Magnetic Resonance in Human Cells Using Fluorescent Polarizing Agents. *Biochemistry* 57, 4741–4746. doi:10.1021/acs.biochem.8b00257

Burmann, B. M., Gerez, J. A., Matečko-Burmann, I., Campioni, S., Kumari, P., Ghosh, D., et al. (2020). Regulation of α-synuclein by Chaperones in Mammalian Cells. *Nature* 577, 127–132. doi:10.1038/s41586-019-1808-9

Costello, W. N., Xiao, Y., and Frederick, K. K. (2019). DNP-assisted NMR Investigation of Proteins at Endogenous Levels in Cellular Milieu. *Methods Enzymol.* 615, 373–406. doi:10.1016/bs.mie.2018.08.023

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a Multidimensional Spectral Processing System Based on UNIX Pipes. *J. Biomol. NMR* 6, 277–293. doi:10.1007/BF00197809

Frederick, K. K., Michaelis, V. K., Corzilius, B., Ong, T.-C., Jacavone, A. C., Griffin, R. G., et al. (2015). Sensitivity-enhanced NMR Reveals Alterations in Protein Structure by Cellular Milieus. *Cell* 163, 620–628. doi:10.1016/j.cell.2015.09.024

Ghosh, R., Kragelj, J., Xiao, Y., and Frederick, K. K. (2020). Cryogenic Sample Loading into a Magic Angle Spinning Nuclear Magnetic Resonance Spectrometer that Preserves Cellular Viability. *J. Visualized Experiments* 163, e61733. doi:10.3791/61733

Ghosh, R., Xiao, Y., Kragelj, J., and Frederick, K. K. (2021). In-Cell Sensitivity-Enhanced NMR of Intact Viable Mammalian Cells. *J. Am. Chem. Soc.* 143, 18454–18466. doi:10.1021/jacs.1c06680

Inomata, K., Ohno, A., Tochio, H., Isogai, S., Tenno, T., Nakase, I., et al. (2009). High-resolution Multi-Dimensional NMR Spectroscopy of Proteins in Human Cells. *Nature* 458, 106–109. doi:10.1038/nature07839

Jagtap, A. P., Krstic, I., Kunjir, N. C., Hänsel, R., Prisner, T. F., and Sigurdsson, S. T. (2015). Sterically Shielded Spin Labels for In-Cell EPR Spectroscopy: Analysis of Stability in Reducing Environment. *Free Radic. Res.* 49, 78–85. doi:10.3109/10715762.2014.979409

Jaroniec, C. P., Filip, C., and Griffin, R. G. (2002). 3D TEDOR NMR Experiments for the Simultaneous Measurement of Multiple Carbon–Nitrogen Distances in Uniformly 13C,15N-Labeled Solids. *J. Am. Chem. Soc.* 124, 10728–10742. doi:10.1021/ja026385y

Karthikeyan, G., Bonucci, A., Casano, G., Gerbaud, G., Abel, S., Thomé, V., et al. (2018). A Bioresistant Nitroxide Spin Label for In-Cell EPR Spectroscopy: *In Vitro* and in Oocytes Protein Structural Dynamics Studies. *Angew. Chem. Int. Ed.* 57, 1366–1370. doi:10.1002/anie.201710184

Lovelock, J. E., and Bishop, M. W. H. (1959). Prevention of Freezing Damage to Living Cells by Dimethyl Sulphoxide. *Nature* 183, 1394–1395. doi:10.1038/1831394a0

Luchinat, E., Barbieri, L., Campbell, T. F., and Banci, L. (2020). Real-Time Quantitative In-Cell NMR: Ligand Binding and Protein Oxidation Monitored in Human Cells Using Multivariate Curve Resolution. *Anal. Chem.* 92, 9997–10006. doi:10.1021/acs.analchem.0c01677

Lund, A., Casano, G., Menzildjian, G., Kaushik, M., Stevanato, G., Yulikov, M., et al. (2020). TinyPols: a Family of Water-Soluble Binitroxides Tailored for Dynamic Nuclear Polarization Enhanced NMR Spectroscopy at 18.8 and 21.1 T. *Chem. Sci.* 11, 2810–2818. doi:10.1039/c9sc05384k

Majumder, S., Xue, J., DeMott, C. M., Reverdatto, S., Burz, D. S., and Shekhtman, A. (2015). Probing Protein Quinary Interactions by In-Cell Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* 54, 2727–2738. doi:10.1021/acs.biochem.5b00036

McCoy, K. M., Rogawski, R., Stovicek, O., and McDermott, A. E. (2019). Stability of Nitroxide Biradical TOTAPOL in Biological Samples. *J. Magn. Reson.* 303, 115–120. doi:10.1016/j.jmr.2019.04.013

McDaniel, R. V., McIntosh, T. J., and Simon, S. A. (1983). Nonelectrolyte Substitution for Water in Phosphatidylcholine Bilayers. *Biochim. Biophys. Acta (Bba) - Biomembranes* 731, 97–108. doi:10.1016/0005-2736(83)90402-9

Narasimhan, S., Scherpe, S., Lucini Paioni, A., van der Zwan, J., Folkers, G. E., Ovaa, H., et al. (2019). DNP-supported Solid-State NMR Spectroscopy of Proteins inside Mammalian Cells. *Angew. Chem.* 131, 13103–13107. doi:10.1002/ange.201903246

Ni, Q. Z., Daviso, E., Can, T. V., Markhasin, E., Jawla, S. K., Swager, T. M., et al. (2013). High Frequency Dynamic Nuclear Polarization. *Acc. Chem. Res.* 46, 1933–1941. doi:10.1021/ar300348n

Nielsen, L., Khurana, R., Coats, A., Frokjaer, S., Brange, J., Vyas, S., et al. (2001). Effect of Environmental Factors on the Kinetics of Insulin Fibril Formation: Elucidation of the Molecular Mechanism. *Biochemistry* 40, 6036–6046. doi:10.1021/bi002555c

Notman, R., Noro, M., O'Malley, B., and Anwar, J. (2006). Molecular Basis for Dimethylsulfoxide (DMSO) Action on Lipid Membranes. *J. Am. Chem. Soc.* 128, 13982–13983. doi:10.1021/ja063363t

Overall, S. A., Price, L. E., Albert, B. J., Gao, C., Alaniva, N., Judge, P. T., et al. (2020). *In Situ* Detection of Endogenous HIV Activation by Dynamic Nuclear Polarization NMR and Flow Cytometry. *Int. J. Mol. Sci.* 21, 4649. doi:10.3390/ijms21134649

Overall, S. A., and Barnes, A. B. (2021). Biomolecular Perturbations in In-Cell Dynamic Nuclear Polarization Experiments. *Front. Mol. Biosci.* 8, 743829. doi:10.3389/fmolb.2021.743829

Pegg, D. E. (2007). Principles of Cryopreservation. *Methods Mol. Biol.* 368, 39–57. doi:10.1007/978-1-59745-362-2_3

Pinon, A. C., Schlagnitweit, J., Berruyer, P., Rossini, A. J., Lelli, M., Socie, E., et al. (2017). Measuring Nano- to Microstructures from Relayed Dynamic Nuclear Polarization NMR. *J. Phys. Chem. C* 121, 15993–16005. doi:10.1021/acs.jpcc.7b04438

Rankin, A. G. M., Trébosc, J., Pourpoint, F., Amoureux, J.-P., and Lafon, O. (2019). Recent Developments in MAS DNP-NMR of Materials. *Solid State. Nucl. Magn. Reson.* 101, 116–143. doi:10.1016/j.ssnmr.2019.05.009

Renault, M., Tommassen-van Boxtel, R., Bos, M. P., Post, J. A., Tommassen, J., and Baldus, M. (2012). Cellular Solid-State Nuclear Magnetic Resonance Spectroscopy. *Proc. Natl. Acad. Sci.* 109, 4863–4868. doi:10.1073/pnas.1116478109

Sakakibara, D., Sasaki, A., Ikeya, T., Hamatsu, J., Hanashima, T., Mishima, M., et al. (2009). Protein Structure Determination in Living Cells by in-cell NMR Spectroscopy. *Nature* 458, 102–105. doi:10.1038/nature07814

Sauvée, C., Rosay, M., Casano, G., Aussenac, F., Weber, R. T., Ouari, O., et al. (2013). Highly Efficient, Water-Soluble Polarizing Agents for Dynamic Nuclear Polarization at High Frequency. *Angew. Chem.* 125, 11058–11061. doi:10.1002/ange.201304657

Schlagnitweit, J., Friebe Sandoz, S., Jaworski, A., Guzzetti, I., Aussenac, F., Carbajo, R. J., et al. (2019). Observing an Antisense Drug Complex in Intact Human Cells by In-Cell NMR Spectroscopy. *Chembiochem* 20, 2474–2478. doi:10.1002/cbic.201900297

Schrader, A. M., Cheng, C.-Y., Israelachvili, J. N., and Han, S. (2016). Communication: Contrasting Effects of Glycerol and DMSO on Lipid Membrane Surface Hydration Dynamics and Forces. *J. Chem. Phys.* 145, 041101. doi:10.1063/1.4959904

Selenko, P., Serber, Z., Gadea, B., Ruderman, J., and Wagner, G. (2006). Quantitative NMR Analysis of the Protein G B1 Domain in Xenopus laevis Egg Extracts and Intact Oocytes. *Proc. Natl. Acad. Sci.* 103, 11904–11909. doi:10.1073/pnas.0604667103

Shaiken, T. E., and Opekun, A. R. (2014). Dissecting the Cell to Nucleus, Perinucleus and Cytosol. *Sci. Rep.* 4, 4923. doi:10.1038/srep04923

Stevanato, G., Casano, G., Kubicki, D. J., Rao, Y., Esteban Hofer, L., Menzildjian, G., et al. (2020). Open and Closed Radicals: Local Geometry Around Unpaired Electrons Governs Magic-Angle Spinning Dynamic Nuclear Polarization Performance. *J. Am. Chem. Soc.* 142, 16587–16599. doi:10.1021/jacs.0c04911

Theillet, F.-X., Binolfi, A., Bekei, B., Martorana, A., Rose, H. M., Stuiver, M., et al. (2016). Structural Disorder of Monomeric α-synuclein Persists in Mammalian Cells. *Nature* 530, 45–50. doi:10.1038/nature16531

Theillet, F.-X., Rose, H. M., Liokatis, S., Binolfi, A., Thongwichian, R., Stuiver, M., et al. (2013). Site-specific NMR Mapping and Time-Resolved Monitoring of Serine and Threonine Phosphorylation in Reconstituted Kinase Reactions and Mammalian Cell Extracts. *Nat. Protoc.* 8, 1416–1432. doi:10.1038/nprot.2013.083

Williams, A. C., and Barry, B. W. (2004). Penetration Enhancers. *Adv. Drug Deliv. Rev.* 56, 603–618. doi:10.1016/j.addr.2003.10.025

Yu, Z. W., and Quinn, P. J. (1995). Phase Stability of Phosphatidylcholines in Dimethylsulfoxide Solutions. *Biophysical J.* 69, 1456–1463. doi:10.1016/s0006-3495(95)80015-9

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

BRIEF RESEARCH REPORT
published: 14 February 2022
doi: 10.3389/fmolb.2022.841790

Influence of the Dynamically Disordered N-Terminal Tail Domain on the Amyloid Core Structure of Human Y145Stop Prion Protein Fibrils

*Zhe Qi[1], Krystyna Surewicz[2], Witold K. Surewicz[2] and Christopher P. Jaroniec[1]**

[1]Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH, United States, [2]Department of Physiology and Biophysics, Case Western Reserve University, Cleveland, OH, United States

The Y145Stop mutant of human prion protein (huPrP23-144) is associated with a familial prionopathy and provides a convenient *in vitro* model for investigating amyloid strains and cross-seeding barriers. huPrP23-144 fibrils feature a compact and relatively rigid parallel in-register β-sheet amyloid core spanning ~30 C-terminal amino acid residues (~112–141) and a large ~90-residue dynamically disordered N-terminal tail domain. Here, we systematically evaluate the influence of this dynamic domain on the structure adopted by the huPrP23-144 amyloid core region, by investigating using magic-angle spinning solid-state nuclear magnetic resonance (NMR) spectroscopy a series of fibril samples formed by huPrP23-144 variants corresponding to deletions of large segments of the N-terminal tail. We find that deletion of the bulk of the N-terminal tail, up to residue 98, yields amyloid fibrils with native-like huPrP23-144 core structure. Interestingly, deletion of additional flexible residues in the stretch 99–106 located outside of the amyloid core yields shorter heterogenous fibrils with fingerprint NMR spectra that are clearly distinct from those for full-length huPrP23-144, suggestive of the onset of perturbations to the native structure and degree of molecular ordering for the core residues. For the deletion variant missing residues 99–106 we show that native huPrP23-144 core structure can be "restored" by seeding the fibril growth with preformed full-length huPrP23-144 fibrils.

Keywords: amyloid, prion, octarepeat, intrinsically disordered region/protein, magic angle spinning (MAS) solid-state NMR

OPEN ACCESS

Edited by:
Vincenzo Venditti,
Iowa State University, United States

Reviewed by:
Benjamin James Wylie,
Texas Tech University, United States
Marvin J. Bayro,
University of Puerto Rico, Puerto Rico

****Correspondence:***
Christopher P. Jaroniec
jaroniec.1@osu.edu

Specialty section:
This article was submitted to Biophysics, a section of the journal Frontiers in Molecular Biosciences

Received: *22 December 2021*
Accepted: *26 January 2022*
Published: *14 February 2022*

Citation:
Qi Z, Surewicz K, Surewicz WK and Jaroniec CP (2022) Influence of the Dynamically Disordered N-Terminal Tail Domain on the Amyloid Core Structure of Human Y145Stop Prion Protein Fibrils. Front. Mol. Biosci. 9:841790. doi: 10.3389/fmolb.2022.841790

INTRODUCTION

Most peptide and protein molecules are capable of undergoing conformational conversion from their native state into highly ordered, β-sheet rich amyloid fibrils (Dobson, 1999), and for ~50 human proteins such misfolding and amyloid formation can occur under physiological conditions *in vivo* leading to development of disease (Chiti and Dobson, 2006). A number of amyloids have been found to contain large dynamically disordered domains flanking the structured fibril core (Heise et al., 2005; Siemer et al., 2006; Loquet et al., 2009; Helmus et al., 2010; Bibow et al., 2011; Li et al., 2012; Raveendra et al., 2013; Frederick et al., 2014; Isas et al., 2015; Cervantes et al., 2016; Lin et al., 2017; Murray et al., 2017; Caulkins et al., 2018; Dregni et al., 2019; Fonda et al., 2021), and it has been suggested that the presence of these conformationally flexible domains may be of pathological or functional significance by stabilizing fibril structures and mediating interactions involving

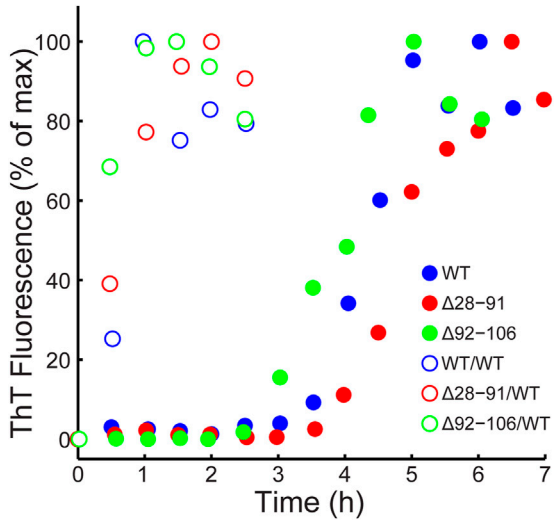


FIGURE 1 | Kinetics of amyloid formation monitored by thioflavin T fluorescence in the absence (filled circles) and presence (open circles) of 2% (mole fraction) of WT huPrP23-144 fibril seeds for WT (blue), Δ28-91 (red) and Δ92-106 (green) huPrP23-144.

protofilaments (Uversky and Fink, 2004; Chiti and Dobson, 2006; Tompa, 2009; van der Wel, 2017; Siemer, 2020). The detailed characterization of dynamically disordered regions in amyloids (and in other large biomacromolecular assemblies) has generally been pursued by multidimensional magic-angle spinning (MAS) nuclear magnetic resonance (NMR) techniques, which are able to visualize these domains directly in hydrated samples at ambient temperature by using experiments based on scalar coupling mediated polarization transfers (Tycko, 2006; van der Wel, 2017; Jaroniec, 2019; Siemer, 2020).

The C-terminally truncated Y145Stop prion protein (PrP23-144) variant is associated with a hereditary prionopathy in humans (Ghetti et al., 1996), and mouse PrP23-144 amyloid fibrils have recently been shown to cause transmissible prion disease in mice (Choi et al., 2016). Importantly, the highly homologous human (hu), mouse (mo) and Syrian hamster (Sha) PrP23-144 proteins (pairwise amino acid, aa, sequence identities of ~90–95%) have also been shown to provide a valuable *in vitro* model for detailed investigation of the structural basis of amyloid strains and transmissibility barriers (Kundu et al., 2003; Vanik et al., 2004; Jones and Surewicz, 2005; Surewicz et al., 2006). Our previous structural and dynamic solid-state NMR studies of huPrP23-144 fibrils revealed the presence of a structured ~30-residue parallel in-register β-amyloid core (aa ~112–141) exhibiting limited protein backbone motions on the ~0.1–1 ms time scale located near the C-terminus and a large dynamically disordered ~90-residue N-terminal tail domain (aa ~23–110) (Helmus et al., 2008; Helmus et al., 2010; Helmus et al., 2011; Theint et al., 2018; Aucoin et al., 2019; Shannon et al., 2019). Additional studies of PrP23-144 amyloids containing mutations and deletions corresponding to different huPrP23-144 core residues enabled these sequence modifications to be correlated with structural and dynamic changes in the PrP23-144 amyloid core and provided initial insights into mammalian PrP23-144 cross-seeding specificities (Jones et al., 2011; Theint et al., 2017a; Theint et al., 2017b; Dao et al., 2021).

Previous studies of amyloids formed by full-length prion protein (PrP23-231) and the proteinase-K resistant 90–231 fragment of transmissible spongiform encephalopathy associated mammalian PrP deposits (Prusiner, 1998) suggest that the flexible N-terminal domain may play a role in PrP aggregation properties, and prion structure and pathogenesis (Weissmann, 1999; Lawson et al., 2004; Baskakov and Bocharova, 2005; Frankenfield et al., 2005). The present study aims to assess the influence of the dynamic huPrP23-144 N-terminal region on amyloid assembly and resulting β-core conformation. This is achieved by performing systematic solid-state NMR, atomic force microscopy (AFM) and thioflavin T (ThT) fluorescence studies on fibril samples formed *in vitro* from recombinant huPrP23-144 variants corresponding to deletions of large segments of the N-terminal tail. Overall, we find that the majority of dynamically disordered N-terminal tail residues, including the octarepeat region (aa 51–91) implicated in copper binding and homeostasis (Millhauser, 2007; Aguzzi et al., 2008), have little impact on the fibril assembly kinetics and ability of the deletion variants to adopt the native huPrP23-144 amyloid core structure. However, we also find that a stretch of ~10 conformationally flexible residues that precede the amyloid core region in huPrP23-144 fibrils appears to play a role in the ability to adopt the native core structure and the degree of molecular ordering within the core.

RESULTS

Previous solid-state NMR studies indicate that the relatively rigid β-core region of huPrP23-144 fibrils consists of residues 112–141 (Helmus et al., 2008; Helmus et al, 2010; Helmus et al, 2011; Shannon et al., 2019). In contrast residues 23–111 and 142–144 are not observable in conventional cross-polarization magic angle spinning (CP-MAS) solid-state NMR spectra that utilize dipolar coupling-based polarization transfers, consistent with their increased mobility (van der Wel, 2017; Siemer, 2020) while most of these residues can be detected in MAS NMR spectra utilizing polarization transfers mediated via J-couplings (van der Wel, 2017; Siemer, 2020). To assess the potential influence of the dynamically disordered N-terminal domain of huPrP23-144 on the conformation adopted by the amyloid core region we generated a series of fibril samples from large N-terminal domain deletion variants of huPrP23-144 and examined their fibrillization kinetics, morphologies and molecular conformations by using ThT fluorescence, AFM and solid-state NMR, respectively.

The huPrP23-144 deletion variants employed in these studies spanned residues 28–106—note that the short segment (aa 23–27) containing multiple lysine and arginine residues (as well as the N-terminal GDSP extension present in our huPrP23-144 construct (Helmus et al., 2008) was not deleted in order to ensure the solubility of the different deletion variants. Initially, we investigated the following huPrP23-144 variants:

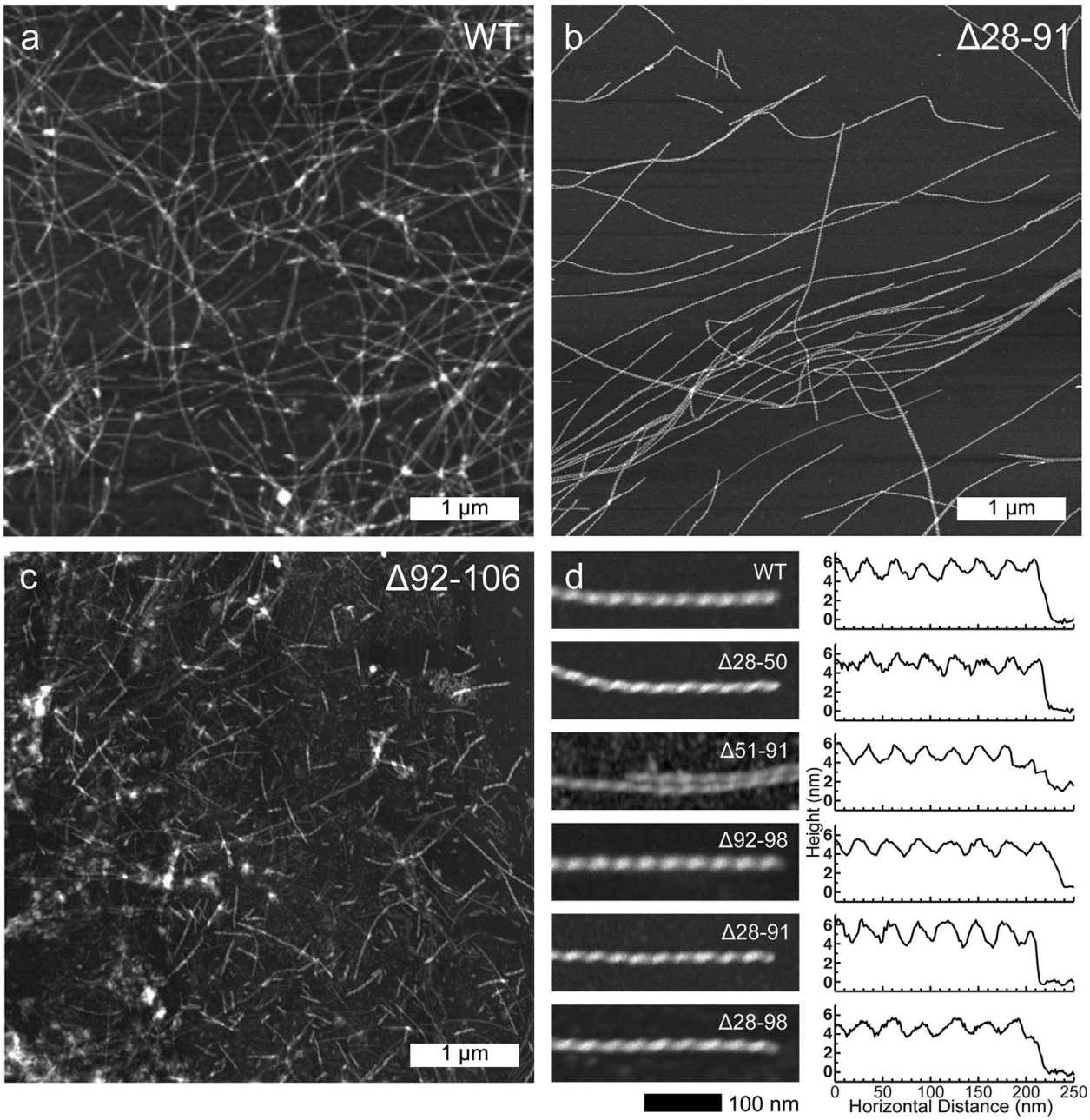


FIGURE 2 | Representative low-resolution atomic force microscopy images for **(A)** WT, **(B)** Δ28-91 and **(C)** Δ92-106 huPrP23-144 fibrils (scale bar 1 μm for all panels). **(D)** High-resolution atomic force microscopy images (left; scale bar 100 nm for all panels) and corresponding height profiles (right) for WT, Δ28-50, Δ51-91, Δ92-98, Δ28-91, and Δ28-98 huPrP23-144 fibrils as indicated in the insets.

