



Advances in RNA 3D Structure Modeling Using Experimental Data

Bing Li¹, Yang Cao¹, Eric Westhof² and Zhichao Miao^{3,4,5*}

¹ Center of Growth, Metabolism and Aging, Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu, China, ² Architecture et Réactivité de l'ARN, Institut de Biologie Moléculaire et Cellulaire du CNRS, Université de Strasbourg, Strasbourg, France, ³ Translational Research Institute of Brain and Brain-Like Intelligence, Department of Anesthesiology, Shanghai Fourth People's Hospital Affiliated to Tongji University School of Medicine, Shanghai, China, ⁴ Newcastle Fibrosis Research Group, Institute of Cellular Medicine, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom, ⁵ European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Cambridge, United Kingdom

OPEN ACCESS

Edited by:

Philipp Kapranov,
Huaqiao University, China

Reviewed by:

Yiliang Ding,
John Innes Centre, United Kingdom
Peter F. Stadler,
Leipzig University, Germany
Francesco Stellato,
University of Rome Tor Vergata, Italy

*Correspondence:

Zhichao Miao
zmiao@ebi.ac.uk

Specialty section:

This article was submitted to
RNA,
a section of the journal
Frontiers in Genetics

Received: 19 June 2020

Accepted: 02 September 2020

Published: 26 October 2020

Citation:

Li B, Cao Y, Westhof E and Miao Z
(2020) Advances in RNA 3D Structure
Modeling Using Experimental Data.
Front. Genet. 11:574485.
doi: 10.3389/fgene.2020.574485

RNA is a unique bio-macromolecule that can both record genetic information and perform biological functions in a variety of molecular processes, including transcription, splicing, translation, and even regulating protein function. RNAs adopt specific three-dimensional conformations to enable their functions. Experimental determination of high-resolution RNA structures using x-ray crystallography is both laborious and demands expertise, thus, hindering our comprehension of RNA structural biology. The computational modeling of RNA structure was a milestone in the birth of bioinformatics. Although computational modeling has been greatly improved over the last decade showing many successful cases, the accuracy of such computational modeling is not only length-dependent but also varies according to the complexity of the structure. To increase credibility, various experimental data were integrated into computational modeling. In this review, we summarize the experiments that can be integrated into RNA structure modeling as well as the computational methods based on these experimental data. We also demonstrate how computational modeling can help the experimental determination of RNA structure. We highlight the recent advances in computational modeling which can offer reliable structure models using high-throughput experimental data.

Keywords: RNA structure, chemical probing, 3D shape, structure prediction, RNA-puzzles

THE PROBLEM OF RNA 3D STRUCTURE DETERMINATION AND ITS HISTORY

Ribonucleic acids or RNAs play significant roles in a great variety of biological processes throughout the central dogma (Sharp, 2009; Cech and Steitz, 2014; Strobel et al., 2016), ranging from transcription regulation (Long et al., 2017), to RNA splicing (Buratti and Baralle, 2004; Luco and Misteli, 2011), and protein synthesis (Valencia-Sanchez et al., 2006). A recent study revealed the fact that RNA may act as a riboregulator of autophagy through the regulation of protein polymerization (Horos et al., 2019), which indicates the function of the proteins are not only regulated by transcription and translation but also by interaction with the RNA structure. The functional diversity of RNA arises from its ability to form specific 3D structures that can function in response to cellular signals (Al-Hashimi and Walter, 2008; Mustoe et al., 2014, 2018). An

example in disease mechanism from recent research (Cammass and Millevoi, 2017) shows RNA G-quadruplexes, folded into a four-stranded conformation, revealing new mechanisms in disease. According to these examples, determining and modeling the RNA structure can substantially contribute to our understanding of biological processes, disease mechanisms, and RNA therapies.

As early as 1969, the first manually predicted tertiary structure of tRNA (Levitt, 1969) was regarded as a milestone in the emergence of bioinformatics. That structure was based on available tRNA sequences and some scattered experimental data [like the cross-link between positions 8 and 13 (Yaniv et al., 1969)]. In 1989, the model of the core of group I intron (Michel and Westhof, 1990) was based on extensive sequence comparisons, clustering secondary structures into distinct classes, and the published experimental data based on mutagenesis. Since then, the RNA structure modeling approaches, including secondary and tertiary structure modeling, have been in intense development. A variety of useful programs were produced in this period [for recent reviews see (Miao and Westhof, 2017; Ponce-Salvatierra et al., 2019)]. With the development of whole-genome sequencing techniques and the availability of various metagenome sequences, hundreds of RNA families have been discovered (Weinberg et al., 2007, 2009, 2010) which greatly expanded our knowledge of the RNA sequence space. Almost all of these RNA sequence families have been described in the database of Rfam (Kalvari et al., 2018). However, up to January 2020, only 99 RNA families in the Rfam database have experimentally determined structures available in the Protein Data Bank (PDB) (Berman et al., 2000) or the Nucleic Acids Database (NDB) (Berman et al., 1992; Coimbatore Narayanan et al., 2014).

The experimental methods of RNA structure determination can be classified as biophysical and biochemical methods. Biophysical experiments such as x-ray crystallography (Suddala and Zhang, 2019), small-angle scattering (SAS) (Hura et al., 2009), and cryogenic electron microscopy (cryo-EM) (Zhang et al., 2019) are uncovering the structural basis of RNA functions at nanometer- or angstrom-level resolutions. Alternatively, biochemical approaches [e.g., chemical probing (Peattie and Herr, 1981)], have been used systematically to validate RNA structures. The recent coupling of RNA structure probing with high-throughput sequencing (Strobel et al., 2018) has changed the estimation from the electrophoresis on denaturing gels (Lucks et al., 2011) to the omics techniques, allowing higher for throughput in RNA structure characterization. Instead of the atomistic models, biochemical approaches promote the experimental flexibility and throughput by sacrificing the resolution. They determine RNA structure using the computational approaches based on the restraints obtained from experiments. Structure dynamics, folding, and *in vivo* structure determination have become a new field known as conformational ensembles, or structure ensembles. The ensembles are the set of all dynamic structure conformations, capturing the structure motions in a large range of energy landscapes and timescales. Conformational ensembles are critical