Δ28-50, Δ51-91 (corresponding to deletion of the entire octarepeat region) and Δ92-106 (see **Figure 3A** for the huPrP23-144 protein sequence and summary of the deletion variants studied). Briefly, fibrils generated from the Δ28-50 and Δ51-91 constructs were found to exhibit wild-type (WT) like morphologies and molecular conformations, while significant differences relative to WT were observed for the Δ92-106 fibrils. To investigate this further we prepared the Δ92-98 variant, which was found to form WT-like fibrils suggesting that deletion of huPrP23-144 N-terminal residues up to aa 98 does not have a significant impact on formation of the native huPrP23-144 amyloid core structure. Based on these findings we then generated the Δ28-98 variant corresponding to the deletion of nearly the entire huPrP23-144 N-terminal tail. While the Δ28-98 construct expressed at reasonable level in rich medium it was found not to express at sufficiently high level in ^{13}C and ^{15}N isotope enriched minimal medium to permit multidimensional solid-state NMR studies, which led us to generate an additional,

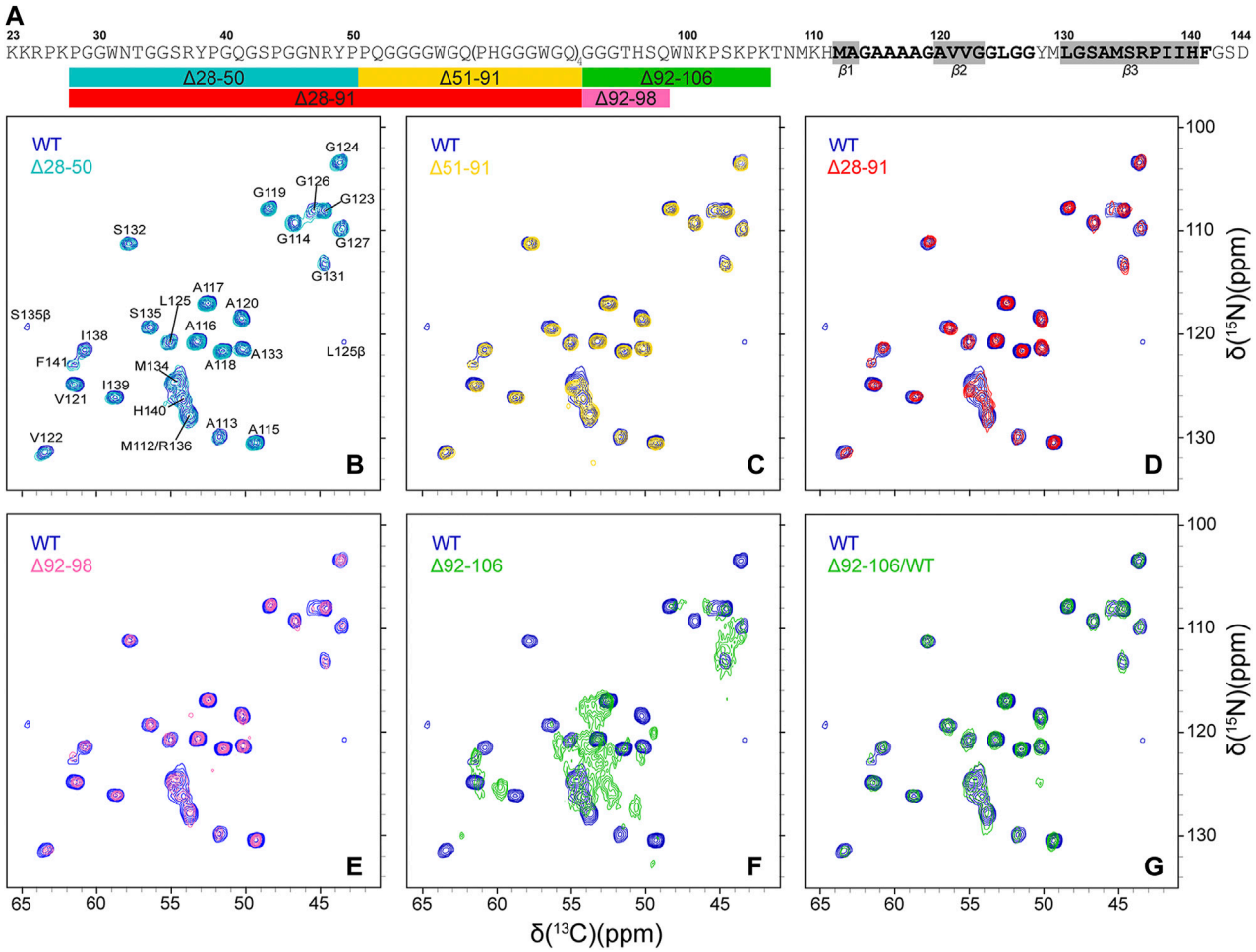


FIGURE 3 | (A) Amino acid sequence of huPrP23-144. Relatively rigid residues comprising the amyloid core detected in conventional solid-state NMR experiments are shown in bold black font, with regions having highest β-strand propensity highlighted in grey rectangles (Theint et al., 2017b), and conformationally flexible residues (Helmus et al., 2008; Helmus et al., 2010) are shown in grey font. The rectangles below the amino acid sequence schematically show the key huPrP23-144 large deletion variants that were investigated in this study, with the colors of the rectangles corresponding to the colors of the contours in the solid-state NMR spectra in panels **(B)**-**(G)** **(B–G)** 2D ^{15}N-^{13}Cα chemical shift correlation solid-state NMR spectra of Δ28-50 (b, cyan), Δ51-91 (c, gold), Δ28-91 (d, red), Δ92-98 (e, magenta), Δ92-106 (f, green) huPrP23-144 fibrils and Δ92-106 fibrils seeded with WT huPrP23-144 amyloid (g, green), overlaid with the reference spectrum for WT huPrP23-144 fibrils (blue).

Δ28-91, deletion variant. The studies of all the aforementioned huPrP23-144 N-terminal deletion variants are described in additional detail below.

As noted above, all the huPrP23-144 deletion variants investigated in this study (Δ28-50, Δ28-91, Δ28-98, Δ51-91, Δ92-98 and Δ92-106) readily converted to amyloid fibrils in autocatalytic, unseeded reactions carried out in potassium phosphate buffer at pH 6.4. The kinetics of fibril formation were monitored by the standard ThT binding assay (Naiki et al., 1989) revealing a nearly identical ~3-4 h lag phase for WT huPrP23-144 and all the deletion variants at 400 μM protein concentration, in line with the value reported previously for WT huPrP23-144 (Kundu et al., 2003). Representative data for WT, Δ28-91 and Δ92-106 huPrP23-144 are shown in **Figure 1**. Furthermore, we found that addition of a small amount of pre-formed WT huPrP23-144 fibril seeds to the reaction resulted in complete elimination of the lag phase for all deletion variants studied (**Figure 1**).

Atomic force microscopy was then used to investigate the morphologies of the resulting amyloid fibrils. With exception of the Δ92-106 fibrils, all the other deletion variants displayed morphologies that were similar to one another as well as to the morphology of WT huPrP23-144 amyloid. Specifically, as shown in **Figure 2**, these fibrils had highly uniform, micron long, threadlike morphologies with left-handed twist characterized by heights of ~5–6 nm and periodicities of ~30 nm. In contrast, the Δ92-106 fibril sample was far more heterogeneous, containing shorter fibrils of varying lengths in the range of tens to hundreds of nanometers as well as considerable amounts of apparently amorphous, non-fibrillar aggregates (**Figure 2C**).

Finally, in order to compare the protein conformations and degree of molecular ordering for the different fibril samples at the atomic level, we recorded two-dimensional (2D) fingerprint ^{15}N-^{13}Cα chemical shift correlation solid-state NMR spectra for wild-type huPrP23-144 amyloid and all the deletion variants except Δ28-98.

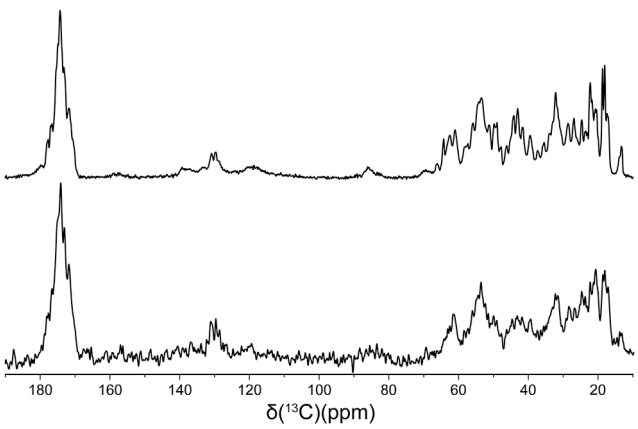


FIGURE 4 | ^{13}C CP-MAS spectra of ^{13}C,^{15}N WT (top) and natural abundance Δ28-98 (bottom) huPrP23-144 amyloid fibrils.

Figure 3 shows the resulting 2D ^{15}N-^{13}Cα spectra, which were found to be effectively identical in terms of NMR signal frequencies, linewidths and relative intensities for WT huPrP23-144 and all deletion variants with exception of Δ92-106 fibrils; the latter exhibited considerable spectral differences relative to the other samples (**Figure 3F**). Note that for fibrils formed from the Δ28-98 variant, which did not express at a sufficiently high level in ^{13}C and ^{15}N isotope labeled minimal media to permit collection of a ^{15}N-^{13}Cα correlation spectrum, a 1D ^{13}C CP-MAS spectrum could be recorded for unlabeled fibrils (**Figure 4**). Apart from the obviously lower sensitivity associated with the use of a natural abundance sample, this spectrum showed considerable similarity to reference ^{13}C CP-MAS spectrum recorded for ^{13}C,^{15}N-labeled WT huPrP23-144 fibrils. Altogether these data indicate that other than Δ92-106, all the huPrP23-144 deletion variants studied form amyloids that are highly ordered at the atomic level and possess WT-like core structures and core residue backbone motions. On the other hand, the spectra of Δ92-106 fibrils are indicative of structural perturbation/polymorphism for the core residues and a higher degree of molecular disorder, consistent with the higher degree of sample heterogeneity observed by AFM. Comparison of 2D ^{15}N-^{13}Cα spectra for WT and Δ92-106 fibrils reveals that they key chemical shift perturbations involve residues ~112–115. This suggests that conformations of the β1-strand and several following core residues are primarily impacted in the deletion variant, in line with earlier studies (Theint et al., 2017a) indicating that several residues in this regime appear to be key for stabilizing the core fold of WT huPrP23-144 amyloid. Remarkably, however, we also find that seeding Δ92-106 amyloid formation with pre-formed WT huPrP23-144 fibrils results in the structural adaptation of the Δ92-106 protein to the WT huPrP23-144 amyloid core fold as evidenced by their indistinguishable 2D ^{15}N-^{13}Cα spectra (**Figure 3G**).

DISCUSSION

Collectively, on the basis of effectively identical fibril assembly kinetics, morphologies and fingerprint solid-state NMR spectra for all huPrP23-144 variants studied containing large deletions up to residue 98, our results indicate that the bulk of the dynamically disordered N-terminal tail domain of huPrP23-144 is not essential for amyloid formation under autocatalytic conditions and ability of the resulting β-core region to adopt a WT-like structure. Remarkably, this finding is strongly correlated with our previously reported data showing that proteinase-K resistant fragments of WT huPrP23-144 amyloid fibrils span residues 97–144, 98–144 and 99–144 (Jones et al., 2011).

Combined with the finding that the protein conformation and degree of molecular ordering appear to be significantly perturbed for the Δ92-106 deletion variant relative to WT huPrP23-144 and the different large huPrP23-144 deletion variants up to residue 98, our data indicate that aa ~99–106 (and presumably several additional amino acids preceding the structured and relatively rigid fibril core beginning around residue Met-112) play a key role in stabilizing the formation of the characteristic huPrP23-144 β-core structure in spite of their flexible nature and location outside of the structured amyloid core region. Given that the 100–110 stretch of huPrP23-144 contains four positively charged lysine residues (at positions 101, 104, 106 and 110) we speculate that the stabilization of the amyloid core structure in WT huPrP23-144 occurs via electrostatic interactions with the negatively charged C-terminal aspartate residue.

While deletion of amino acids in the 99–106 regime clearly impacts the amyloid core structure and molecular ordering of huPrP23-144 within fibrils formed under autocatalytic conditions as discussed above, we also find that the seeding of fibril formation by the Δ92-106 deletion variant with pre-formed WT huPrP23-144 fibrils at low concentration leads to the Δ92-106 proteins adopting a WT-like core structure at the atomic level as revealed by the fingerprint ^{15}N-^{13}Cα NMR spectrum that is effectively identical to the corresponding spectrum for WT huPrP23-144 amyloid. Interestingly, this structural templating process that yields a WT-like fold for Δ92-106 amyloid was found to not appreciably alter the overall morphology of the Δ92-106 fibrils as viewed by AFM, with the majority of the sample consisting of relatively short fibrils similar to those observed in the unseeded reaction.

In summary, we systematically evaluated the influence of the dynamically disordered N-terminal tail domain of huPrP23-144 on the structure adopted by the amyloid core region by using deletion mutagenesis combined with magic-angle spinning solid-state NMR spectroscopy. We find that N-terminal huPrP23-144 residues up to aa 98, which coincide with the protein segment most susceptible to proteinase-K digestion in mature fibrils, can be deleted without impacting the core structure formed in an autocatalytic fibril assembly process. Remarkably, deletion of additional flexible residues (aa 99–106) located outside the amyloid core leads to the formation of amyloid fibrils with a perturbed core structure and reduced degree of molecular ordering in the fibril lattice, most likely caused by the disruption of stabilizing electrostatic interactions involving several lysine side-chains found in this region and the C-terminal aspartate residue. These structural perturbations, however, can be alleviated by catalyzing the amyloid formation with preformed WT huPrP23-144 fibril seeds.

MATERIALS AND METHODS

Protein Expression and Purification

The plasmid encoding human PrP23-144 was described previously (Kundu et al., 2003) and plasmids encoding huPrP23-144 deletion variants were generated similarly to our previous study (Jones et al., 2011) by deletion mutagenesis using a QuikChange kit (Stratagene). Uniformly ^{13}C,^{15}N labeled huPrP23-144 and deletion variants were expressed in *E. coli* BL21 (DE3) cells and purified by nickel affinity chromatography as described in detail in previous studies (Theint et al., 2017b; Dao et al., 2021). The identities and purities of the resulting proteins were routinely confirmed by SDS/PAGE and MALDI mass spectrometry.

Amyloid Fibril Formation

Amyloid fibrils were prepared under quiescent conditions at 25°C as described in previous studies (Theint et al., 2017a; Theint et al., 2017b; Dao et al., 2021). Lyophilized huPrP23-144 variants were dissolved in ultrapure water at ~400 μM concentration and 1 M potassium phosphate buffer at pH 6.4 was added to a final concentration of 50 mM. For samples seeded with WT huPrP23-144 amyloid, 2% (mole fraction) of preformed fibrils was added immediately to the reaction after addition of phosphate buffer.

Thioflavin T Fluorescence and Atomic Force Microscopy

Kinetics of fibril formation were monitored via the standard ThT fluorescence assay (Naiki et al., 1989) as described in detail in previous studies (Theint et al., 2018). Fibril morphologies were assessed by atomic force microscopy (AFM) as follows. Fibril suspensions were diluted 50-fold in ultrapure water, deposited on freshly cleaved mica substrates (Ted Pella Inc.) for 5 min, rinsed with three 50 μL aliquots of ultrapure water, and allowed to air dry for 1-2 h prior to imaging. Images were collected using a Bruker Dimension Icon AFM in PeakForce quantitative nanomechanical mapping mode with a ScanAsyst-Air probe and processed with the Bruker NanoScope Analysis software.

Solid-State NMR Spectroscopy

Fibril suspensions for solid-state NMR analysis were incubated as described above for 48 h and centrifuged. The fibril pellets were washed three times with 50 mM potassium phosphate pH 6.4 buffer and packed into 3.2 mm zirconia rotors by centrifugation with the final samples containing ~5–10 mg of fibrils. Standard 1D ^{13}C CP-MAS and 2D ^{15}N-^{13}Cα solid-state NMR spectra were recorded on 500 MHz Varian and 800 MHz Bruker spectrometers, equipped with 3.2 mm triple resonance (^{1}H-^{13}C-^{15}N) BioMAS and E^{free} probes, respectively. The sample temperature and MAS frequencies were actively regulated at 5°C and 11.111 kHz, respectively, and the experimental parameters were similar to those used in our previous solid-state NMR studies of WT huPrP23-144 fibrils (Helmus et al., 2008) with data acquisition times for the 2D ^{15}N-^{13}Cα NMR spectra of 8–16 h per sample. NMR spectra were processed and analyzed using NMRPipe (Delaglio et al., 1995) and Sparky (Goddard and Kneller, 2006), respectively.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

WS and CJ designed the study. ZQ and KS prepared samples and performed experiments. ZQ prepared figures. ZQ, WS, and CJ wrote the manuscript.

FUNDING

This work was supported by grants from NIH (R01GM094357 to CPJ and WKS and S10OD012303 to CPJ) and NSF (MCB-1715174 to CPJ).

ACKNOWLEDGMENTS

We thank Dr. Sean Smrt for stimulating discussions and insightful comments on the manuscript.

REFERENCES

Aguzzi, A., Baumann, F., and Bremer, J. (2008). The Prion's Elusive Reason for Being. *Annu. Rev. Neurosci.* 31, 439–477. doi:10.1146/annurev.neuro.31.060407.125620

Aucoin, D., Xia, Y., Theint, T., Nadaud, P. S., Surewicz, K., Surewicz, W. K., et al. (2019). Protein-solvent Interfaces in Human Y145Stop Prion Protein Amyloid Fibrils Probed by Paramagnetic Solid-State NMR Spectroscopy. *J. Struct. Biol.* 206 (1), 36–42. doi:10.1016/j.jsb.2018.04.002

Baskakov, I. V., and Bocharova, O. V. (2005). *In Vitro* conversion of Mammalian Prion Protein into Amyloid Fibrils Displays Unusual Features. *Biochemistry* 44 (7), 2339–2348. doi:10.1021/bi048322t

Bibow, S., Mukrasch, M. D., Chinnathambi, S., Biernat, J., Griesinger, C., Mandelkow, E., et al. (2011). The Dynamic Structure of Filamentous Tau. *Angew. Chem. Int. Ed.* 50 (48), 11520–11524. doi:10.1002/anie.201105493

Caulkins, B. G., Cervantes, S. A., Isas, J. M., and Siemer, A. B. (2018). Dynamics of the Proline-Rich C-Terminus of Huntingtin Exon-1 Fibrils. *J. Phys. Chem. B* 122 (41), 9507–9515. doi:10.1021/acs.jpcb.8b09213

Cervantes, S. A., Bajakian, T. H., Soria, M. A., Falk, A. S., Service, R. J., Langen, R., et al. (2016). Identification and Structural Characterization of the N-Terminal Amyloid Core of Orb2 Isoform A. *Sci. Rep.* 6, 38265. doi:10.1038/srep38265

Chiti, F., and Dobson, C. M. (2006). Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* 75, 333–366. doi:10.1146/annurev.biochem.75.101304.123901

Choi, J.-K., Cali, I., Surewicz, K., Kong, Q., Gambetti, P., and Surewicz, W. K. (2016). Amyloid Fibrils from the N-Terminal Prion Protein Fragment Are Infectious. *Proc. Natl. Acad. Sci. USA* 113 (48), 13851–13856. doi:10.1073/pnas.1610716113

Dao, H. H., Hlaing, M. Z., Ma, Y., Surewicz, K., Surewicz, W. K., and Jaroniec, C. P. (2021). 13C and 15N Chemical Shift Assignments of A117V and M129V Human Y145Stop Prion Protein Amyloid Fibrils. *Biomol. NMR Assign.* 15 (1), 45–51. doi:10.1007/s12104-020-09981-4

Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *J. Biomol. NMR* 6 (3), 277–293. doi:10.1007/BF00197809

Dobson, C. M. (1999). Protein Misfolding, Evolution and Disease. *Trends Biochem. Sci.* 24 (9), 329–332. doi:10.1016/s0968-0004(99)01445-0

Dregni, A. J., Mandala, V. S., Wu, H., Elkins, M. R., Wang, H. K., Hung, I., et al. (2019). *In Vitro* 0N4R Tau Fibrils Contain a Monomorphic β-sheet Core Enclosed by Dynamically Heterogeneous Fuzzy Coat Segments. *Proc. Natl. Acad. Sci. USA* 116 (33), 16357–16366. doi:10.1073/pnas.1906839116

Fonda, B. D., Jami, K. M., Boulos, N. R., and Murray, D. T. (2021). Identification of the Rigid Core for Aged Liquid Droplets of an RNA-Binding Protein Low Complexity Domain. *J. Am. Chem. Soc.* 143 (17), 6657–6668. doi:10.1021/jacs.1c02424

Frankenfield, K. N., Powers, E. T., and Kelly, J. W. (2005). Influence of the N-Terminal Domain on the Aggregation Properties of the Prion Protein. *Protein Sci.* 14 (8), 2154–2166. doi:10.1110/ps.051434005

Frederick, K. K., Debelouchina, G. T., Kayatekin, C., Dorminy, T., Jacavone, A. C., Griffin, R. G., et al. (2014). Distinct Prion Strains Are Defined by Amyloid Core Structure and Chaperone Binding Site Dynamics. *Chem. Biol.* 21 (2), 295–305. doi:10.1016/j.chembiol.2013.12.013

Ghetti, B., Piccardo, P., Spillantini, M. G., Ichimiya, Y., Porro, M., Perini, F., et al. (1996). Vascular Variant of Prion Protein Cerebral Amyloidosis with Tau-Positive Neurofibrillary Tangles: the Phenotype of the Stop Codon 145 Mutation in PRNP. *Proc. Natl. Acad. Sci.* 93 (2), 744–748. doi:10.1073/pnas.93.2.744

Goddard, T., and Kneller, D. (2006). *SPARKY 3*. San Francisco: University of Califronia.

Heise, H., Hoyer, W., Becker, S., Andronesi, O. C., Riedel, D., and Baldus, M. (2005). Molecular-level Secondary Structure, Polymorphism, and Dynamics of Full-Length -synuclein Fibrils Studied by Solid-State NMR. *Proc. Natl. Acad. Sci.* 102 (44), 15871–15876. doi:10.1073/pnas.0506109102

Helmus, J. J., Surewicz, K., Apostol, M. I., Surewicz, W. K., and Jaroniec, C. P. (2011). Intermolecular Alignment in Y145Stop Human Prion Protein Amyloid Fibrils Probed by Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 133 (35), 13934–13937. doi:10.1021/ja206469q

Helmus, J. J., Surewicz, K., Nadaud, P. S., Surewicz, W. K., and Jaroniec, C. P. (2008). Molecular Conformation and Dynamics of the Y145Stop Variant of Human Prion Protein in Amyloid Fibrils. *Proc. Natl. Acad. Sci.* 105 (17), 6284–6289. doi:10.1073/pnas.0711716105

Helmus, J. J., Surewicz, K., Surewicz, W. K., and Jaroniec, C. P. (2010). Conformational Flexibility of Y145Stop Human Prion Protein Amyloid Fibrils Probed by Solid-State Nuclear Magnetic Resonance Spectroscopy. *J. Am. Chem. Soc.* 132 (7), 2393–2403. doi:10.1021/ja909827v

Isas, J. M., Langen, R., and Siemer, A. B. (2015). Solid-state Nuclear Magnetic Resonance on the Static and Dynamic Domains of Huntingtin Exon-1 Fibrils. *Biochemistry* 54 (25), 3942–3949. doi:10.1021/acs.biochem.5b00281

Jaroniec, C. P. (2019). Two Decades of Progress in Structural and Dynamic Studies of Amyloids by Solid-State NMR. *J. Magn. Reson.* 306, 42–47. doi:10.1016/j.jmr.2019.07.015

Jones, E. M., and Surewicz, W. K. (2005). Fibril Conformation as the Basis of Species- and Strain-dependent Seeding Specificity of Mammalian Prion Amyloids. *Cell* 121 (1), 63–72. doi:10.1016/j.cell.2005.01.034

Jones, E. M., Wu, B., Surewicz, K., Nadaud, P. S., Helmus, J. J., Chen, S., et al. (2011). Structural Polymorphism in Amyloids. *J. Biol. Chem.* 286 (49), 42777–42784. doi:10.1074/jbc.M111.302539

Kundu, B., Maiti, N. R., Jones, E. M., Surewicz, K. A., Vanik, D. L., and Surewicz, W. K. (2003). Nucleation-dependent Conformational Conversion of the Y145Stop Variant of Human Prion Protein: Structural Clues for Prion Propagation. *Proc. Natl. Acad. Sci.* 100 (21), 12069–12074. doi:10.1073/pnas.2033281100

Lawson, V. A., Priola, S. A., Meade-White, K., Lawson, M., and Chesebro, B. (2004). Flexible N-Terminal Region of Prion Protein Influences Conformation of Protease-Resistant Prion Protein Isoforms Associated with Cross-Species Scrapie Infection *In Vivo* and *In Vitro*. *J. Biol. Chem.* 279 (14), 13689–13695. doi:10.1074/jbc.M303697200

Li, J., McQuade, T., Siemer, A. B., Napetschnig, J., Moriwaki, K., Hsiao, Y.-S., et al. (2012). The RIP1/RIP3 Necrosome Forms a Functional Amyloid Signaling Complex Required for Programmed Necrosis. *Cell* 150 (2), 339–350. doi:10.1016/j.cell.2012.06.019

Lin, H.-K., Boatz, J. C., Krabbendam, I. E., Kodali, R., Hou, Z., Wetzel, R., et al. (2017). Fibril Polymorphism Affects Immobilized Non-amyloid Flanking Domains of Huntingtin Exon1 rather Than its Polyglutamine Core. *Nat. Commun.* 8, 15462. doi:10.1038/ncomms15462

Loquet, A., Bousset, L., Gardiennet, C., Sourigues, Y., Wasmer, C., Habenstein, B., et al. (2009). Prion Fibrils of Ure2p Assembled under Physiological Conditions Contain Highly Ordered, Natively Folded Modules. *J. Mol. Biol.* 394 (1), 108–118. doi:10.1016/j.jmb.2009.09.016

Millhauser, G. L. (2007). Copper and the Prion Protein: Methods, Structures, Function, and Disease. *Annu. Rev. Phys. Chem.* 58, 299–320. doi:10.1146/annurev.physchem.58.032806.104657

Murray, D. T., Kato, M., Lin, Y., Thurber, K. R., Hung, I., McKnight, S. L., et al. (2017). Structure of FUS Protein Fibrils and its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell* 171 (3), 615–627. e16. doi:10.1016/j.cell.2017.08.048

Naiki, H., Higuchi, K., Hosokawa, M., and Takeda, T. (1989). Fluorometric Determination of Amyloid Fibrils *In Vitro* Using the Fluorescent Dye, Thioflavine T. *Anal. Biochem.* 177 (2), 244–249. doi:10.1016/0003-2697(89)90046-8

Prusiner, S. B. (1998). Nobel Lecture: Prions. *Proc. Natl. Acad. Sci.* 95 (23), 13363–13383. doi:10.1073/pnas.95.23.13363

Raveendra, B. L., Siemer, A. B., Puthanveettil, S. V., Hendrickson, W. A., Kandel, E. R., and McDermott, A. E. (2013). Characterization of Prion-like Conformational Changes of the Neuronal Isoform of Aplysia CPEB. *Nat. Struct. Mol. Biol.* 20 (4), 495–501. doi:10.1038/nsmb.2503

Shannon, M. D., Theint, T., Mukhopadhyay, D., Surewicz, K., Surewicz, W. K., Marion, D., et al. (2019). Conformational Dynamics in the Core of Human Y145Stop Prion Protein Amyloid Probed by Relaxation Dispersion NMR. *ChemPhysChem* 20 (2), 311–317. doi:10.1002/cphc.201800779

Siemer, A. B. (2020). Advances in Studying Protein Disorder with Solid-State NMR. *Solid State. Nucl. Magn. Reson.* 106, 101643. doi:10.1016/j.ssnmr.2020.101643

Siemer, A. B., Arnold, A. A., Ritter, C., Westfeld, T., Ernst, M., Riek, R., et al. (2006). Observation of Highly Flexible Residues in Amyloid Fibrils of the HET-S Prion. *J. Am. Chem. Soc.* 128 (40), 13224–13228. doi:10.1021/ja063639x

Surewicz, W. K., Jones, E. M., and Apetri, A. C. (2006). The Emerging Principles of Mammalian Prion Propagation and Transmissibility Barriers: Insight from Studies *In Vitro*. *Acc. Chem. Res.* 39 (9), 654–662. doi:10.1021/ar050226c

Theint, T., Nadaud, P. S., Aucoin, D., Helmus, J. J., Pondaven, S. P., Surewicz, K., et al. (2017b). Species-dependent Structural Polymorphism of Y145Stop Prion Protein Amyloid Revealed by Solid-State NMR Spectroscopy. *Nat. Commun.* 8 (1), 753. doi:10.1038/s41467-017-00794-z

Theint, T., Nadaud, P. S., Surewicz, K., Surewicz, W. K., and Jaroniec, C. P. (2017a). 13C and 15N Chemical Shift Assignments of Mammalian Y145Stop Prion Protein Amyloid Fibrils. *Biomol. NMR Assign.* 11 (1), 75–80. doi:10.1007/s12104-016-9723-6

Theint, T., Xia, Y., Nadaud, P. S., Mukhopadhyay, D., Schwieters, C. D., Surewicz, K., et al. (2018). Structural Studies of Amyloid Fibrils by Paramagnetic Solid-State Nuclear Magnetic Resonance Spectroscopy. *J. Am. Chem. Soc.* 140 (41), 13161–13166. doi:10.1021/jacs.8b06758

Tompa, P. (2009). Structural Disorder in Amyloid Fibrils: its Implication in Dynamic Interactions of Proteins. *FEBS J.* 276 (19), 5406–5415. doi:10.1111/j.1742-4658.2009.07250.x

Tycko, R. (2006). Molecular Structure of Amyloid Fibrils: Insights from Solid-State NMR. *Quart. Rev. Biophys.* 39 (1), 1–55. doi:10.1017/S0033583506004173

Uversky, V. N., and Fink, A. L. (2004). Conformational Constraints for Amyloid Fibrillation: The Importance of Being Unfolded. *Biochim. Biophys. Acta (Bba) - Proteins Proteomics* 1698 (2), 131–153. doi:10.1016/j.bbapap.2003.12.008

van der Wel, P. C. A. (2017). Insights into Protein Misfolding and Aggregation Enabled by Solid-State NMR Spectroscopy. *Solid State. Nucl. Magn. Reson.* 88, 1–14. doi:10.1016/j.ssnmr.2017.10.001

Vanik, D. L., Surewicz, K. A., and Surewicz, W. K. (2004). Molecular Basis of Barriers for Interspecies Transmissibility of Mammalian Prions. *Mol. Cel* 14 (1), 139–145. doi:10.1016/s1097-2765(04)00155-8

Weissmann, C. (1999). Molecular Genetics of Transmissible Spongiform Encephalopathies. *J. Biol. Chem.* 274 (1), 3–6. doi:10.1074/jbc.274.1.3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 29 March 2022
doi: 10.3389/fmolb.2022.828785

NMR Assignment of Methyl Groups in Immobilized Proteins Using Multiple-Bond ^{13}C Homonuclear Transfers, Proton Detection, and Very Fast MAS

*Piotr Paluch[1,2], Rafal Augustyniak[1], Mai-Liis Org[3], Kalju Vanatalu[3], Ats Kaldma[3], Ago Samoson[3] and Jan Stanek[1]**

[1]Faculty of Chemistry, University of Warsaw, Warsaw, Poland, [2]Centre of Molecular and Macromolecular Studies, Polish Academy of Sciences, Łódź, Poland, [3]Tallin University of Technology, Tallinn, Estonia

OPEN ACCESS

Edited by:
Józef Romuald Lewandowski,
University of Warwick,
United Kingdom

Reviewed by:
Bernd Reif,
Technical University of Munich,
Germany
Thomas Vosegaard,
Aarhus University, Denmark

***Correspondence:**
Jan Stanek
janstanek@chem.uw.edu.pl

†ORCID:
Piotr Paluch
orcid.org/0000-0001-8601-3977
Rafal Augustyniak
orcid.org/0000-0001-6591-7192
Ago Samoson
orcid.org/0000-0001-7757-6356
Jan Stanek
orcid.org/0000-0003-1660-168X

Specialty section:
This article was submitted to
Structural Biology,
a section of the journal
Frontiers in Molecular Biosciences

Received: 03 December 2021
Accepted: 08 February 2022
Published: 29 March 2022

Citation:
Paluch P, Augustyniak R, Org M-L, Vanatalu K, Kaldma A, Samoson A and Stanek J (2022) NMR Assignment of Methyl Groups in Immobilized Proteins Using Multiple-Bond ^{13}C Homonuclear Transfers, Proton Detection, and Very Fast MAS.
Front. Mol. Biosci. 9:828785.
doi: 10.3389/fmolb.2022.828785

In nuclear magnetic resonance spectroscopy of proteins, methyl protons play a particular role as extremely sensitive reporters on dynamics, allosteric effects, and protein–protein interactions, accessible even in high-molecular-weight systems approaching 1 MDa. The notorious issue of their chemical shift assignment is addressed here by a joint use of solid-state ^{1}H-detected methods at very fast (nearly 100 kHz) magic-angle spinning, partial deuteration, and high-magnetic fields. The suitability of a series of RF schemes is evaluated for the efficient coherence transfer across entire ^{13}C side chains of methyl-containing residues, which is key for establishing connection between methyl and backbone ^{1}H resonances. The performance of ten methods for recoupling of either isotropic ^{13}C–^{13}C scalar or anisotropic dipolar interactions (five variants of TOBSY, FLOPSY, DIPSI, WALTZ, RFDR, and DREAM) is evaluated experimentally at two state-of-the-art magic-angle spinning (55 and 94.5 kHz) and static magnetic field conditions (18.8 and 23.5 T). Model isotopically labeled compounds (alanine and Met-Leu-Phe tripeptide) and ILV-methyl and amide-selectively protonated, and otherwise deuterated chicken α-spectrin SH3 protein are used as convenient reference systems. Spin dynamics simulations in SIMPSON are performed to determine optimal parameters of these RF schemes, up to recently experimentally attained spinning frequencies (200 kHz) and B_0 field strengths (28.2 T). The concept of linearization of ^{13}C side chain by appropriate isotope labeling is revisited and showed to significantly increase sensitivity of methyl-to-backbone correlations. A resolution enhancement provided by 4D spectroscopy with non-uniform (sparse) sampling is demonstrated to remove ambiguities in simultaneous resonance assignment of methyl proton and carbon chemical shifts.

Keywords: NMR resonance assignment, methyl groups, solid-state NMR, fast MAS, proton detection, TOCSY, isotope labeling/method

1 INTRODUCTION

For protein studies by nuclear magnetic resonance (NMR), amide and methyl ^{1}H resonances are the most commonly exploited. The latter ones are particularly convenient due to the magnetic equivalence of three ^{1}H spins (thus threefold sensitivity gain), and enhanced longitudinal relaxation, both caused by fast methyl rotation. In solution NMR, multiple-quantum correlations (HMQC) can be employed to select only the slowly relaxing methyl ^{1}H–^{13}C coherences (Tugarinov et al., 2003). When combined with extreme ^{1}H dilution by deuteration and selective methyl amino acid labeling, the approach allows to study the local dynamics and protein interactions for systems close to MDa molecular weight (MW) (Rosenzweig and Kay 2014; Huang and Kalodimos 2017; Boswell and Latham, 2018).

A prerequisite for the interpretation of NMR data at atomic resolution is a unique, site-specific mapping of chemical shifts to individual atoms. However, conventional resonance assignment strategies for proteins are centered around backbone ^{1}H resonances (Sattler, Schleucher, and Griesinger 1999). Chemical shifts of methyl spins, which are peripheral with respect to backbone, are clearly challenging to assign in a systematic way, particularly if detached methyl sites are the only nondeuterated spins. Tailored experiments were developed to correlate methyl to amide frequencies; however, in addition to long coherence transfers involved and thus intrinsically low sensitivity, they require full ^{1}H occupancy of (detected) amide sites (Tugarinov and Kay 2003b). The increased proton density is detrimental to ^{1}H resolution and sensitivity particularly in large-MW proteins. Alternative strategy relies on correlation to backbone ^{13}C′ and ^{13}Cα spins (Tugarinov and Kay 2003a), which can be accomplished in the absence of amide ^{1}H, but the issue of extended coherence transfer pathway persists.

Mutagenesis is commonly employed to address the methyl assignment issue (Amero et al., 2011). In this approach, single point mutations are introduced, and fingerprint ^{1}H–^{13}C HMQC spectra are recorded and compared for as many samples as methyl-containing amino acids. The major disadvantages are the labor and cost of isotope-enriched compounds needed to prepare typically tens of samples. Other pitfalls are ambiguities due to overlap of ^{1}H–^{13}C cross-peaks, or global chemical shift changes induced by mutations (Gorman et al., 2018).

An orthogonal approach is based on the observation of ^{1}H–^{1}H methyl proximities (NOEs), which can aid the assignment given the presence of the 3D structure determined using other techniques (Gorman et al., 2018). These data are interpreted either by a spectroscopist or, more effectively, by dedicated algorithms intensively developed over the last few years (Schmidt and Güntert 2012, 2013; Pritišanac et al., 2019; Pritišanac, Alderson, and Güntert 2020). The approach depends critically on the NOESY data quality and performs well mostly for methyl ^{1}H spins experiencing a dense network of interactions in rigid regions.

Immobilized proteins, such as in amyloid fibrils, sedimented large-MW aggregates, lipid bilayers, or microcrystals, are amenable to solid-state NMR (ssNMR), since this method does not experience high-MW slow-tumbling limitations characteristic of the solution counterpart (Tycko 2011; Miao and Cross 2013; Baker and Baldus 2014; Ladizhansky 2017; Linser 2017; Mandala, Williams, and Hong 2018; van der Wel 2018; Wiegand 2020). The possibility of detection of methyl ^{1}H resonances at high resolution in ssNMR was investigated already at low magic-angle spinning (MAS) frequencies (up to 20 kHz), when coupled to high to extreme ^{1}H dilution by deuteration (Agarwal et al., 2006; Agarwal and Reif 2008; Asami, Schmieder, and Reif 2010; Asami et al., 2012; Asami and Reif 2013; Mainz et al., 2013). The advent of fast-spinning (above 40 kHz) MAS probes allowed narrow methyl ^{1}H lines at significantly higher proton content, either when labile amide ^{1}H sites are fully reprotonated (Lewandowski et al., 2011; Linser et al., 2011), or when Ile, Leu, and Val residues are also selectively and nonrandomly (100% CH_3 or CHD_2) protonated at methyl sites (Huber et al., 2011; Agarwal et al., 2014; Andreas, et al., 2015a; Kurauskas et al., 2016; Gauto et al., 2019), or when Leu and Val residues are reverse-labeled in an otherwise deuterated matrix ("proton clouds") (Sinnige et al., 2014). Quite importantly, Schanda and co-workers recently showed that at 55–57 kHz MAS and B_0 = 14.1 T, the $^{13}CH_3$ isotopomer yields significantly higher sensitivity compared to $^{13}CHD_2$ labeling, and only at a minor loss of ^{1}H resolution (Kurauskas et al., 2016). Complete elimination of detrimental sensitivity and resolution effects of strong ^{1}H–^{1}H dipole interactions by MAS remains a challenge, particularly in methyl-dense protein regions, and would require yet unavailable MAS rates (above 250 kHz) (Kai Xue et al., 2017; K Xue et al., 2018; Kai Xue et al., 2019, Xue et al., 2020). Nevertheless, even in non-deuterated but relatively small proteins, resolved methyl ^{1}H–^{13}C correlations (of ^{1}H linewidths of about 150 Hz) were obtained with 100–111 kHz MAS and high static magnetic fields (23.5 T) (Andreas et al., 2016).

There are few examples of *de novo* assignment of methyl ^{1}H resonances in ssNMR, and frequently the assignment has been aided by either correlations to previously available ^{13}C shifts using either dipolar (Agarwal and Reif 2008; Asami and Reif 2012) or scalar-based (Andreas et al., 2015b) (H)CCH-type spectra, or transferred from solution NMR by comparison to relatively uncomplicated ^{1}H, ^{13}C-CP-HSQC (Huber et al., 2011). In an early work, a strategy was proposed that bases on correlation of methyl resonances to backbone ^{13}Cα and C′ spins using (H)CCH-TOBSY experiment, and detection of dilute methyl ^{1}H spins, but was found challenging for Leu and Ile residues due to low efficiency of multi-^{13}C-bond transfers (Agarwal and Reif 2008). A systematic strategy relying on correlations of side-chain ^{13}C to backbone ^{15}N and ^{1}H, originally proposed by Linser for perdeuterated proteins at slow MAS (25 kHz) (Linser 2011), and more recently at fast MAS (55–60 kHz) with rotor-asynchronous MOCCA mixing (Kulminskaya et al., 2016; Vasa et al., 2018), can be readily adapted for residues selectively 100% reprotonated at methyl sites, but their efficiency with respect to methyl-containing spin systems was not investigated in detail. In the case of non-deuterated ("fully protonated") proteins, such a long coherence transfer

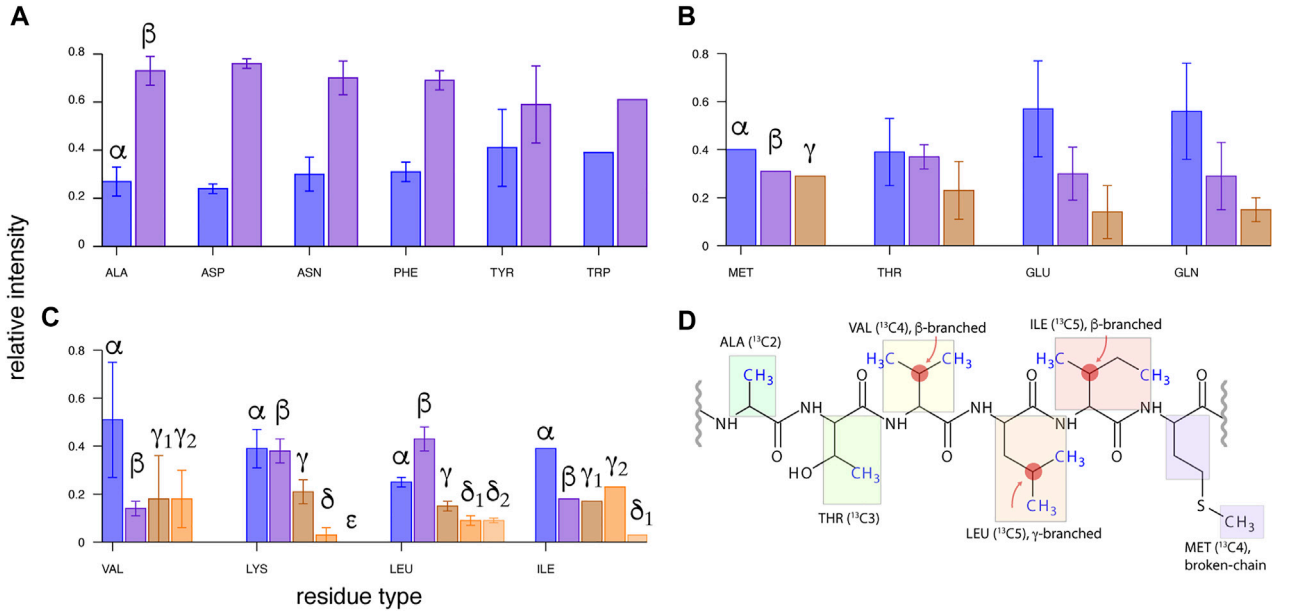


FIGURE 1 | Transfer efficiencies from side-chain ^{13}C (β, γ, . . .) to ^{13}Cα spin for various residues in microcrystalline U-^{13}C,^{15}N-labeled GB1 protein using ^{13}C WALTZ-16 mixing of 15 ms duration, with MAS of ν_R = 111 kHz (in a 0.7-mm rotor) and on a 23.5-T spectrometer. Intensity of ^{13}C^{X}-^{13}Cα-^{1}Hα cross-peaks in (H)CCH spectrum (Andreas et al., 2016; Stanek et al., 2016) was normalized to a sum of 1.0 for each residue, and averaged within each amino acid type. The data were grouped into **(A)** two- and pseudo-two-, **(B)** three- and pseudo-three-, and **(C)** four- and five-aliphatic ^{13}C spin systems (disregarding backbone and side-chain carbonyl and aromatic ^{13}C spins). **(D)** Topologies of ^{13}C side chains for six amino acid residue types containing a methyl group, from left to right: least to most challenging for methyl resonance assignment.

encounters sensitivity limitations, and more practical is a "two-hop" strategy in which side-chain ^{1}H and ^{13}C resonances are first correlated to alpha ^{1}H and ^{13}C spins using (H)CCH and H(C)CH experiments (Andreas et al., 2016; Stanek et al., 2016), and subsequently to amide ^{1}H and ^{15}N shifts using a combination of ^{13}Cα–^{15}N–^{1}H^{N} and ^{15}N–^{13}Cα–^{1}Hα correlations (Zhou et al., 2007; Stanek et al., 2016). Both aforementioned approaches were proposed for a general use (for all amino acid residue types), without a particular emphasis on the methyl assignment. In fact, the literature data on non-deuterated microcrystalline protein GB1 (Andreas et al., 2016) show that WALTZ mixing at MAS rate ν_R = 111 kHz and B_0 = 23.5 T leads to an effective ^{13}C^{x}→^{13}Cα transfer only in two-, pseudo-two-, and three-aliphatic ^{13}C-spin systems (**Figures 1A,B**), and becomes problematic for terminal spins in pseudo-three (Glu, Gln)- and larger ^{13}C spin systems (**Figure 1C**). Unfortunately, the latter include valine, leucine, and isoleucine, which additionally suffer from branching of ^{13}C chain at β- or γ-positions (**Figure 1D**). This raises questions of feasibility of methyl resonance assignment under less sensitivity-favorable sample or hardware conditions, and calls for a careful optimization of ^{13}C mixing scheme at arising new MAS conditions and B_0 fields.

In this work, we critically assess the efficiency of several ^{13}C mixing schemes for methyl ^{1}H and ^{13}C resonance assignment at fast MAS; determine their optimal range of B_0, MAS frequency, and ^{13}C RF strength within and beyond currently available regimes; and explore alternative ^{13}C labeling to further boost sensitivity of the approach.

2 MATERIALS AND METHODS

2.1 Sample Preparation

2,3-^{13}C (99%)-labeled crystalline alanine was purchased from *Eurisotop* (France) and manually packed into a Bruker 1.3-mm MAS rotor.

N-formylated microcrystalline uniformly ^{13}C,^{15}N-labeled Met-Leu-Phe tripeptide ("fMLF") was purchased from *Giotto Biotech* (Sesto Fiorentino, Italy) and manually packed into Darklands OÜ 0.81 mm MAS rotor.

The plasmid coding for the Src homology 3 (SH3) domain (965–1,025) of chicken α-spectrin (gene SPTAN1, Uniprot P07751) in a pET3a vector was a kind gift of Dr. T. Schubeis (High Field NMR Centre in Lyon, France). The insert was subcloned into a modified pET28a vector that resulted in a construct including an N-terminal His$_6$-tag followed by a SUMO solubility tag. This allowed us to express His$_6$-SUMO-SH3 fusion protein.

For preparation of the NMR samples, transformed *Escherichia coli* BL21 (DE3) cells were grown in M9 D_2O media supplemented with 1 g/L of $^{15}NH_4Cl$ (*Cortecnet*, France) and 3 g/L of (^{2}H,^{13}C)-glucose (*Cortecnet*, France) as the sole nitrogen and carbon sources, respectively, following the established procedure (Tugarinov, Kanelis, and Kay 2006). Labeling of the

Ile, Leu, and Val side chains was achieved by the addition of the amino acid precursors 1 h prior to the induction with 1 mM IPTG. For Leu and Val residues, we used two kinds of precursors that yield continuous ^{13}C chains from methyl (either Cδ or Cγ) to Cα and C′ atoms, provided that ^{13}C-enriched glucose is also employed (Goto et al., 1999; Tugarinov and Kay 2003a). The specific labeling patterns and the origin of particular nuclei are shown in **Supplementary Figure S1**. SH3 sample with branched side chains of leucine and valine residues (hereafter referred to as "ILV-C5" sample) was prepared using 100 mg/L of 1,2,3,4,4′-^{13}C-3-^{2}H-labeled α-ketoisovaleric-acid (sodium salt, *Eurisotop*, France, catalogue number CDLM-4418-PK) as a Leu and Val precursor. The corresponding sample with linearized ^{13}C-side chains of Leu/Val residues (referred to as "ILV-C4" sample) required the addition of 100 mg/L 1,2,3,4-^{13}C-3,4′,4′,4′-^{2}H-labeled α-ketoisovaleric acid (sodium salt, *Eurisotop*, France, catalogue number CDLM-8100-PK). For both samples, 60 mg/L of 1,2,3,4-^{13}C-3,3-^{2}H-labeled α-ketobutyric acid (sodium salt, *Eurisotop*, France, catalogue number CDLM-4611-PK) was used as the precursor of isoleucine residues with uniform ^{13}C enrichment (**Supplementary Figure S1**). The cells were grown at 24°C for 18 h after the induction.

The purification protocol of SH3 protein was modified with respect to the original one (Pauli et al., 2000, 2001; Chevelkov et al., 2006). Instead, a standard protocol involving His-trap affinity column and HiLoad 16/60 Superdex 75 gel filtration column (GE Healthcare) was employed. Briefly, we used 50 mM HEPES, 200 mM NaCl, and 1 mM DTT (pH 7.4) supplemented with 20 mM imidazole as a lysis buffer, and the same buffer including additionally 400 mM imidazole was used to elute a protein from the His-trap column. The cleavage of the His_6-SUMO tag was achieved with a custom-made Ulp1 protease (Reverter and Lima 2009) in a lysis buffer lacking imidazole. The reaction was monitored with the SDS-PAGE and, once completed, the protein was passed through the His-trap column again to remove SUMO as well as His-tagged Ulp1 protease. Cleaved SH3 protein was dialyzed at 4°C overnight against the Superdex 75 running buffer (20 mM citric acid, 150 mM NaCl, pH 3.5), concentrated, and purified on a gel filtration column.

To obtain solid-state NMR protein samples, fractions containing pure SH3 were pooled, concentrated to approximately 10 mg/mL using Vivaspin 3-kDa cutoff centrifugal concentrators (*Sartorius*) and extensively dialyzed against 100 mM $(NH_4)_2SO_4$ adjusted to pH 3.5 with sulfuric acid. Finally, to crystallize the protein, ammonia water solution was added dropwise to reach pH 7.5. Obtained turbid solutions were stored in a refrigerator (4°C) for approximately a week for a slow buildup of microcrystals. Crystallization efficiency was estimated to 85% by a spectrophotometric measurement of protein concentration decrease in the supernatant. The suspension was mixed with 1 M $CuNa_2EDTA$ 9:1 v/v (effective c_{Cu2+} = 100 mM), and left for impregnation of SH3 crystals with paramagnetic Cu^{2+} ions for 3 days. In the case of the SH3 "ILV-C4" sample, TSP was added to the supernatant at an effective concentration of 10 mM for chemical shift calibration reference. About 2 mg of protein was transferred to a Bruker 1.3-mm MAS rotor by 30 min of ultracentrifugation at an average acceleration of 96,500 *g*. Each Darklands OÜ 0.81-mm MAS rotor was filled with approximately 0.5 mg of protein in five equal parts by stepwise packing in a tailored ultracentrifuge adapter experiencing an average acceleration of 135,000 *g* for 30 min.

2.2 Nuclear Magnetic Resonance Spectroscopy

The NMR ^{13}C–^{13}C 2D correlation experiments on model compounds (alanine and fMLF) were performed using a standard RF irradiation scheme shown in **Figure 2A**. A collection of homonuclear mixing schemes was employed as detailed below. For each case, a series of 2D spectra with gradually incremented mixing time was acquired, in the range suggested by literature and spin dynamics simulations, if permitted by probe RF circuity.

The experiments for site-specific assignment of methyl ^{13}C and ^{1}H resonances in SH3 protein that provide correlations to backbone amide ^{15}N and ^{1}H chemical shifts [3D (H)C(CC)(CA)NH] were straightforwardly adopted from the literature (Linser 2011) and are shown in **Figure 2B**. Despite long methyl ^{1}H transverse relaxation times observed, the experiments employ ^{1}H–^{13}C cross-polarization (CP) instead of INEPT-type transfer, since the latter would suffer from inefficiency of conversion of anti- to in-phase ^{13}C coherence caused by concurrent evolution of two (passive) $^{1}J_{CH}$ couplings in CH_3 moieties. Implementations of presented experiments for Bruker spectrometers are freely available from the community-based repository Zenodo as detailed below.

For the site-specific evaluation of experimental performance with various ^{13}C mixing schemes, resolution provided by 3D spectra was required, preventing acquisition of mixing time series for SH3 protein. Optimal mixing times were, however, possible to determine using first-increment 1D optimizations prior to 3D data acquisition.

The series of 2D ^{13}C–^{13}C correlation experiments on 2,3-^{13}C-alanine was carried out on a Bruker Avance III spectrometer operating at ^{1}H and ^{13}C frequency of 600.1 and 150.9 MHz, respectively, equipped with a Bruker 1.3-mm H/C/N MAS probehead. The sample has been spun at 55,555 ± 10 Hz and controlled by a Bruker MAS-II unit, without temperature stabilization. ^{1}H, ^{13}C pulses and power during ^{1}H to ^{13}C cross-polarization have been carefully calibrated prior to experiments. Detailed information on pulse lengths, RF amplitude, spectral windows, etc. is provided in **Supplementary Table S1**.

2D experiments on ^{13}C,^{15}N-fMLF tripeptide and 3D and 4D experiments on two SH3 protein samples were performed on a Bruker Avance III HD spectrometer operating at a ^{1}H, ^{13}C, and ^{15}N frequency of 799.7, 201.1, and 81.0 MHz, respectively, equipped with a 0.81-mm H/C/N/D MAS probehead developed by Ago Samoson's group (Darklands OÜ, Estonia). NMR data were acquired at two MAS frequencies, ν_R = 55.5 and 94.5 kHz (98 kHz for fMLF).

3D experiments on "ILV-C5"-labeled SH3 protein were also performed on a Bruker NEO 23.5-T spectrometer (in CRMN

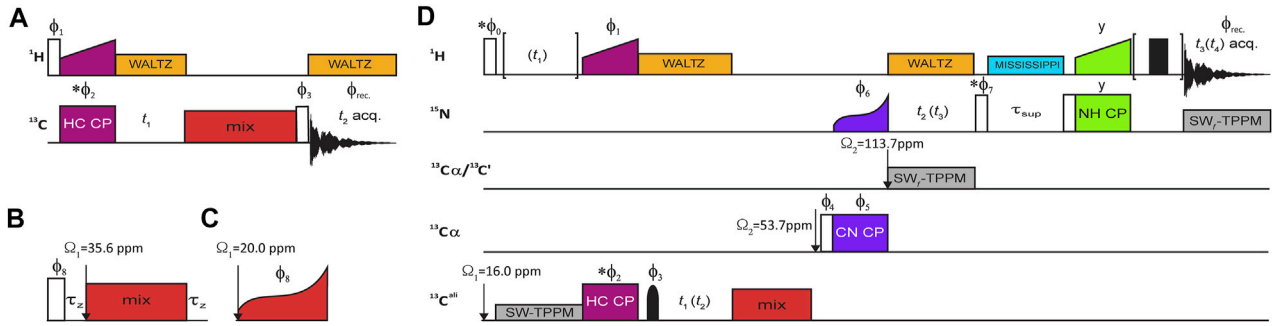


FIGURE 2 | RF irradiation schemes of NMR experiments. **(A)** ^{13}C–^{13}C 2D correlation experiment with mixing two variants of ^{13}C mixing expanded in **(B)** and **(C)**. All schemes act on the ^{13}C magnetization preceded and followed by a *z*-filter (τ_z) **(B)** with the exception of DREAM **(C)**, which operates on the ^{13}C single-quantum coherence. **(D)** Pulse diagram for 3D (H)C^{MET}(CC)(CA)NH, with the optional evolution period for methyl ^{1}H coherences enclosed in brackets (for the 4D variant). The optional ^{1}H spin echo immediately before acquisition was only applied for Bruker 1.3-mm probe to eliminate probe ^{1}H background signal. Open and solid rectangles represent 90° and 180° high-power pulses, respectively, and a solid bell shape denotes a methyl ^{13}C-selective Q3 refocusing pulse (Emsley and Bodenhausen 1992) of 942.7 μs duration and peak RF of ν_1 = 2.56 kHz. ^{13}C mixing is marked in *red*, and cross-polarization transfer steps are labeled and color-coded as *maroon*, *purple*, and *lime green* simultaneous irradiation on pairs of respective channels for H→C, C→N, and N→H transfers. ^{1}H low-power WALTZ (Shaka, Keeler, and Freeman 1983) decoupling is marked in *orange*, while heteronuclear (^{13}C or ^{15}N) frequency-swept TPPM decoupling (Bennett et al., 1995) is marked in *gray*. MISSISSPPI solvent suppression (Zhou and Rienstra 2008) is marked in *cyan*. Changes in ^{13}C offset (Ω) are denoted with vertical arrows. Unless indicated otherwise, all pulse phases are set to *x*. The following phase cycling was employed: $\phi_1 = y$, $\phi_2 = y$, $\phi_3 = \{2(x), 2(-x), 2(y), 2(-y)\}$, $\phi_8 = \{x, -x\}$, $\phi_{rec.} = \{x, -x, -x, x, y, -y, -y, y\}$ for 2D experiment with inset **(B)**; $\phi_1 = \{y, -y\}$, $\phi_2 = \{2(x), 2(-x)\}$, $\phi_3 = \{4(x), 4(-x)\}$, $\phi_8 = \{2(x), 2(-x)\}$, $\phi_{rec.} = \{x, -x, x, -x\}$ for 2D DREAM experiment (A + C); $\phi_1 = \{4(y), 4(-y)\}$, $\phi_2 = x$, $\phi_3 = \{x, x, y, y\}$, $\phi_4 = y$, $\phi_5 = \{8(x), 8-x)\}$, $\phi_6 = \{y, -y\}$, $\phi_7 = x$, $\phi_8 = -y$, $\phi_{rec.} = \{x, -x, -x, x, -x, x, x, -x, -x, x, -x, x, -x, -x, x\}$ for 3D experiment (D) with recoupling block (B); and $\phi_1 = \{8(y), 8(-y)\}$, $\phi_2 = \{4(x), 4(-x)\}$, $\phi_3 = \{2(x), 2(y), 2(-x), 2(-y)\}$, $\phi_4 = x$, $\phi_5 = x$, $\phi_6 = \{y, -y\}$, $\phi_7 = \{x\}$, $\phi_8 = \{4(x), 4(-x)\}$, $\phi_{rec.} = \{x, -x, -x, x, -x, x, x, -x, -x, x, -x, x, -x, -x, x\}$ for 3D DREAM experiment (D + C). Pulse with phase marked with an asterisk are involved in quadrature detection according to States-TPPI procedure.

Lyon, France) equipped with a Bruker 1.3-mm H/C/N/D MAS probe, and MAS-III and BCU-II spinning and temperature stabilization units.

For the experiments on ^{13}C,^{15}N-fMLF magic angle setting has been set using KBr sample prior to the actual series. In this case, no temperature stabilization device was used. ^{1}H and ^{13}C 90° pulse lengths were carefully calibrated using 1D ^{13}C-detected CP experiment.

Prior to experiments on SH3 samples, a careful magnet shimming was performed to maximize reliability of linewidth measurements. The classical shimming protocol employing adamantane sample is very time-consuming in 0.81-mm MAS rotors; thus, a sample of silicon grease was used first, leading to full-width at half-height (FWHH) of ^{1}H line of 16 Hz at 50 kHz MAS. Subsequently, shim currents were refined using adamantane sample resulting in an FWHH of the low-field ^{13}C resonance of 1.1 Hz at 50 kHz MAS under low-power ^{1}H decoupling.

Magic angle was finely adjusted directly on the protein sample at the target MAS frequency by maximizing the intensity of the ^{1}H signal in the (H)NH 1D experiment followed by a 5-ms ^{1}H spin echo. Sample temperature was stabilized to 20°C using Darklands OÜ VT controller. Thermocouple readout was calibrated to sample temperature by measurement of the ^{207}Pb chemical shift of $Pb(NO_3)_2$ at the exactly same cooling gas flow, spinning speed, and thermocouple target temperature. ^{1}H, ^{15}N 90° pulse lengths have been carefully calibrated for each sample and spinning speed using the (H)NH experiment. ^{13}C 90° pulse length was calibrated using the (H)CONH experiment. RF amplitude of hard and soft pulses, as well as of decoupling and recoupling schemes was automatically calculated in a pulse program based on widths and powers of reference high-power 90° pulses. ^{1}H–^{13}C, ^{13}C–^{15}N, and ^{15}N–^{1}H cross-polarization power was optimized directly using the first increment of the (H)C(DIPSI)(CA)NH experiment, and propagated to all 3D (and 4D) experiments at each sample and spinning condition. Fast recycling (0.3 or 0.4 s interscan delay) was applied to improve sensitivity, taking advantage of longitudinal ^{1}H relaxation enhancement by paramagnetic Cu^{2+} ions (Ganapathy et al., 1981; Wickramasinghe et al., 2007, 2009). Details on NMR data acquisition for fMLF tripeptide and SH3 protein are provided in **Supplementary Tables S1, S3**, respectively.

B_0 was not stabilized in either case, but the field drift was monitored using 1D ^{1}H spectra between experiments and had a negligible effect on data (e.g., a total 11 Hz ^{1}H downfield drift over 10 days on an 18.8-T spectrometer for "ILV-C5" SH3 sample at 94.5 kHz MAS). The stability of CP conditions was monitored using 1D (H)(CA)(N)H and (H)(C)(DIPSI)(CA)(N)H (for proteins) or ^{13}C-detected 1D CP (for fMLF) experiments.

2D and 3D data were Fourier processed in Bruker TopSpin using parameters reported in **Supplementary Tables S2, S4**. The non-uniformly sampled 4D data were converted using in-house written script *bruk2ssa* (courtesy of M. Górka), and processed using signal separation algorithm for distortion-free spectral reconstruction (Stanek, Augustyniak, and Koźmiński 2012). All spectra were analyzed in NMRFAM-Sparky (Lee, Tonelli, and Markley 2015).

2.3 Homonuclear ^{13}C Mixing Schemes

The key element determining the performance of experiments shown in **Figure 2** is ^{13}C mixing. For this comparative study, we

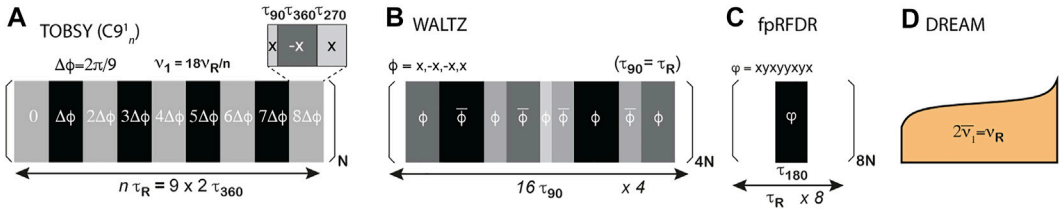


FIGURE 3 | RF schemes of ^{13}C mixings compared in this study. **(A)** TOBSY $C9^1_n$ variants with variable *n*. **(B)** WALTZ-16, a representative of TOCSY class. The primitive element is composed of phase-alternated 90, 180, 270, and 360° pulses, depicted as rectangles color-coded from *light gray* to *black*. **(C,D)** Transfer schemes primarily recoupling ^{13}C dipolar interactions: **(C)** finite-pulse RFDR and **(D)** DREAM.

have made a selection of literature RF designs based on the following criteria: (1) ^{13}C should occur primarily between bonded ^{13}C spins by recoupling of ^{13}C–^{13}C scalar (*J*) or dipolar (*D*) interactions, (2) RF requirements are acceptable for a typical "fast" MAS probe, and (3) they pose no danger to the hydration of fragile protein samples, e.g., by an extended period of high-power ^{1}H irradiation. For the last reason, we did not consider second-order recoupling schemes suitable for fast MAS, such as MIRROR (Scholz et al., 2008), PARIS (Weingarth et al., 2009), SHANGHAI (Hu et al., 2011), or CORD/CORD-RFDR (Hou et al., 2013; Lu et al., 2015). In the presence of relaxation, the efficiency of these sequences deteriorates with faster MAS, concomitantly to increasing efficiency of averaging of H–H dipolar interactions. Protein deuteration, which is highly recommended for resolution in large proteins, is also clearly incompatible with second-order recoupling mechanism (De Paëpe, Lewandowski, and Griffin 2008) as it prohibitively dilutes the proton interaction network.

2.3.1 TOBSY

TOBSY was designed based on average Hamiltonian theory (AHT) and symmetry properties of particular interactions (Levitt 2007), imposing that chemical shift anisotropy (CSA), isotropic chemical shifts (CS), and dipolar interaction terms are suppressed in the zeroth- and first-order Hamiltonians, and the zeroth-order term stems only from the isotropic scalar interaction (Baldus and Meier 1996; Hardy, Verel, and Meier 2001). In a general design denoted CN^{ν}_{n} (Edén and Levitt 1999), a primitive multi-pulse block C is repeated N times over *n* rotor cycles with the gradual phase incrementation in $\Delta\phi = 2\pi\nu/N$ steps (**Figure 3A**). Low-power TOBSY schemes suitable for fast MAS employ POST primitive block ($90_{\phi}360_{\phi+\pi}270_{\phi}$), and use $N = 9$, $\nu = 1$, and $n = 9p \pm 3$ (where *p* is an integer), e.g., $C9^1_{21}$, $C9^1_{24}$, and $C9^1_{30}$, to retain only assumed interaction (Tan et al., 2018). The schemes with particular *n* differ in robustness to dipolar C–H interactions and susceptibility to CSA (in higher-order AHT terms). TOBSY requires an RF strength of $\nu_1 = 2N/n\ \nu_R$ and thus is readily applicable in typical fast MAS probes for $n > 18$. Here, we employed sequences with $n = 24$, 30, 33, 39, and 48, with mixing time either varied between 0 and 50 ms (in 2D series) or optimized for maximum transfer (in 3D experiments).

2.3.2 TOCSY

Isotropic mixing schemes developed for ^{13}C mixing in solution NMR (Bax, Clore, and Gronenborn 1990; Kay et al., 1993), hereafter referred to using a general term TOCSY, retain only the Hamiltonian term associated with scalar (*J*) interaction between ^{13}C nuclei (in the presence of the overall tumbling). For solids under MAS, all isotropic interactions are preserved; thus, these RF schemes induce the evolution of J_{CC} interactions while suppressing the evolution of J_{HC} and isotropic ^{13}C chemical shifts. The proper treatment of anisotropic interactions is not ensured at all; however, at sufficiently fast MAS, their contribution is *supposedly* small. Simple arguments advocated for the use of WALTZ-16 (Shaka, Keeler, and Freeman 1983) mixing synchronized with rotation ($\tau_{90} = \tau_R$; **Figure 3B**), and basic properties were verified using 2-^{13}C-spin SIMPSON simulations (Andreas et al., 2016). Classical TOCSY mixing schemes that are robust with respect to large chemical shift offsets, namely, DIPSI-3 (Shaka, Lee, and Pines 1988) and FLOPSY-16 (Kadkhodaie et al., 1991), were also employed here. In solution NMR, TOCSY designs are ranked according to the figure of merit, i.e., the bandwidth with respect to applied RF strength, but such a ranking is of limited relevance here since MAS probes can easily generate sufficient ν_1. Although not required, we retained rotor synchronization (ν_1 = ¼ ν_R) in experiments for three selected TOCSY sequences, but tested other RF conditions in spin dynamics simulations. Mixing time was varied (for alanine and fMLF) or optimized in 1D experiments (for proteins) in steps of 188.448, 217.32, and 96.0 τ_R (assuming $\tau_R = \tau^C_{90}$) for FLOPSY, DIPSI, and WALTZ, respectively.

2.3.3 RFDR

Finite-pulse ^{13}C RFDR (Bennett et al., 1998) recouples dipolar ^{13}C interactions in the first-order average Hamiltonian by the application of high-power π pulse each rotor cycle (**Figure 3C**). Although the sequence is known to suffer from dipolar truncation (Griffin 1998), this is not necessarily a disadvantage for intraresidue ^{13}C transfers. RFDR requires $\nu_1 > ½\ \nu_R$ for pulses partially covering the mixing time; in practice, ν_1 on the order of ν_R is preferred. The ratio of τ_{180}/τ_R determines the scaling factor of dipolar interaction (nominally of about 2 kHz for a bonded ^{13}C–^{13}C spin pair), and in this study, we used ν_1 = 150 (for fMLF) and 100 kHz (otherwise). The following is the phase cycle (xy8) of the π pulse: *x*, *y*, *x*, *y*, *y*, *x*, *y*, *x*, as suggested previously (Shen et al., 2012). Mixing time was optimized in steps of multiples of $8\tau_R$ between 0 and 50 ms.

2.3.4 DREAM

DREAM recouples dipolar ^{13}C interactions in the first-order average Hamiltonian using an adiabatic pulse for introducing HORROR condition ($\nu_1 = ½\ \nu_R$, **Figure 3D**) (Verel et al., 1998; Verel, Ernst, and Meier 2001). DREAM is commonly employed with slow (Pauli et al., 2001) and fast MAS (Penzel et al., 2015) to trigger ^{13}Cα→^{13}Cβ or ^{13}C′→^{13}Cα coherence transfer, for which RF and offset conditions can be straightforwardly determined. Transfers in multi-spin systems were studied and conditions were optimized at "slow" MAS (Westfeld et al., 2012); however, the proper order of HORROR conditions cannot in general be satisfied due to characteristic chemical shifts in ^{13}C side chains. Additionally, by its nature, DREAM cross-peak intensity is negative with respect to the origin coherence, which potentially causes destructive interferences within a spin system or due to spectral overlaps. The RF shape was defined as usual, $\nu(t) = \overline{\nu_1} + d^{eff}\tan(\frac{2}{\tau}\tan^{-1}(\Delta/d^{eff})(t-\tau/2))$, where d^{eff}, Δ, τ, and $\overline{\nu_1}$ denote the effective dipolar coupling (averaged over crystal orientations and decreased due to dynamics), modulation depth, mixing time, and average RF strength, respectively. Directionality of the transfer was selected here with increasing RF over mixing time. We employed the parameters recommended for DREAM adiabatic RF modulation at fast MAS, namely, the modulation depth of 1/5 ν_R, dipolar C–C coupling of 1 kHz, and average RF $\bar{\nu}_1$ to nominally ½ ν_R, which potentially allows DREAM to cover the entire aliphatic ^{13}C band at typical high fields (e.g., 18.8 T). In addition to the mixing time (varied up to 10 or 25 ms for alanine, or fMLF and SH3, respectively), ^{13}C offset for DREAM was also optimized for best ^{13}C^{met}→^{13}Cα transfer in the 2D series for alanine and fMLF, and using 1D first-FID of sequence shown in **Figure 2D**.

2.4 Spin Dynamics Simulations Using SIMPSON

Spin dynamics simulations have been performed using SIMPSON (version 4.0.0c) (Bak, Rasmussen, and Nielsen 2011; Tošner et al., 2014). Powder averaging has been performed with 1,848 {α, β, γ} Euler angles that describe the orientation of the molecule in the rotor frame. A total of 168 {α, β} angle pairs have been selected using REPULSION algorithms (Bak and Nielsen 1997) and 11 γ angles have been regularly sampled from 0 to 360°. Spin systems have been generated using the SIMMOL package (Bak et al., 2002).

For simulations in the case of a model four-spin system (Hα–Cα–Cβ–Hβ), dipolar spin couplings have been generated based on distances in L-alanine structure (ref. code LALNIN61 in CCS database), and only α and β carbons and protons have been considered. $^1J_{CC}$ and $^1J_{HC}$ couplings have been set to 33 and 145 Hz, respectively. In addition, only one proton from the CH_3 group has been considered with the H–C dipolar coupling value reduced by three due to fast rotation around the C–C axis. Other relevant parameters of the spin system are reported in the **Supplementary Material**.

For simulation in the case of C_6 and C_5 spin systems, dipolar couplings have been calculated for the geometry of leucine-8 in the X-ray structure of chicken SH3 protein (PDB 1SHG). Experimentally determined chemical shifts in fMLF have been assumed, and CSA parameters were calculated in Gaussian. $^1J_{CC}$ couplings were set to 50 Hz for the ^{13}Cα–^{13}C′ pair and to 33 Hz between aliphatic ^{13}C spins. $^2J_{CC}$ was set to 3 Hz, and longer-distance *J* couplings were neglected. Examples of SIMPSON input files are provided in the **Supplementary Material**.

3 RESULTS AND DISCUSSION

3.1 Homonuclear Mixing in Model Two-^{13}C Spin Systems

Efficiency of ^{13}C mixing varies in general with MAS frequency ν_R, RF strength ν_1 (if not imposed by ν_R), B_0 field (through scaling of isotropic chemical shift differences and CSA), presence of ^{1}H–^{13}C interactions, and extent of dynamics as well as other design-specific parameters detailed above. To mitigate the complexity of the problem, one usually refers to model two-^{13}C spin systems, such as, e.g., in the acetate anion. Here, we used 2,3-$^{13}C_2$,^{15}N-labeled crystalline alanine as a closer analogue of a protein residue, additionally possessing a methyl group, and avoiding interference with large CSA of the (unlabeled) carbonyl carbon. The efficiency of ^{13}Cβ→^{13}Cα transfer was quantified by observation of cross-peak intensity in a τ_{mix} series of 2D ^{13}C–^{13}C spectra at fast MAS (ν_R = 55.5 kHz) and moderate B_0 field (14.1 T).

The results shown in **Figure 4** confirm the high efficiency of selected RF designs. Also, fundamental differences between them exhibit the following: TOBSY and TOCSY sequences recouple J_{CC} in a time close to 1/(2J) (scaling factors might apply to the effective coupling constant, e.g., for FLOPSY). RFDR initially shows rapid oscillations with frequency related to a scaled dipolar coupling (D_{CC} ≈ 2 kHz), stabilizes at approximately τ_{mix} = 5–10 ms, and then decays due to π pulse imperfections and incoherent effects. DREAM shows a steady buildup, reaches a plateau at τ_{mix} ≈ 5 ms, and then decays slowly mostly due to ^{13}C $T_{1\rho}$-like relaxation. The results are generally consistent with the literature, also confirming the increasing robustness of TOBSY $C9^1_n$ to D_{HC} interactions with increasing *n* (**Figure 4A**) (Tan et al., 2018). Although RFDR experimentally showed highest performance, observed differences between considered schemes are rather minor.

Spin dynamics simulations are instrumental for the in-depth understanding of complex RF schemes applied under new MAS conditions. Here, we resorted to SIMPSON as a generally agreed simulation platform (Tošner et al., 2014), and attempted first to reproduce experimental results. We modeled alanine using a minimal system consisting of two ^{13}C and ^{1}H spins as described above. Results shown in **Figures 4E–H** reproduce well the experimental buildup curves, optimal mixing times, and scaling factors of recoupled interactions. Differences in efficiency between TOBSY $C9^1_n$ sequences with *n* = 48, 39, 33, and 24 (**Figure 4E**) are more pronounced than in the experiment (**Figure 4A**). $C9^1_{48}$ allows for an almost complete transfer for τ_{mix}

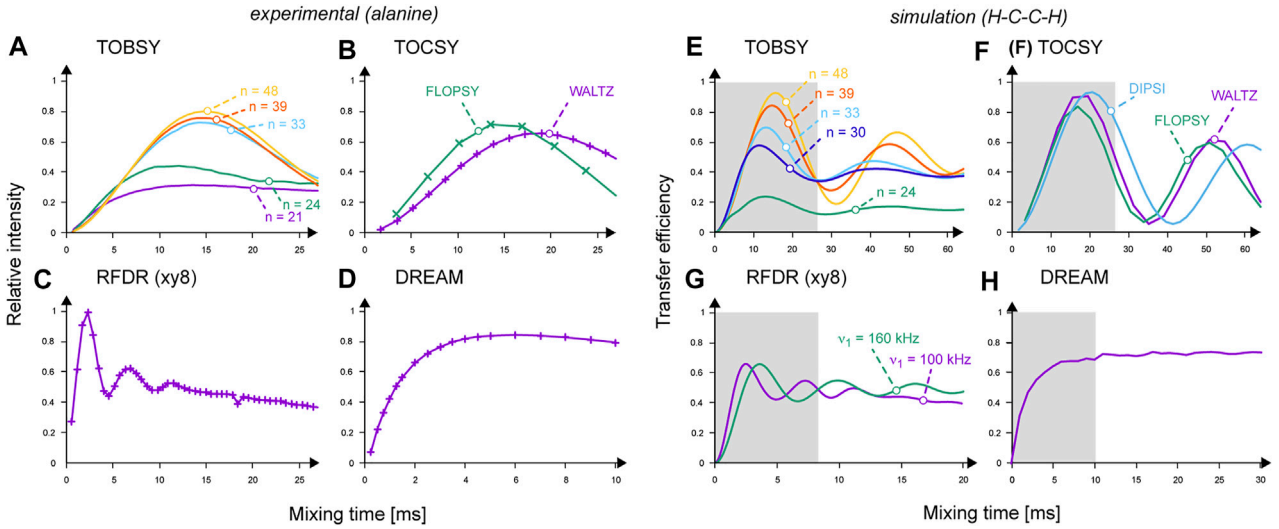


FIGURE 4 | (A–D) Performance of ^{13}C mixing schemes observed experimentally in ^{13}C–^{13}C 2D spectra on 2,3-$^{13}C_2$-labeled alanine: **(A)** TOBSY C9^1_n with n = 21, 24, 33, 39, and 48 are labeled and color-coded as *purple*, *green*, *cyan*, *red*, and *orange* curves, respectively; **(B)** FLOPSY-16 (in *green*) and WALTZ-16 (in *purple*); **(C)** RFDR at ν_1 = 100 kHz; **(D)** DREAM. Cross-peak intensities were normalized to the most intense point in the RFDR series. Experimental points are connected with lines for eye guidance only. **(E–H)** Buildup of transferred ^{13}C magnetization (absolute transfer efficiency) in a two ^{13}C- and two ^{1}H-spin system simulated in SIMPSON. **(E)** TOBSY C9^1_n with n = 24, 30, 33, 39, and 48 are labeled and color-coded as *green*, *dark blue*, *cyan*, *red*, and *orange* curves, respectively; **(F)** FLOPSY-16 (in *green*), DIPSI-3 (in *light blue*) and WALTZ-16 (in *purple*); **(G)** RFDR at ν_1 = 100 (in *purple*) and 160 kHz (in *green*); **(H)** DREAM. Gray boxes indicate the mixing time ranges sampled experimentally on alanine.

≈ 16 ms. For longer mixing times, recoupling of higher-order AHT terms occurs (mostly the cross-terms between ^{13}C isotropic and anisotropic chemical shifts, and D_{HC} interactions), leading to coherence dephasing and decreased intensity of subsequent maximums. Somewhat surprisingly, all TOCSY sequences perform excellently (**Figure 4F**), despite no deliberate treatment of anisotropic interactions. RFDR performance is relatively worse than in the experiment, and equalizes ^{13}C magnetization between both coupled ^{13}C spins. Supposedly, incoherent effects and B_1 field inhomogeneity in the MAS coil affect RFDR to the smallest extent among the considered mixing types. In simulations, DREAM does not show an optimum, but a steadily increasing coherence transfer.

Having validated the simulation platform, we attempted to differentiate ^{13}C mixing schemes with respect to destructive interferences (cross-terms) with interactions that are field-dependent (isotropic and anisotropic CS), and proton-content-dependent ($D_{H\text{-}C}$ interactions). All combinations of selectively "activated" interactions were probed, and the effects of most relevant cross-terms are summarized in **Supplementary Figure S2**. We confirmed that TOBSY C9^1_n sequences with n > 30 are quite robust to D_{HC}, with (J_{CC}, D_{HC}), (D_{CC}, D_{HC}), and (D_{CC}, CSA) being the primary cross-terms responsible for non-ideal performance. We additionally confirmed that TOBSY is robust with respect to an increasing chemical shift difference (or B_0 field), e.g., up to $\Delta\Omega$ = 15 kHz for C9$^1_{24}$ at ν_R = 55 kHz (and ν_1 = 2/3 ν_R). Among TOCSY sequences, FLOPSY appeared to be most susceptible to (J_{CC}, D_{HC}) interferences (or, in other words, to the presence of protons). Also, all TOCSY designs are less affected by (D_{CC}, D_{HC}) terms than TOBSY and very robust to CSA (**Supplementary Figure S2**). Thus, WALTZ and DIPSI are good candidates for high B_0 field measurements, and on systems with high proton density. Within the initial buildup, RFDR is extremely robust to all interaction interferences, and in fact even enhanced by recoupling of J_{CC} interactions.

Additionally, we verified the susceptibility of TOBSY and TOCSY sequences to B_1 field inhomogeneity, which can be substantial in MAS RF coils [as large as 20% in the active volume (Tošner et al., 2018)]. As shown in **Supplementary Figure S3**, TOBSY sequences only tolerate maximum 5% RF deviations (or miscalibration), and TOCSY designs, DIPSI in particular, are significantly more robust with this respect.

However, multiple limitations of spin dynamics simulations must be considered when their results are interpreted quantitatively: (1) no incoherent effects (relaxation) are included; (2) structural variability of dipolar interactions, chemical shifts, and CSA are tedious to replicate; and (3) pulse transients, limited short-term rotation stability, and other experimental deficiencies are neglected.

3.2 Homonuclear ^{13}C Transfer Across Multi-Spin Side Chains in Model Systems

Subsequently, we performed an analogous series of 2D ^{13}C–^{13}C correlation spectra on a sample of crystalline tripeptide fMLF, where the leucine residue serves as a realistic and convenient reference multi-spin system containing a methyl group. Spectral excerpts in **Figure 5** demonstrate that, while one-bond transfers remain effective in all cases, efficiencies of multi-bond ones are low and differ dramatically between particular RF schemes (see

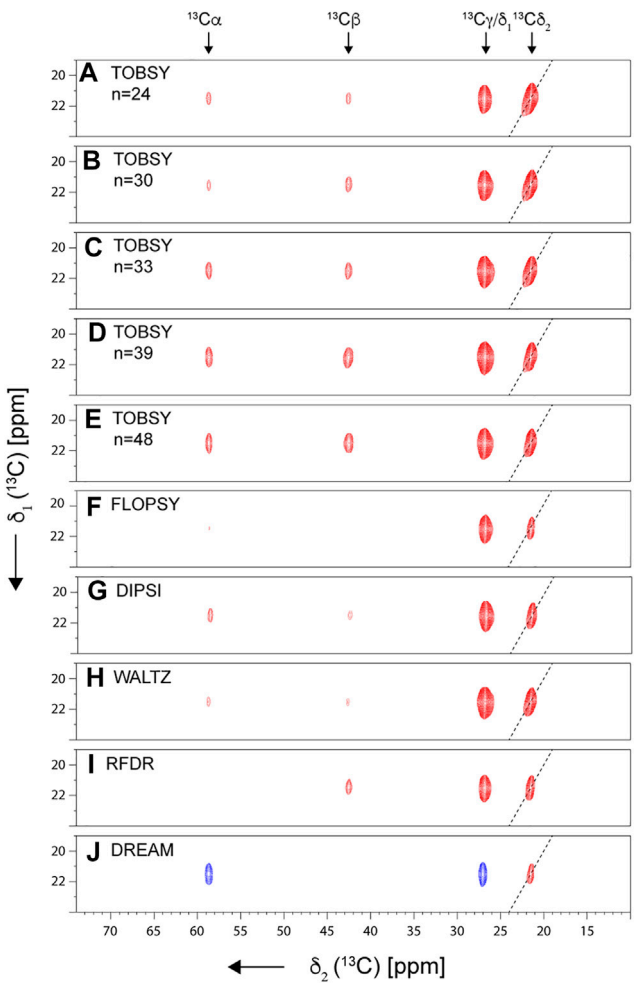


FIGURE 5 | Strips from 2D ^{13}C–^{13}C spectra recorded for fMLF at ν_R = 55.5 kHz on an 18.8-T spectrometer. Presented are diagonal and cross-peaks originating from Leu ^{13}Cδ_1 magnetization (prior to ^{13}C mixing). Spectrum diagonal is indicated with a dashed line. Spectra with a mixing time optimal for Cδ_1→ Cα transfer are shown: 34.6, 32.0, 21.4, 25.3, and 36.3 ms for TOBSY C9^1_n, with n = 24, 30, 33, 39, and 48, respectively **(A–E)**; 23.7, 27.4, and 20.7 for FLOPSY-16 **(F)**, DIPSI-3 **(G)**, and WALTZ-16 **(H)**, respectively; and 11.2 and 14 ms for RFDR (ν_1 = 150 kHz) **(I)** and DREAM **(J)**, respectively.

Supplementary Figure S4 for mixing time dependencies). In contrast to the case of alanine, RFDR performs worst in this respect, emphasizing the limited utility of simplistic two-^{13}C spin systems for evaluation of ^{13}C mixing performance.

It was noted previously in both solution (Tugarinov and Kay 2003a) and solid-state NMR (Agarwal and Reif 2008) that a combination of a branched ^{13}C chain, large number of ^{13}C spins, and large chemical shift differences makes the leucine spin system extremely challenging for any ^{13}C mixing. In the following, we will deliberately focus on the most demanding Cδ_1→Cα transfer as the sensitivity-limiting step, regardless of which side-chain assignment strategy mentioned above [the one-step side-chain to backbone, (H)C(CC)(CA)NH, or the two-step approach using (H)CCH and (H)NCAHA spectra] is employed. (Another potentially relevant transfer Cδ_1→C′ is expected to be even more problematic due to a large CSA and distinct chemical shift.)

The entire experimental series was repeated at very fast spinning conditions (ν_R = 98 kHz) to probe sensitivity of selected mixing schemes to coherent and incoherent effects of dipolar C–H and H–H interactions. Absolute efficiencies of Cδ_1→Cα transfer were quantified as described in **Supplementary Material**, and presented in **Figure 6**. Strikingly, all RF schemes greatly benefit from increased MAS rate (up to a factor of 4.5–5 for three TOCSY representatives). At ν_R = 55.5 kHz DREAM performs best, and the most efficient scheme at ν_R = 98 kHz is DIPSI, with TOBSY C9$^1_{48}$ being close to the leaders at both MAS frequencies. In both cases, RFDR and TOBSY C9^1_n with n < 30 remain evident outliers.

Given such a dramatic response to MAS frequency, we attempted to verify capabilities of spin dynamics simulation for reproducing the behavior of a complex, but extremely useful leucine spin system, and possibly extrapolate it to different spinning and field conditions. Unfortunately, despite intense optimization of SIMPSON routines (e.g., the reuse of propagators) and the use of the state-of-the-art high-performance clusters, inclusion of any meaningful number of ^{1}H spins ($L \approx$ 5–6) turned out unfeasible due to the exponential scaling of required computational resources (4^L). Nevertheless, many features of the transfers were quite well reproduced in the exclusively ^{13}C 6-spin system, in particular (1) the optimal Cδ_1 → Cα mixing times (**Supplementary Figure S5**), (2) the shape of build-up curves for individual transfers (**Supplementary Figure S6**), and (3) poor performance of RFDR (≈5%) at both spinning speeds. The results also corroborate TOBSY C9$^1_{48}$ and DIPSI as the performance leaders at ν_R ≈ 100 kHz (**Supplementary Figure S7**). The exact quantification of DREAM poses difficulties due to the absence of an optimum in simulations, but efficiency at τ_{mix} = 25 ms, as used in the experiments, places this mixing scheme closer to the best than to mediocre designs.

The absence of protons in the spin system used in simulations resulted in almost identical performance of all TOBSY schemes at ν_R = 100 kHz, which clearly disagrees with the experiments. Likely for the same reason, the experimentally observed relative order of performance of TOCSY sequences was not reproduced. To some extent, one can still rely on the analysis of cross-term importance for four spin system (**Supplementary Figure S2**) to predict the impact of protons also in leucine residues. However, absolute efficiencies predicted in simulations (**Supplementary Figure S7**) and observed experimentally (**Figure 6**) show discrepancies from roughly 25% (for TOBSY C9$^1_{48}$ and DIPSI) to more than 100% (e.g., for FLOPSY and WALTZ) at ν_R = 98–100 kHz, and even larger ones at the lower spinning speed (ν_R = 55.5 kHz). These disagreements arise not only due to the (unaccounted) coherent effects of ^{1}H–^{13}C interactions, but also due to ^{13}C relaxation, suggesting the ultimate need of experimental verification for a quantitative comparison at specific ν_R and B_0.

Despite apparent limitations, SIMPSON simulations still provide upper limits for transfer efficiencies, which is useful to predict performance trends at different experimental conditions.

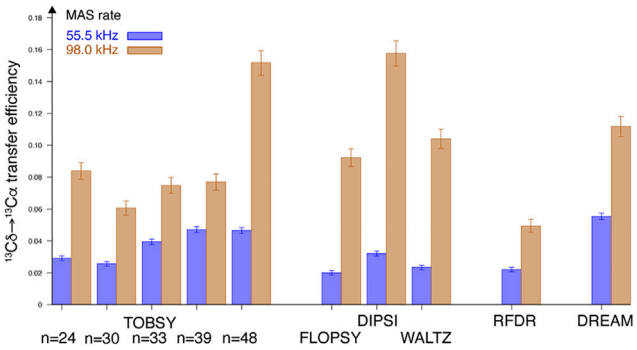


FIGURE 6 | Maximum absolute efficiency of $C\delta_1 \rightarrow C\alpha$ magnetization transfer observed in leucine residue of fMLF at B_0 = 18.8 T field and at ν_R = 55.5 (*blue* bars, left in each pair) and 98 kHz (*tan* bars, right in each pair).

For example, ν_R dependence sampled between 30 and 200 kHz (**Supplementary Figure S8**) shows an optimal MAS range of 80–100 kHz for TOBSY and TOCSY schemes. The penalty observed at ν_R > 100 kHz is surprising, and in fact related to RF strength (ν_1), which was fixed in a constant proportion to ν_R. To isolate ν_1 dependence without traversing through recoupling conditions, we performed additional simulations at *constant* ν_1 = 55.5 kHz and ν_1 = ¼ ν_R, but indirectly varied RF bandwidth of sequences *via* modulation of resonance frequencies and CSA with the strength of B_0 field (**Figure 7**). At high field (above $\nu_{0,H}$ = 800–1,000 MHz), most sequences show insufficient ^{13}C bandwidth, thus decreasing the performance of a multibond transfer. At low field, the transfer efficiency degrades as well, with a very similar behavior to that observed in **Supplementary Figure S8** for large rotation rates ν_R. This suggests that the effect actually relates only to the effective ^{13}C bandwidth, and we confirmed by observation of C′ magnetization that an excessive RF strength is detrimental to ^{13}Cα magnetization due to the concurrent drainage to C′.

This leads to a paradox that increased MAS rates might not necessarily be beneficial if RF strength is bound to ν_R by the mixing design. Contrary to the case of TOBSY, TOCSY schemes tolerate variable ν_1/ν_R ratios in a wide range below ¼ (**Supplementary Figure S10**), which gives an additional flexibility for their application at moderate-to-high fields and very fast spinning. The simulation does not account for RF-dependent ^{13}C $T_{1\rho}$, and very small RF should obviously also be avoided, particularly in systems showing increased microsecond local dynamics.

3.3 Methyl Resonance Assignment in Proteins by Correlation to Backbone Spins

Results obtained for fMLF strongly suggest proton dilution for the effective ^{13}C mixing in extended spin systems in proteins. Fortunately, with fast MAS, this can be accomplished without compromising the occupancy of methyl ^{1}H sites, e.g., by selective labeling of Ile, Leu, and Val residues from suitable precursors coupled to expression in D_2O media (Tugarinov and Kay 2004). Importantly, ^{13}C (and ^{2}H)-enriched glucose must be used to preserve the continuity of Ile and Leu ^{13}C side chains (**Supplementary Figure S1**), since α and β of the former and α and carbonyl carbons of the latter residue are incorporated from the medium rather than from the precursor (Lundström et al., 2007).

To demonstrate the efficacy of methyl ^{1}H assignment, we resorted to a model SH3 domain of α-spectrin in a microcrystalline state. Given the selective ^{1}H labeling, a magnetization transfer *from* methyl *to* backbone amide ^{1}H (and *not* a reverse one) is strongly preferred for the sensitivity reasons. In addition to 3-fold larger occupancy of methyl ^{1}H, the initial polarization is enhanced by fast T_1 relaxation of methyl protons, while NMR signal detection is performed on well-dispersed, generally narrower, and, thus, more sensitive amide ^{1}H protons. Naturally, a protocol for complete reprotonation of labile amide ^{1}H sites after expression in D_2O is a prerequisite for this approach. While extensive deuteration is not a necessity for proteins as simple as SH3 at fast MAS (>50 kHz) and high field (here 18.8 T), it would certainly be required for resolution and sensitivity reasons in more challenging applications. We employed a slightly adapted literature RF irradiation scheme (shown already in **Figure 2D**), with various ^{13}C mixings as in the previous cases.

3.3.1 Joint Effect of the Rotation Frequency and RF Strength

For SH3 sample we acquired 3D spectra with site-specific resolution at a single optimized mixing time for each ^{13}C mixing. Two spinning conditions, ν_R = 55.5 and 94.5 kHz, were applied using the same probe and rotor for the maximum cohesion of the data, and the representative strips are presented in **Figure 8** (for ν_R = 94.5 kHz).

The experiment inherently shows only the transferred signal, thus poses difficulties to the rigorous quantification of ^{13}C transfer efficiency. Nevertheless, we attempted to correct for uneven efficiency of the remaining part of the pulse sequence (notably CP steps) at different MAS conditions by normalizing the cross-peak intensities to those observed in the 3D (H)CANH experiment, which shares majority of the coherence pathway. The observed relative signal-to-noise ratio was averaged over 24 individual strong and resolved correlations. Results shown in **Figure 9** (for all residue types) and **Supplementary Figure S11** (separately for Val, Leu, and Ile) show a dramatic increase of ^{13}C mixing efficiency with fast MAS (94.5 w. r. t. 55.5 kHz), with a similar effect (a factor of 3.2–3.5 for TOCSY and $C9^1_{48}$) to that observed on nondeuterated fMLF. Despite a dilute network of ^{1}H interactions in the SH3 sample, and averaging also over valine and isoleucine residues, the best-performing RF schemes are virtually the same as for the leucine residue in fMLF (**Figure 6**), namely, DREAM at 55.5 kHz MAS, and DREAM, TOBSY $C9^1_{48}$, and three TOCSY sequences at 94.5 kHz MAS. The relevance of ^{1}H–^{1}H and ^{1}H–^{13}C interactions might be due to the fact that, despite partial sample deuteration, the local ^{1}H environment of leucine, valine, and isoleucine residues is quite dense, with

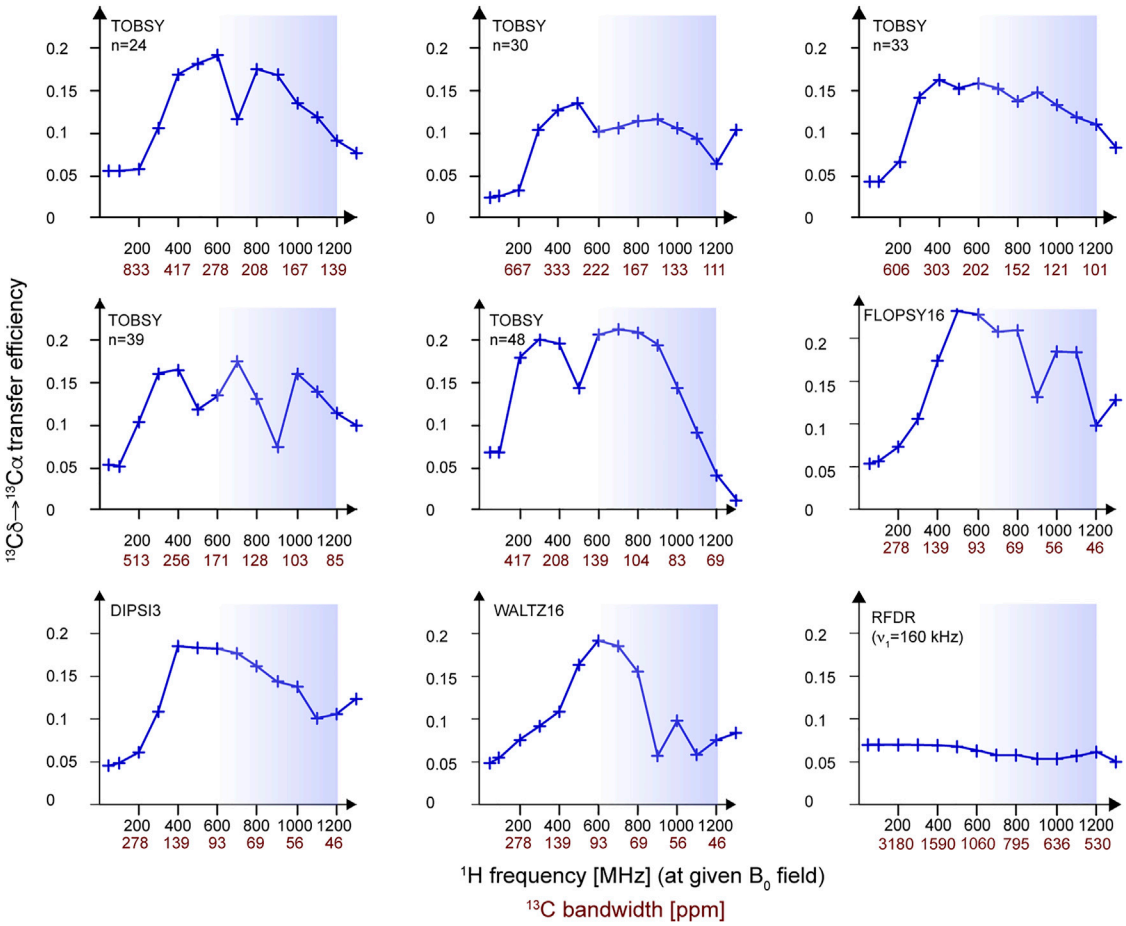


FIGURE 7 | Efficiency of $C\delta_1 \rightarrow C\alpha$ transfer in a $^{13}C_6$ spin system simulated in SIMPSON for a range of magnetic fields B_0 (corresponding to ^{1}H frequency between 50 and 1,300 MHz), MAS frequency of 55.5 kHz, and a constant RF strength of $\nu_1 = ¼\ \nu_R \approx 13.9$ kHz. Note that, in essence, the dependence on ^{13}C bandwidth (ν_1 expressed in ^{13}C ppm) of a particular mixing is investigated here. For each B_0 point, a full buildup was performed, and an optimum was picked at the mixing times that are reported in **Supplementary Figure S9**. All RF schemes except the particularly computationally demanding DREAM were evaluated. Blue boxes indicate the B_0 field range available for ^{1}H-detected MAS NMR at present.

only 4 out of 11, 2 of 9, and 7 of 11 protons replaced by deuterons, respectively. As for the overall proton density, 14 ILV residues (out of 62 residues in SH3) contain 81 protons at methyl sites. These, in addition to 116 backbone and side-chain labile protons, yield a moderate rather than a low protonation level (approximately 38%).

Compared to fMLF, spins in SH3 protein undergo higher-amplitude dynamics, and certainly there is a larger contribution of incoherent effects, such as leading to ^{13}C T_2 and $T_{1\rho}$ relaxation and ^{1}H spin-diffusion. $T_{1\rho}$ relaxation is further suppressed by stronger RF at increased MAS frequency (particularly for DREAM, which effectively spin-locks the ^{13}C coherence), as it is bound to ν_R for all sequences except RFDR. As for the coherent effects, increased MAS frequency also more effectively averages D_{CH} interactions (i.e., a fast rotation suppresses second-order AHT cross-terms, deleterious, e.g., for TOBSY with low n). Definitely, there is an additional S/N benefit from line narrowing of amide ^{1}H resonances at 94.5 kHz MAS (separable under a few of assumptions), but, as described below, it is estimated to contribute no more than 20% to the overall S/N gain.

In view of the apparent relevance of H–C and H–H interactions in deuterated SH3, additional efficiency gains are expected with the use of high-power ^{1}H decoupling ($\nu_{1,H} > 3\nu_R$) during ^{13}C mixing; however, this likely endangers sample hydration and thus spectral quality. Overall, fast MAS seems to be an attractive route to amplify the signal, and, in our comparison, the gain in efficiency well compensates the 2- to 3-fold sensitivity losses entailed by active volume reduction from a 1.3-mm rotor (optimal for $\nu_R \approx 55$ kHz) to a 0.7- or 0.81-mm rotor ($\nu_R \approx 100$ kHz). We speculate that for methyl resonance assignment using the presented approach, *non*deuterated proteins would benefit from fast MAS to an even larger extent.

3.3.2 Effect of the Static Magnetic Field

Our experiments were performed at a typical high-field spectrometer (18.8 T) used commonly in protein studies by

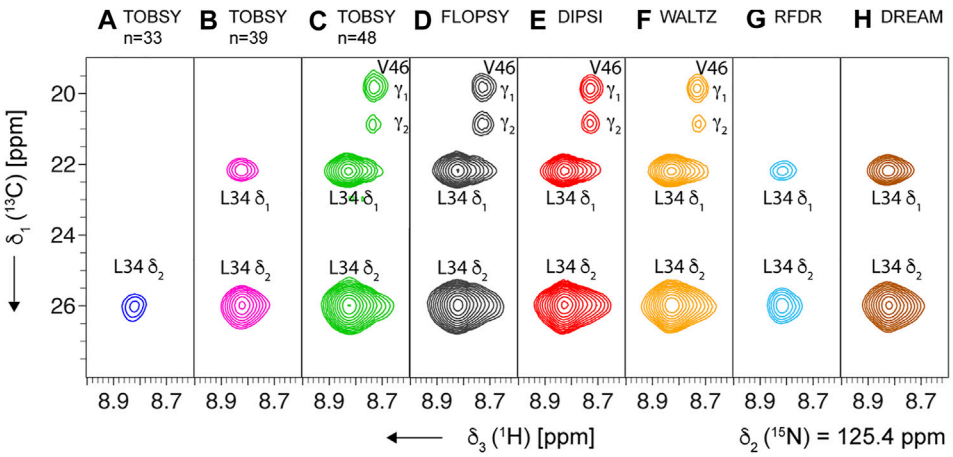


FIGURE 8 | Representative strips from 3D (H)C(CC)(CA)NH experiment for SH3 protein recorded at MAS frequency of 94.5 kHz and on an 18.8-T spectrometer, using the following mixing schemes at their optimal mixing time (see **Supplementary Table S3**): **(A–C)** TOBSY $C9^1_n$, with n = 33, 39, and 48, respectively, **(D)** FLOPSY-16, **(E)** DIPSI-3, **(F)** WALTZ-16, **(G)** RFDR (ν_1 = 100 kHz), and **(H)** DREAM. All planes show the cross-section at δ_2 (^{15}N) = 125.4 ppm, and show two cross-peaks for Leu-34 residue (L34 $^1H^N$,^{15}N intra-residue correlations to δ_1 and δ_2 ^{13}C). The most intense spectra (panels **C–F**) additionally show cross-peaks of Val-46 residue due to partial overlap of L34 and V46 $^1H^N$ and ^{15}N frequencies.

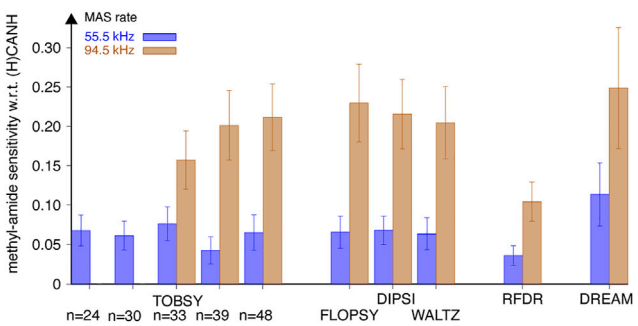


FIGURE 9 | Sensitivity (S/N per unit time) of methyl-to-amide correlations in 3D (H)C(CC)(CA)NH spectra normalized to sensitivity of respective intraresidue correlations in 3D (H)CANH spectra, recorded for SH3 protein at a MAS frequency of 55.5 (*blue* bars, left in each pair) and 94.5 kHz (*tan* bars, right in each pair) on an 18.8-T spectrometer using various ^{13}C mixing schemes. Relative intensity of 24 cross-peaks [for 6 valine (all except Val-46), 1 isoleucine, and 6 leucine residues] was averaged with equal weights. Data for TOBSY $C9^1_{24}$ and $C9^1_{30}$ were not collected at ν_R = 94.5 kHz. The error bars reflect the standard deviation (scatter) of values observed for a set of residues, not the experimental error of average sensitivity.

^{1}H-detected MAS. Stronger B_0 fields are increasingly accessible; therefore, it is relevant to explore the performance of ^{13}C mixing at such conditions. For example, for a transition to a 23.5-T field, spin dynamics simulations on a six-^{13}C spin system suggest only a minor and negative impact on transfer efficiency (results in **Figure 7** are summarized for two considered B_0 fields in **Supplementary Figure S12**). For experimental verification, we performed a full series of 3D experiments on SH3 protein (from the same crystallization batch) in a 1.3-mm rotor spun at ν_R = 55.5 kHz on a 1,000-MHz ^{1}H spectrometer. Data were analyzed with a site-specific resolution and averaged over all reliable cross-peaks of Leu, Val, and Ile residues as mentioned above. Also, as before, we normalized the obtained S/N ratios of cross-peaks with respect to peaks in 3D (H)CANH to eliminate the effects of larger Boltzmann polarization, rotor active volume, possibly different sample packing density, and specific RF coil efficiency (see **Supplementary Material**). Results compared in **Figure 10** illustrate that only selected ^{13}C mixing schemes, notably, FLOPSY, RFDR, and DIPSI, actually profit from higher B_0 field. Note that for consistency with the data acquired at 18.8 T, we did not optimize RF strength (ν_1) for TOCSY sequences (in the range below ¼ ν_R), despite potential gains in efficiency (see **Figure 7** and **Supplementary Figure S10**). The rationale for observed efficiency changes are again coherent and incoherent effects of H–H and H–C interactions, which evade proper treatment in our spin dynamics simulations.

3.3.3 Linearization of the ^{13}C Side Chain

One of the reasons for a low transfer efficiency between methyl and alpha ^{13}C spins using J_{CC}-based isotropic mixing in Leu, Val, and Ile is a branched topology of the ^{13}C chain. A 2-fold degradation of FLOPSY-8 efficiency for Ile residues was identified by Kay and co-workers, and remedied with relay-type (COSY) experiments in which ^{13}C coherence is sequentially transferred with a careful manipulation of individual ^{13}C spins using very frequency selective pulses (Tugarinov and Kay 2003b). This approach is not immediately applicable in MAS NMR, since required transfer delays are prohibitively long compared to ^{13}C T_2 lifetimes. For Leu residues, the overlap between chemical shifts of δ and γ ^{13}C precludes the elimination of detrimental passive coupling to the second methyl ^{13}C; thus, in solution NMR, the issue was addressed by a tailored amino acid labeling, in which one of the methyl groups was labeled as $^{12}CD_3$ (Tugarinov and Kay 2003a). Since the labeling of the α-

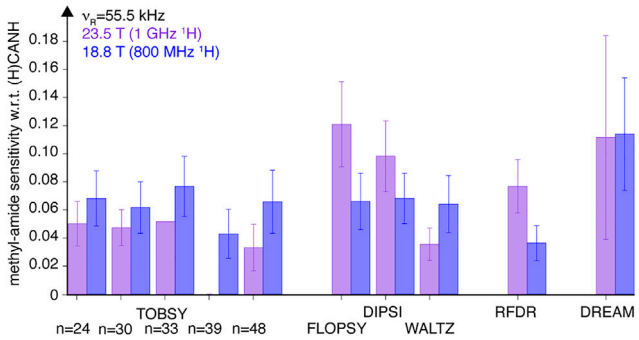


FIGURE 10 | Efficiency of various ^{13}C mixing schemes compared between 18.8- and 23.5-T fields [*purple* (left) and *blue* (right) bars, respectively]. Presented are sensitivity of correlation peaks in 3D (H)C(CC)(CA)NH spectra of SH3 protein obtained at 55.5 kHz MAS, with respect to sensitivity of correlations in (H)CANH experiment at both conditions. Relative sensitivity is averaged equally over 24 cross-peaks of Val, Leu, and Ile residues. Data for TOBSY $C9^1_{39}$ were not collected at 23.5 T field. The error bars reflect the standard deviation (scatter) of values observed for a set of residues, not the experimental error of average sensitivity.

ketoisovalerate (precursor) is not stereospecific, a 1:1 mixture of Leu ($^{13}CH_3$, $^{12}CD_3$) and Leu ($^{12}CD_3$,$^{13}CH_3$) isotopomers is obtained (and analogously for valines), entailing a 50% loss of methyl ^{1}H occupancy. Nevertheless, this sensitivity loss was compensated by larger efficiency of complex out-and-back transfer schemes proposed by Kay and co-workers for methyl ^{1}H assignment in solution.

To verify the utility of the linearization of ^{13}C side chains of Leu and Val residues for MAS NMR, we first performed SIMPSON simulations, and compared the behavior of 5-^{13}C spin (linear) and 6-^{13}C spin (branched) spin systems that mimic the leucine residue with respective isotope labeling patterns (**Supplementary Figures S13** and **Supplementary Figure S14**). Indeed, for all *J*-based isotropic mixing schemes (TOBSY and TOCSY), a 2- to 3-fold efficiency gain is predicted at ν_R = 100 kHz, and comparable at ν_R = 55.5 kHz. It is thus expected that, in this case, the 50% loss of initial signal is at least compensated by increased Cδ→Cα (Cγ→Cα for valines) transfer efficiency. Dipolar-based mixing (RFDR and DREAM) marginally profits from the simplification of ^{13}C chain, likely only due to dispersion of ^{13}C magnetization over a smaller number of ^{13}C spins.

The sample of SH3 with linearized side chains of Leu and Val [formally {I (δ_1), L ($^{13}CH_3$, $^{12}CD_3$), V ($^{13}CH_3$,$^{12}CD_3$)} U-(^{15}N, ^{13}C, ^{2}H, $^1H^N$)-labeled, with Ile-$^{13}C_6$, Leu-$^{13}C_5$,Val-$^{13}C_4$, here referred to as "ILV-C4" for brevity] was prepared by following carefully the expression protocol of the previous ("ILV-C5", in fact Ile-$^{13}C_6$, Leu-$^{13}C_6$, and Val-$^{13}C_5$) sample, but using the 1,2,3,4-^{13}C-3,4′,4′,4′-^{2}H-labeled α-ketoisovalerate as a precursor (Tugarinov and Kay 2003a). The microcrystals were paramagnetically doped with EDTA-chelated Cu^{II} ions, and transferred into MAS 0.81-mm rotor with comparable density (see ^{13}C 1D spectra of both samples in **Supplementary Figure S15A**). As expected, the methyl ^{1}H signal in direct-excitation spectra decreased approximately twice (**Supplementary Figure S15B**). We repeated the entire series of 3D (H)C(CC)(CA)NH experiments at both spinning frequencies ν_R = 55.5 and 94.5 kHz, using virtually the same previous experimental settings. Differences in CP efficiency and rotor packing density were compensated by normalization to the respective intraresidue peak intensities in (H)CANH spectra. Finally, for each condition, the relative cross-peak intensities that approximate Cδ (or Cγ)→Cα transfer efficiencies were averaged over 24 correlation peaks. The comparison in **Figure 11** provides evidence that transfer efficiency increased well beyond the factor of 2, yielding improved sensitivity for "ILV-C4" sample (with the exception of FLOPSY), and unexpectedly also manifested for dipolar-based mixings, for which a total sensitivity loss was expected based on SIMPSON simulations. Despite the 50% dilution of methyl ^{1}H spins, "ILV-C4"-labeling clearly yields superior results, with DIPSI, WALTZ, and TOBSY $C9^1_{48}$ as the methods of first choice. Interestingly, if the stereospecifically labeled $^{13}C_4$-α-ketoisovalerate was commercially available, S/N ratios could double, corresponding to a further sensitivity gain (i.e., per square root of time) of $\sqrt{2}$ (accounting for the need for two separate acquisition series).

The unexpectedly high gains observed with "ILV-C4" labeling can be explained by (1) decreased proton density, particularly in the local environments of leucine and valine ^{13}C spin systems, and (2) increased detection sensitivity due to amide ^{1}H line narrowing. The relevance of the first effect on the efficiency of ^{13}C transfer can be appreciated based on the MAS frequency dependence discussed above for the "ILV-C5" sample. Here, we illustrated the effect by measurement of methyl ^{1}H linewidth change upon additional proton dilution, as observed in ^{1}H, ^{13}C-CP-HSQC spectra as cross-peak raw full-width at half-height. Indeed, a 2-fold linewidth reduction is observed at ν_R = 55.5 kHz (**Figure 12A** and **Supplementary Figure S16A, B**), surpassing the effect of faster rotation for the "ILV-C5" sample (the ratio of 1.3 between 55.5 and 94.5 kHz). For the "ILV-C4" sample, both spinning conditions lead to similar linewidths (46 ± 11 Hz) with only few exceptions.

The second effect was carefully quantified by ^{1}H amide linewidth measurements in fingerprint ^{1}H, ^{15}N-CP-HSQC spectra (also without application of a window function in ^{1}H dimension). As expected, proton dilution in the methyl side chain of Leu and Val residues has a relatively smaller impact on amide ^{1}H coherence lifetimes (**Figure 12B**). At ν_R = 94.5 kHz, both samples show similar linewidths of 40 Hz (see **Supplementary Figures S16C,D** for a per-peak comparison), but approximately 9% difference is observed at ν_R = 55.5 kHz. The line-narrowing effect of rotation frequency increase is of a factor of 1.4 and 1.3 for "ILV-C5" and "ILV-C4" samples, respectively. Overall, the use of "ILV-C4" labeling is beneficial for resolution of ^{1}H, ^{13}C methyl resonances, and sensitivity of methyl-to-amide correlations.

3.3.4 Assignment of Protons: 4D Correlations

As shown in **Figure 13**, even in proteins as small as SH3, a single 3D ($^{13}C^{met}$-edited) spectrum can result in massive

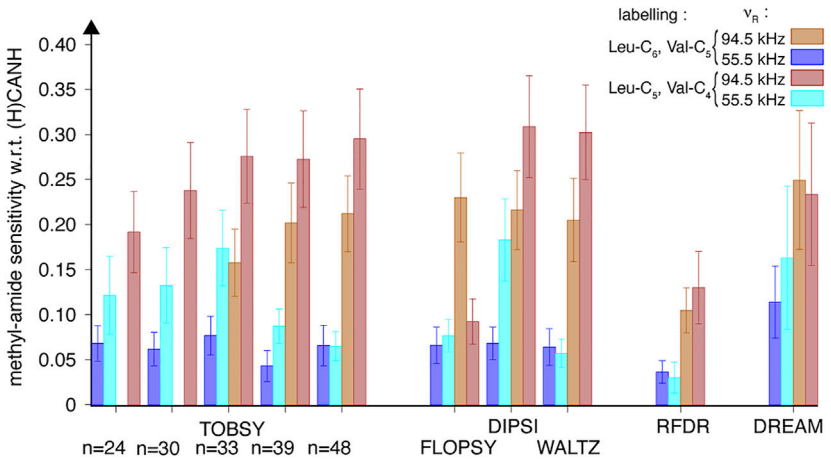


FIGURE 11 | Relative efficiency of ^{13}C mixing schemes compared between two SH3 protein samples: with branched ("ILV-C5", *tan* filled bars) and linear ("ILV-C4", *red* filled bars) ^{13}C spin systems of leucine and valine residues, quantified using 24 cross-peak intensities in 3D (H)C(CC)(CA)NH spectra recorded at a MAS frequency of 94.5 kHz on an 18.8-T spectrometer. Data for 55.5 kHz MAS is shown as *blue* and *cyan* bars for "ILV-C5" and "ILV-C4" SH3 samples, respectively. Sensitivity of each cross-peak was normalized to respective intra-residue peak in 3D (H)CANH spectrum at each sample and spinning condition, and subsequently averaged over Val, Leu, and Ile residues.

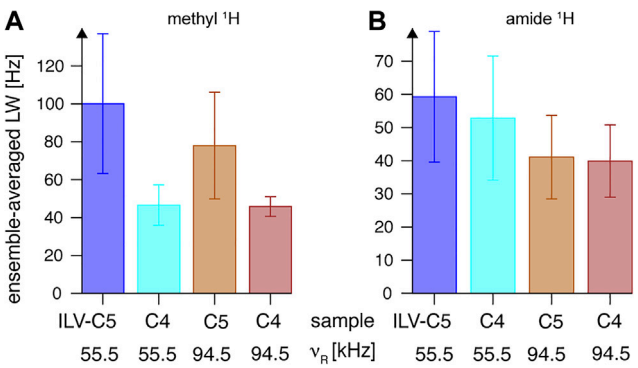


FIGURE 12 | Linewidths (full widths at half height) of **(A)** methyl and **(B)** backbone amide ^{1}H resonances observed in 2D ^{13}C-^{1}H and ^{15}N–^{1}H CP-HSQC spectra, respectively, for "ILV-C4" (at ν_R = 55.5 kHz in *cyan*, and at ν_R = 94.5 kHz in *red*) and "ILV-C5" (at ν_R = 55.5 kHz in *blue*, and at ν_R = 94.5 kHz in *tan*) SH3 samples on an 18.8-T spectrometer. Linewidths observed in each case were averaged over all resonances, and the error bars indicate the standard deviation (the dispersion of values observed for all residues), not the precision of linewidth determination.

ambiguities. For example, V53 γ_2, V58 γ_2, L8 δ_1, and L34 δ_1 ^{1}H assignment cannot be established from (H)C(CC)(CA)NH spectrum alone, since these methyl sites exhibit close ^{13}C shifts (highlighted by a *dashed* line in ^{13}C-CP-HSQC, panel C). The joint analysis of a pair of 3D (H)C(CC)(CA)NH and H(C)(CC)(CA)NH spectra could help to correctly match ^{1}H to ^{13}C frequencies within a spin system (given sufficient resolution of each spectrum); however, if both methyl sites in a residue show either close ^{1}H or close ^{13}C shifts, the ambiguity remains (e.g., in Val-58). The stereospecific labeling, i.e., using leucine and valine pro-*S*-C_4 and pro-*R*-C_4 precursors (yet unavailable for the presented approach) would be a costly solution.

Alternatively, one can resort to high-resolution 4D spectroscopy with non-uniform sampling, increasingly popular for MAS NMR (Huber et al., 2011; Linser et al., 2011; Paramasivam et al., 2012; Linser et al., 2014; Xiang et al., 2014; Fraga et al., 2017; Sergeyev et al., 2017; Marchanka et al., 2018; Vasa et al., 2018). Indeed, the use of a 4D HC(CC)(CA)NH spectrum (**Figure 13B**) allows disambiguating the assignment as demonstrated for L34 and V58 spin systems (close ^{1}H shifts), which additionally suffer from overlap of L34 δ_1 and V58 γ_2 peaks in ^{1}H, ^{13}C-CP-HSQC. Given that a 4D NUS spectrum can be acquired in a similar time to a pair of 3D ones, this approach is recommended for all but the smallest proteins.

In the presented 4D NUS experiment, the average S/N of cross-peaks, normalized to 24 h of acquisition, was 25 ± 7 (**Supplementary Figure S17**). One can safely interpolate that a minimal S/N ratio of 10 is obtained in 3.8 h for SH3 (62 aa). If we limit the discussion to the class of microcrystalline samples of comparable spectral properties (CP efficiency and linewidths) and crystal packing density, S/N of a single peak scales inversely with the number of amino acids *K* (to account for the smaller number of molecules in a rotor). Thus, the time needed for the same minimal S/N ratio scales as K^2, yielding reasonable acquisition times of 0.42, 1.7, 3.7, 6.7, 10.4, and 15 days needed at 18.8 T for proteins of *K* = 100, 200, 300, 400, 500, and 600 residues, respectively. It is noteworthy that these times would be approximately 2- and 3.4-fold shorter, respectively, on 23.5- and 28.2-T (1,000 and 1,200 MHz ^{1}H) spectrometers available nowadays (assuming $B_0^{3/2}$ scaling of S/N ratio). Our results thus suggest feasibility of site-specific assignment of methyl resonances in sizable proteins for which the crystallization conditions and deuteration protocol have been

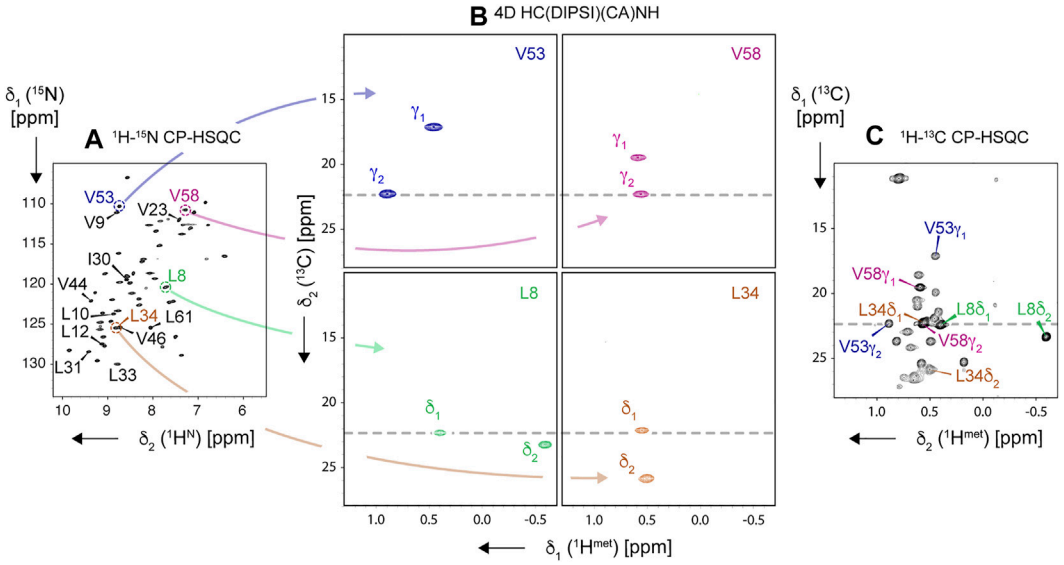


FIGURE 13 | Resolving ambiguity of methyl ^{1}H and ^{13}C resonance assignment using 4D HC(DIPSI)(CA)NH experiment, demonstrated on the "ILV-C4" sample of SH3 protein, spun at 94.5 kHz on an 18.8-T spectrometer. Navigation in the 4D spectral matrix is based on peak positions in 2D ^{1}H, ^{15}N-CP-HSQC **(A)**, allowing to display intraresidue pairs of correlations in ^{1}H, ^{13}C-cross sections of the 4D cube **(B)**. Obtained methyl resonance assignments for four representative residues (L8, L34, V53, and V58, color-coded) are readily transferable to 2D ^{1}H, ^{13}C-CP-HSQC spectrum **(C)**. In **(A)**, amide correlations of all Leu, Val, and Ile residues are labeled. All correlations except for residue Val-46 were observed in the 4D experiment.

already established, and can potentially be of great value to further solution or solid-state NMR studies. Whether methyl chemical shifts are readily transferable to other sample conditions is yet to be investigated. However, provided no significant alteration to protein fold, aliphatic ^{13}C should not experience large deviations as demonstrated recently for maltose binding protein (Stanek et al., 2020).

4 CONCLUSION

We showed that a careful selection of ^{13}C homonuclear mixing allows one to obtain sensitive correlations of methyl-to-amide ^{1}H chemical shifts under fast MAS and high B_0 field conditions. For highly deuterated proteins on an 18.8-T spectrometer, the best performance was obtained with DIPSI-3, WALTZ-16, and TOBSY $C9^{1}_{48}$ at ν_R = 94.5 kHz, and with DREAM at ν_R = 55.5 kHz. Dramatic improvement of the multi-bond (methyl to alpha ^{13}C) transfer efficiency was observed upon increase of MAS frequency from 55.5 to 94.5 kHz, which is attributed to the suppression of incoherent effects of H–H and H–C dipolar interactions. Further significant performance increase was obtained by the linearization of ^{13}C side chains of leucine and valine residues, with additional gains in resolution of methyl ^{1}H resonances. We demonstrated that unambiguous assignment of methyl ^{1}H and ^{13}C resonances is feasible for microcrystalline proteins by a combination of protein deuteration, paramagnetic T_1 relaxation enhancement, suitable ^{13}C isotope labeling, ultra-fast MAS, and 4D spectroscopy with non-uniform sampling. These findings pave the way for efficient assignment complementary to the labor-intensive mutagenesis-based strategy, with protein mass limitations largely mitigated by the favorable scaling of sensitivity in MAS NMR.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://doi.org/10.5281/zenodo.5911897.

AUTHOR CONTRIBUTIONS

PP and JS designed and performed the research and analyzed data. RA prepared protein samples. M-LO, KV, and AK designed, constructed, and tested the MAS 100-kHz system (probe, rotation, and temperature controllers and rotor tools) under supervision of AS. PP, RA, and JS wrote the manuscript.

FUNDING

This research was funded by the Polish National Science Centre with grant No. 2019/33/B/ST4/02021. JS and PP thank the Polish National Agency for Academic Exchange for the generous support of *Polish Return* programme (Contract No. PPN/PPO/2018/1/00098). The computational resources for this work were provided by the Polish Infrastructure for Supporting Computational Science in the

European Research Space (PLGRID, grant ID: plgnmrsi2, plgnmrsi3). We gratefully acknowledge financial support from the TGIR-RMN-THC Fr3050 CNRS for the access to the 23.5-T NMR spectrometer in CRMN (Lyon, France).

ACKNOWLEDGMENTS

We are grateful to Tobias Schubeis (CRMN Lyon) for sharing the SH3 plasmid and the original sample preparation protocol. The help of Tanguy Le Marchand in setting up the experiments on the 23.5-T spectrometer (CRMN Lyon) is greatly appreciated. The authors thank Morgane Callon (ETH Zurich) and Alons Lends (CBMN—IECB Bordeaux) for hints on rotor packing. We gratefully acknowledge M.J. Potrzebowski for providing access to the 14.1-T spectrometer at CBMM PAS in Łódź (Poland). The authors wish to express gratitude to Rector of Tallin University of Technology for financial support, and Darklands OÜ for technical support. We would like to thank W. Koźmiński for technical advice and maintenance of NMR system at CNBCh of University of Warsaw.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2022.828785/full#supplementary-material

REFERENCES

Agarwal, V., Diehl, A., Skrynnikov, N., and Reif, B. (2006). High Resolution ^{1}H Detected ^{1}H,^{13}C Correlation Spectra in MAS Solid-State NMR Using Deuterated Proteins with Selective ^{1}H,^{2}H Isotopic Labeling of Methyl Groups. *J. Am. Chem. Soc.* 128 (39), 12620–12621. doi:10.1021/ja064379m

Agarwal, V., Penzel, S., Szekely, K., Cadalbert, R., Testori, E., Oss, A., et al. (2014). De Novo 3D Structure Determination from Sub-milligram Protein Samples by Solid-State 100 kHz MAS NMR Spectroscopy. *Angew. Chem. Int. Ed.* 53 (45), 12253–12256. doi:10.1002/anie.201405730

Agarwal, V., and Reif, B. (2008). Residual Methyl Protonation in Perdeuterated Proteins for Multi-Dimensional Correlation Experiments in MAS Solid-State NMR Spectroscopy. *J. Magn. Reson.* 194 (1), 16–24. doi:10.1016/j.jmr.2008.05.021

Amero, C., Asunción Durá, M., Noirclerc-Savoye, M., Perollier, A., Gallet, B., Plevin, M. J., et al. (2011). A Systematic Mutagenesis-Driven Strategy for Site-Resolved NMR Studies of Supramolecular Assemblies. *J. Biomol. NMR* 50 (3), 229–236. doi:10.1007/s10858-011-9513-5

Andreas, L. B., Jaudzems, K., Stanek, J., Lalli, D., Bertarello, A., Le Marchand, T., et al. (2016). Structure of Fully Protonated Proteins by Proton-Detected Magic-Angle Spinning NMR. *Proc. Natl. Acad. Sci. USA* 113 (33), 9187–9192. doi:10.1073/pnas.1602248113

Andreas, L. B., Le Marchand, T., Jaudzems, K., and Pintacuda, G. (2015b). High-Resolution Proton-Detected NMR of Proteins at Very Fast MAS. *J. Magn. Reson.* 253, 36–49. doi:10.1016/j.jmr.2015.01.003

Andreas, L. B., Reese, M., Eddy, M. T., Gelev, V., Ni, Q. Z., Miller, E. A., et al. (2015a). Structure and Mechanism of the Influenza A M218-60 Dimer of Dimers. *J. Am. Chem. Soc.* 137 (47), 14877–14886. doi:10.1021/jacs.5b04802

Asami, S., and Reif, B. (2012). Assignment Strategies for Aliphatic Protons in the Solid-State in Randomly Protonated Proteins. *J. Biomol. NMR* 52 (1), 31–39. doi:10.1007/s10858-011-9591-4

Asami, S., and Reif, B. (2013). Proton-Detected Solid-State NMR Spectroscopy at Aliphatic Sites: Application to Crystalline Systems. *Acc. Chem. Res.* 46 (9), 2089–2097. doi:10.1021/ar400063y

Asami, S., Schmieder, P., and Reif, B. (2010). High Resolution ^{1}H-Detected Solid-State NMR Spectroscopy of Protein Aliphatic Resonances: Access to Tertiary Structure Information. *J. Am. Chem. Soc.* 132 (43), 15133–15135. doi:10.1021/ja106170h

Asami, S., Szekely, K., Schanda, P., Meier, B. H., and Reif, B. (2012). Optimal Degree of Protonation for ^{1}H Detection of Aliphatic Sites in Randomly Deuterated Proteins as a Function of the MAS Frequency. *J. Biomol. NMR* 54 (2), 155–168. doi:10.1007/s10858-012-9659-9

Bak, M., and Nielsen, N. C. (1997). REPULSION, A Novel Approach to Efficient Powder Averaging in Solid-State NMR. *J. Magn. Reson.* 125 (1), 132–139. doi:10.1006/jmre.1996.1087

Bak, M., Rasmussen, J. T., and Nielsen, N. C. (2011). SIMPSON: A General Simulation Program for Solid-State NMR Spectroscopy. *J. Magn. Reson.* 213 (2), 366–400. doi:10.1016/j.jmr.2011.09.008

Bak, M., Schultz, R., Vosegaard, T., and Nielsen, N. C. (2002). Specification and Visualization of Anisotropic Interaction Tensors in Polypeptides and Numerical Simulations in Biological Solid-State NMR. *J. Magn. Reson.* 154 (1), 28–45. doi:10.1006/jmre.2001.2454

Baker, L. A., and Baldus, M. (2014). Characterization of Membrane Protein Function by Solid-State NMR Spectroscopy. *Curr. Opin. Struct. Biol.* 27 (August), 48–55. doi:10.1016/j.sbi.2014.03.009

Baldus, M., and Meier, B. H. (1996). Total Correlation Spectroscopy in the Solid State. The Use of Scalar Couplings to Determine the Through-Bond Connectivity. *J. Magn. Reson. Ser. A* 121 (1), 65–69. doi:10.1006/jmra.1996.0137

Bax, A., Clore, G. M., and Gronenborn, A. M. (1990). ^{1}H^{1}H Correlation via Isotropic Mixing of ^{13}C Magnetization, a New Three-Dimensional Approach for Assigning ^{1}H and ^{13}C Spectra of ^{13}C-Enriched Proteins. *J. Magn. Reson. (1969)* 88 (2), 425–431. doi:10.1016/0022-2364(90)90202-K

Bennett, A. E., Rienstra, C. M., Auger, M., Lakshmi, K. V., and Griffin, R. G. (1995). Heteronuclear Decoupling in Rotating Solids. *J. Chem. Phys.* 103 (16), 6951–6958. doi:10.1063/1.470372

Bennett, A. E., Rienstra, C. M., Griffiths, J. M., Zhen, W., Lansbury, P. T., and Griffin, R. G. (1998). Homonuclear Radio Frequency-Driven Recoupling in Rotating Solids. *J. Chem. Phys.* 108 (22), 9463–9479. doi:10.1063/1.476420

Boswell, Z. K., and Latham, M. P. (2018). Methyl-Based NMR Spectroscopy Methods for Uncovering Structural Dynamics in Large Proteins and Protein Complexes. *Biochemistry* 58 (3), 144–155. doi:10.1021/acs.biochem.8b00953

Chevelkov, V., Rehbein, K., Diehl, A., and Reif, B. (2006). Ultrahigh Resolution in Proton Solid-State NMR Spectroscopy at High Levels of Deuteration. *Angew. Chem. Int. Ed.* 45 (23), 3878–3881. doi:10.1002/anie.200600328

De Paëpe, G., Lewandowski, J. R., and Griffin, R. G. (2008). Spin Dynamics in the Modulation Frame: Application to Homonuclear Recoupling in Magic Angle Spinning Solid-State NMR. *J. Chem. Phys.* 128 (12), 124503. doi:10.1063/1.2834732

Edén, M., and Levitt, M. H. (1999). Pulse Sequence Symmetries in the Nuclear Magnetic Resonance of Spinning Solids: Application to Heteronuclear Decoupling. *J. Chem. Phys.* 111 (4), 1511–1519. doi:10.1063/1.479410

Emsley, L., and Bodenhausen, G. (1992). Optimization of Shaped Selective Pulses for NMR Using a Quaternion Description of Their Overall Propagators. *J. Magn. Reson. (1969)* 97 (1), 135–148. doi:10.1016/0022-2364(92)90242-Y

Fraga, H., Arnaud, C. A., Gauto, D. F., Audin, M., Kurauskas, V., Macek, P., et al. (2017). Solid-State NMR H–N–(C)–H and H–N–C–C 3D/4D Correlation Experiments for Resonance Assignment of Large Proteins. *ChemPhysChem* 18 (19), 2697–2703. doi:10.1002/cphc.201700572

Ganapathy, S., Naito, A., and McDowell, C. A. (1981). Paramagnetic Doping as an Aid in Obtaining High-Resolution Carbon-13 NMR Spectra of Biomolecules in the Solid State. *J. Am. Chem. Soc.* 103 (20), 6011–6015. doi:10.1021/ja00410a003

Gauto, D. F., Estrozi, L. F., Schwieters, C. D., Effantin, G., Macek, P., Sounier, R., et al. (2019). Integrated NMR and Cryo-EM Atomic-Resolution Structure Determination of a Half-Megadalton Enzyme Complex. *Nat. Commun.* 10 (1), 2697. doi:10.1038/s41467-019-10490-9

Gorman, S. D., Sahu, D., O'Rourke, K. F., and Boehr, D. D. (2018). Assigning Methyl Resonances for Protein Solution-State NMR Studies. *Methods* 148 (September), 88–99. doi:10.1016/j.ymeth.2018.06.010

Goto, N. K., Gardner, K. H., Mueller, G. A., Willis, R. C., and Kay, L. E. (1999). A Robust and Cost-Effective Method for the Production of Val, Leu, Ile (δ1) Methyl-Protonated ^{15}N-, ^{13}C-, ^{2}H-Labeled Proteins. *J. Biomol. NMR* 13 (4), 369–374. doi:10.1023/A:1008393201236

Griffin, R. (1998). Dipolar Recoupling in MAS Spectra of Biological Solids. *Nat. Struct. Mol. Biol.* 5 (7), 508–512. doi:10.1038/749

Hardy, E. H., Verel, R., and Meier, B. H. (2001). Fast MAS Total Through-Bond Correlation Spectroscopy. *J. Magn. Reson.* 148 (2), 459–464. doi:10.1006/jmre.2000.2258

Hou, G., Yan, S., Trébosc, J., Amoureux, J.-P., and Polenova, T. (2013). Broadband Homonuclear Correlation Spectroscopy Driven by Combined $R2^{v}_{n}$ Sequences under Fast Magic Angle Spinning for NMR Structural Analysis of Organic and Biological Solids. *J. Magn. Reson.* 232, 18–30. doi:10.1016/j.jmr.2013.04.009

Hu, B., Lafon, O., Trébosc, J., Chen, Q., and Amoureux, J.-P. (2011). Broad-Band Homo-Nuclear Correlations Assisted by ^{1}H Irradiation for Bio-Molecules in Very High Magnetic Field at Fast and Ultra-fast MAS Frequencies. *J. Magn. Reson.* 212 (2), 320–329. doi:10.1016/j.jmr.2011.07.011

Huang, C., and Kalodimos, C. G. (2017). Structures of Large Protein Complexes Determined by Nuclear Magnetic Resonance Spectroscopy. *Annu. Rev. Biophys.* 46 (1), 317–336. doi:10.1146/annurev-biophys-070816-033701

Huber, M., Hiller, S., Schanda, P., Ernst, M., Böckmann, A., Verel, R., et al. (2011). A Proton-Detected 4D Solid-State NMR Experiment for Protein Structure Determination. *ChemPhysChem* 12 (5), 915–918. doi:10.1002/cphc.201100062

Kadkhodaie, M., Rivas, O., Tan, M., Mohebbi, A., and Shaka, A. J. (1991). Broadband Homonuclear Cross Polarization Using Flip-Flop Spectroscopy. *J. Magn. Reson. (1969)* 91 (2), 437–443. doi:10.1016/0022-2364(91)90210-K

Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R., and Forman-Kay, J. D. (1993). A Gradient-Enhanced HCCH-TOCSY Experiment for Recording Side-Chain ^{1}H and ^{13}C Correlations in H_2O Samples of Proteins. *J. Magn. Reson. Ser. B* 101 (3), 333–337. doi:10.1006/jmrb.1993.1053

Kulminskaya, N., Vasa, S. K., Giller, K., Becker, S., Kwan, A., Sunde, M., et al. (2016). Access to Side-Chain Carbon Information in Deuterated Solids under Fast MAS through Non-rotor-synchronized Mixing. *Chem. Commun.* 52 (2), 268–271. doi:10.1039/C5CC07345F

Kurauskas, V., Crublet, E., Macek, P., Kerfah, R., Gauto, D. F., Boisbouvier, J., et al. (2016). Sensitive Proton-Detected Solid-State NMR Spectroscopy of Large Proteins with Selective CH3 Labelling: Application to the 50S Ribosome Subunit. *Chem. Commun.* 52 (61), 9558–9561. doi:10.1039/C6CC04484K

Ladizhansky, V. (2017). Applications of Solid-State NMR to Membrane Proteins. *Biochim. Biophys. Acta (Bba) - Proteins Proteomics* 1865 (11), 1577–1586. doi:10.1016/j.bbapap.2017.07.004

Lee, W., Tonelli, M., and Markley, J. L. (2015). NMRFAM-SPARKY: Enhanced Software for Biomolecular NMR Spectroscopy. *Bioinformatics* 31 (8), 1325–1327. doi:10.1093/bioinformatics/btu830

Levitt, M. H. (2007). "eMagRes,". Editors R. K. Harris and R. L. Wasylishen (Chichester, UK: John Wiley & Sons), 9. doi:10.1002/9780470034590

Lewandowski, J. R., Dumez, J.-N., Akbey, Ü., Lange, S., Emsley, L., and Oschkinat, H. (2011). Enhanced Resolution and Coherence Lifetimes in the Solid-State NMR Spectroscopy of Perdeuterated Proteins under Ultrafast Magic-Angle Spinning. *J. Phys. Chem. Lett.* 2 (17), 2205–2211. doi:10.1021/jz200844n

Linser, R., Bardiaux, B., Andreas, L. B., Hyberts, S. G., Morris, V. K., Pintacuda, G., et al. (2014). Solid-State NMR Structure Determination from Diagonal-Compensated, Sparsely Nonuniform-Sampled 4D Proton-Proton Restraints. *J. Am. Chem. Soc.* 136 (31), 11002–11010. doi:10.1021/ja504603g

Linser, R., Bardiaux, B., Higman, V., Fink, U., and Reif, B. (2011). Structure Calculation from Unambiguous Long-Range Amide and Methyl ^{1}H–^{1}H Distance Restraints for a Microcrystalline Protein with MAS Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 133 (15), 5905–5912. doi:10.1021/ja110222h

Linser, R. (2011). Side-Chain to Backbone Correlations from Solid-State NMR of Perdeuterated Proteins through Combined Excitation and Long-Range Magnetization Transfers. *J. Biomol. NMR* 51 (3), 221–226. doi:10.1007/s10858-011-9531-3

Linser, R. (2017). Solid-State NMR Spectroscopic Trends for Supramolecular Assemblies and Protein Aggregates. *Solid State. Nucl. Magn. Reson.* 87 (August), 45–53. doi:10.1016/j.ssnmr.2017.08.003

Lu, X., Guo, C., Hou, G., and Polenova, T. (2015). Combined Zero-Quantum and Spin-Diffusion Mixing for Efficient Homonuclear Correlation Spectroscopy under Fast MAS: Broadband Recoupling and Detection of Long-Range Correlations. *J. Biomol. NMR* 61 (1), 7–20. doi:10.1007/s10858-014-9875-6

Lundström, P., Teilum, K., Carstensen, T., Bezsonova, I., Wiesner, S., Hansen, D. F., et al. (2007). Fractional ^{13}C Enrichment of Isolated Carbons Using [1-^{13}C]- or [2-^{13}C]-Glucose Facilitates the Accurate Measurement of Dynamics at Backbone Cα and Side-Chain Methyl Positions in Proteins. *J. Biomol. NMR* 38 (3), 199–212. doi:10.1007/s10858-007-9158-6

Mainz, A., Religa, T. L., Sprangers, R., Linser, R., Kay, L. E., and Reif, B. (2013). NMR Spectroscopy of Soluble Protein Complexes at One Mega-Dalton and beyond. *Angew. Chem. Int. Ed.* 52 (33), 8746–8751. doi:10.1002/anie.201301215

Mandala, V. S., Williams, J. K., and Hong, M. (2018). Structure and Dynamics of Membrane Proteins from Solid-State NMR. *Annu. Rev. Biophys.* 47 (1), 201–222. doi:10.1146/annurev-biophys-070816-033712

Marchanka, A., Stanek, J., Pintacuda, G., and Carlomagno, T. (2018). Rapid Access to RNA Resonances by Proton-Detected Solid-State NMR at >100 kHz MAS. *Chem. Commun.* 54 (65), 8972–8975. doi:10.1039/c8cc04437f

Miao, Y., and Cross, T. A. (2013). Solid State NMR and Protein-Protein Interactions in Membranes. *Curr. Opin. Struct. Biol.* 23 (6), 919–928. doi:10.1016/j.sbi.2013.08.004

Paramasivam, S., Suiter, C. L., Hou, G., Sun, S., Palmer, M., Hoch, J. C., et al. (2012). Enhanced Sensitivity by Nonuniform Sampling Enables Multidimensional MAS NMR Spectroscopy of Protein Assemblies. *J. Phys. Chem. B* 116 (25), 7416–7427. doi:10.1021/jp3032786

Pauli, J., Baldus, M., Van Rossum, B., De Groot, H., and Oschkinat, H. (2001). Backbone and Side-Chain ^{13}C and ^{15}N Signal Assignments of the α-Spectrin SH3 Domain by Magic Angle Spinning Solid-State NMR at 17.6 Tesla. *ChemBioChem* 2 (4), 272–281. doi:10.1002/1439-7633(20010401)2:4<272::aid-cbic272>3.0.co;2-2

Pauli, J., van Rossum, B., Förster, H., de Groot, H. J. M., and Oschkinat, H. (2000). Sample Optimization and Identification of Signal Patterns of Amino Acid Side Chains in 2D RFDR Spectra of the α-Spectrin SH3 Domain. *J. Magn. Reson.* 143 (2), 411–416. doi:10.1006/jmre.2000.2029

Penzel, S., Smith, A. A., Agarwal, V., Hunkeler, A., Org, M.-L., Samoson, A., et al. (2015). Mai-Liis Org, Ago Samoson, Anja Böckmann, Matthias Ernst, and Beat H MeierProtein Resonance Assignment at MAS Frequencies Approaching 100 kHz: A Quantitative Comparison of J-Coupling and Dipolar-Coupling-Based Transfer Methods. *J. Biomol. NMR* 63 (2), 165–186. doi:10.1007/s10858-015-9975-y

Pritišanac, I., Alderson, T. R., and Güntert, P. (2020). Automated Assignment of Methyl NMR Spectra from Large Proteins. *Prog. Nucl. Magn. Reson. Spectrosc.* 118-119 (June), 54–73. doi:10.1016/j.pnmrs.2020.04.001

Pritišanac, I., Würz, J. M., Alderson, T. R., and Güntert, P. (2019). Automatic Structure-Based NMR Methyl Resonance Assignment in Large Proteins. *Nat. Commun.* 10 (1), 1–12. doi:10.1038/s41467-019-12837-8

Reverter, D., and Lima, C. D. (2009). Preparation of SUMO Proteases and Kinetic Analysis Using Endogenous Substrates. *Methods Mol. Biolin* 497, 225–239. doi:10.1007/978-1-59745-566-4_15

Rosenzweig, R., and Kay, L. E. (2014). Bringing Dynamic Molecular Machines into Focus by Methyl-TROSY NMR. *Annu. Rev. Biochem.* 83 (1), 291–315. doi:10.1146/annurev-biochem-060713-035829

Sattler, M., Schleucher, J., and Griesinger, C. (1999). Heteronuclear Multidimensional NMR Experiments for the Structure Determination of Proteins in Solution Employing Pulsed Field Gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* 34 (2), 93–158. doi:10.1016/S0079-6565(98)00025-9

Schmidt, E., and Güntert, P. (2012). A New Algorithm for Reliable and General NMR Resonance Assignment. *J. Am. Chem. Soc.* 134 (30), 12817–12829. doi:10.1021/ja305091n

Schmidt, E., and Güntert, P. (2013). Reliability of Exclusively NOESY-Based Automated Resonance Assignment and Structure Determination of Proteins. *J. Biomol. NMR* 57 (2), 193–204. doi:10.1007/s10858-013-9779-x

Scholz, I., Huber, M., Manolikas, T., Meier, B. H., and Ernst, M. (2008). MIRROR Recoupling and its Application to Spin Diffusion under Fast Magic-Angle Spinning. *Chem. Phys. Lett.* 460 (1), 278–283. doi:10.1016/j.cplett.2008.05.058

Sergeyev, I. V., Itin, B., Rogawski, R., Day, L. A., and McDermott, A. E. (2017). Efficient Assignment and NMR Analysis of an Intact Virus Using Sequential

Side-Chain Correlations and DNP Sensitization. *Proc. Natl. Acad. Sci. USA* 114 (20), 5171–5176. doi:10.1073/pnas.1701484114

Shaka, A. J., Keeler, J., and Freeman, R. (1983). Evaluation of a New Broadband Decoupling Sequence: WALTZ-16. *J. Magn. Reson. (1969)* 53 (2), 313–340. doi:10.1016/0022-2364(83)90035-5

Shaka, A. J., Lee, C. J., and Pines, A. (1988). Iterative Schemes for Bilinear Operators; Application to Spin Decoupling. *J. Magn. Reson. (1969)* 77 (2), 274–293. doi:10.1016/0022-2364(88)90178-3

Shen, M., Hu, B., Lafon, O., Trébosc, J., Chen, Q., and Amoureux, J.-P. (2012). Broadband Finite-Pulse Radio-Frequency-Driven Recoupling (Fp-RFDR) with (XY8)41 Super-cycling for Homo-Nuclear Correlations in Very High Magnetic Fields at Fast and Ultra-fast MAS Frequencies. *J. Magn. Reson.* 223 (October), 107–119. doi:10.1016/j.jmr.2012.07.013

Sinnige, T., Daniëls, M., Baldus, M., and Weingarth, M. (2014). Proton Clouds to Measure Long-Range Contacts between Nonexchangeable Side Chain Protons in Solid-State NMR. *J. Am. Chem. Soc.* 136 (12), 4452–4455. doi:10.1021/ja412870m

Stanek, J., Andreas, L. B., Jaudzems, K., Cala, D., Lalli, D., Bertarello, A., et al. (2016). NMR Spectroscopic Assignment of Backbone and Side-Chain Protons in Fully Protonated Proteins: Microcrystals, Sedimented Assemblies, and Amyloid Fibrils. *Angew. Chem. Int. Ed.* 55 (50), 15504–15509. doi:10.1002/anie.201607084

Stanek, J., Augustyniak, R., and Koźmiński, W. (2012). Suppression of Sampling Artefacts in High-Resolution Four-Dimensional NMR Spectra Using Signal Separation Algorithm. *J. Magn. Reson.* 214 (January), 91–102. doi:10.1016/j.jmr.2011.10.009

Stanek, J., Schubeis, T., Paluch, P., Güntert, P., Andreas, L. B., and Pintacuda, G. (2020). Automated Backbone NMR Resonance Assignment of Large Proteins Using Redundant Linking from a Single Simultaneous Acquisition. *J. Am. Chem. Soc.* 142 (12), 5793–5799. doi:10.1021/jacs.0c00251

Tan, K. O., Agarwal, V., Lakomek, N.-A., Penzel, S., Meier, B. H., and Ernst, M. (2018). Efficient Low-Power TOBSY Sequences for Fast MAS. *Solid State. Nucl. Magn. Reson.* 89, 27–34. doi:10.1016/j.ssnmr.2017.11.003

Tošner, Z., Andersen, R., Stevensson, B., Edén, M., Nielsen, N. C., and Vosegaard, T. (2014). Computer-Intensive Simulation of Solid-State NMR Experiments Using SIMPSON. *J. Magn. Reson.* 246 (September), 79–93. doi:10.1016/j.jmr.2014.07.002

Tošner, Z., Sarkar, R., Becker-Baldus, J., Glaubitz, C., Wegner, S., Engelke, F., et al. (2018). Overcoming Volume Selectivity of Dipolar Recoupling in Biological Solid-State NMR Spectroscopy. *Angew. Chem. Int. Ed.* 57 (44), 14514–14518. doi:10.1002/anie.201805002

Tugarinov, V., Hwang, P. M., Ollerenshaw, J. E., and Kay, L. E. (2003). Cross-Correlated Relaxation Enhanced ^{1}H–^{13}C NMR Spectroscopy of Methyl Groups in Very High Molecular Weight Proteins and Protein Complexes. *J. Am. Chem. Soc.* 125 (34), 10420–10428. doi:10.1021/ja030153x

Tugarinov, V., Kanelis, V., and Kay, L. E. (2006). Isotope Labeling Strategies for the Study of High-Molecular-Weight Proteins by Solution NMR Spectroscopy. *Nat. Protoc.* 1 (2), 749–754. doi:10.1038/nprot.2006.101

Tugarinov, V., and Kay, L. E. (2004). An Isotope Labeling Strategy for Methyl TROSY Spectroscopy. *J. Biomol. NMR* 28 (2), 165–172. doi:10.1023/B:JNMR.0000013824.93994.1f

Tugarinov, V., and Kay, L. E. (2003a). Ile, Leu, and Val Methyl Assignments of the 723-Residue Malate Synthase G Using a New Labeling Strategy and Novel NMR Methods. *J. Am. Chem. Soc.* 125 (45), 13868–13878. doi:10.1021/ja030345s

Tugarinov, V., and Kay, L. E. (2003b). Side Chain Assignments of Ile δ1 Methyl Groups in High Molecular Weight Proteins: an Application to a 46 ns Tumbling Molecule. *J. Am. Chem. Soc.* 125 (19), 5701–5706. doi:10.1021/ja021452+

Tycko, R. (2011). Solid-State NMR Studies of Amyloid Fibril Structure. *Annu. Rev. Phys. Chem.* 62 (1), 279–299. doi:10.1146/annurev-physchem-032210-103539

van der Wel, P. C. A. (2018). New Applications of Solid-State NMR in Structural Biology. *Emerging Top. Life Sci.* 2 (1), 57–67. doi:10.1042/ETLS20170088

Vasa, S. K., Singh, H., Rovó, P., Linser, R., and Linser, R. (2018). Dynamics and Interactions of a 29 KDa Human Enzyme Studied by Solid-State NMR. *J. Phys. Chem. Lett.* 9 (6), 1307–1311. doi:10.1021/acs.jpclett.8b00110

Verel, R., Baldus, M., Ernst, M., Beat, H., and Meier, H. (1998). A Homonuclear Spin-Pair Filter for Solid-State NMR Based on Adiabatic-Passage Techniques. *Chem. Phys. Lett.* 287 (3–4), 421–428. doi:10.1016/S0009-2614(98)00172-9

Verel, R., Ernst, M., and Meier, B. H. (2001). Adiabatic Dipolar Recoupling in Solid-State NMR: The DREAM Scheme. *J. Magn. Reson.* 150 (1), 81–99. doi:10.1006/jmre.2001.2310

Weingarth, M., Demco, D. E., Bodenhausen, G., and Tekely, P. (2009). Improved Magnetization Transfer in Solid-State NMR with Fast Magic Angle Spinning. *Chem. Phys. Lett.* 469 (4), 342–348. doi:10.1016/j.cplett.2008.12.084

Westfeld, T., Verel, R., Ernst, M., Böckmann, A., and Meier, B. H. (2012). Properties of the DREAM Scheme and its Optimization for Application to Proteins. *J. Biomol. NMR* 53 (2), 103–112. doi:10.1007/s10858-012-9627-4

Wickramasinghe, N. P., Kotecha, M., Samoson, A., Past, J., and Ishii, Y. (2007). Sensitivity Enhancement in ^{13}C Solid-State NMR of Protein Microcrystals by Use of Paramagnetic Metal Ions for Optimizing ^{1}H T_1 Relaxation. *J. Magn. Reson.* 184 (2), 350–356. doi:10.1016/j.jmr.2006.10.012

Wickramasinghe, N. P., Parthasarathy, S., Jones, C. R., Bhardwaj, C., Long, F., Kotecha, M., et al. (2009). Nanomole-Scale Protein Solid-State NMR by Breaking Intrinsic ^{1}H T_1 Boundaries. *Nat. Methods* 6 (3), 215–218. doi:10.1038/nmeth.1300

Wiegand, T. (2020). A Solid-State NMR Tool Box for the Investigation of ATP-Fueled Protein Engines. *Prog. Nucl. Magn. Reson. Spectrosc.* 117 (April), 1–32. doi:10.1016/j.pnmrs.2020.02.001

Xiang, S., Chevelkov, V., Becker, S., and Lange, A. (2014). Towards Automatic Protein Backbone Assignment Using Proton-Detected 4D Solid-State NMR Data. *J. Biomol. NMR* 60 (2–3), 85–90. doi:10.1007/s10858-014-9859-6

Xue, K., Motz, C., Asami, S., Decker, V., Wegner, S., Tosner, Z., et al. (2018). Magic-Angle Spinning Frequencies beyond 300 kHz Are Necessary to Yield Maximum Sensitivity in Selectively Methyl Protonated Protein Samples in Solid-State NMR. *J. Phys. Chem. C* 122, 16437–16442. doi:10.1021/acs.jpcc.8b05600

Xue, K., Sarkar, R., Lalli, D., Koch, B., Pintacuda, G., Tosner, Z., et al. (2020). Impact of Magnetic Field Strength on Resolution and Sensitivity of Proton Resonances in Biological Solids. *J. Phys. Chem. C* 124 (41), 22631–22637. doi:10.1021/acs.jpcc.0c05407

Xue, K., Sarkar, R., Motz, C., Asami, S., Camargo, D. C. R., Decker, V., et al. (2017). Limits of Resolution and Sensitivity of Proton Detected MAS Solid-State NMR Experiments at 111 kHz in Deuterated and Protonated Proteins. *Sci. Rep.* 7 (1), 7444. doi:10.1038/s41598-017-07253-1

Xue, K., Sarkar, R., Tosner, Z., Lalli, D., Motz, C., Koch, B., et al. (2019). MAS Dependent Sensitivity of Different Isotopomers in Selectively Methyl Protonated Protein Samples in Solid State NMR. *J. Biomol. NMR* 73 (10–11), 625–631. doi:10.1007/s10858-019-00274-0

Zhou, D. H., and Rienstra, C. M. (2008). High-Performance Solvent Suppression for Proton Detected Solid-State NMR. *J. Magn. Reson.* 192 (1), 167–172. doi:10.1016/j.jmr.2008.01.012

Zhou, D. H., Shah, G., Cormos, M., Mullen, C., Sandoz, D., and Rienstra, C. M. (2007). Proton-Detected Solid-State NMR Spectroscopy of Fully Protonated Proteins at 40 kHz Magic-Angle Spinning. *J. Am. Chem. Soc.* 129 (38), 11791–11801. doi:10.1021/ja073462m

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

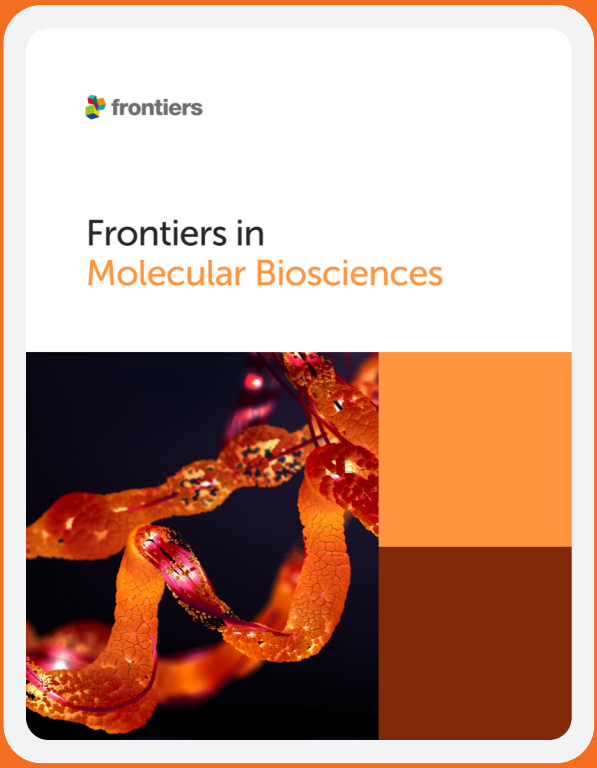
Frontiers in Molecular Biosciences

Explores biological processes in living organisms on a molecular scale

Focuses on the molecular mechanisms underpinning and regulating biological processes in organisms across all branches of life.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact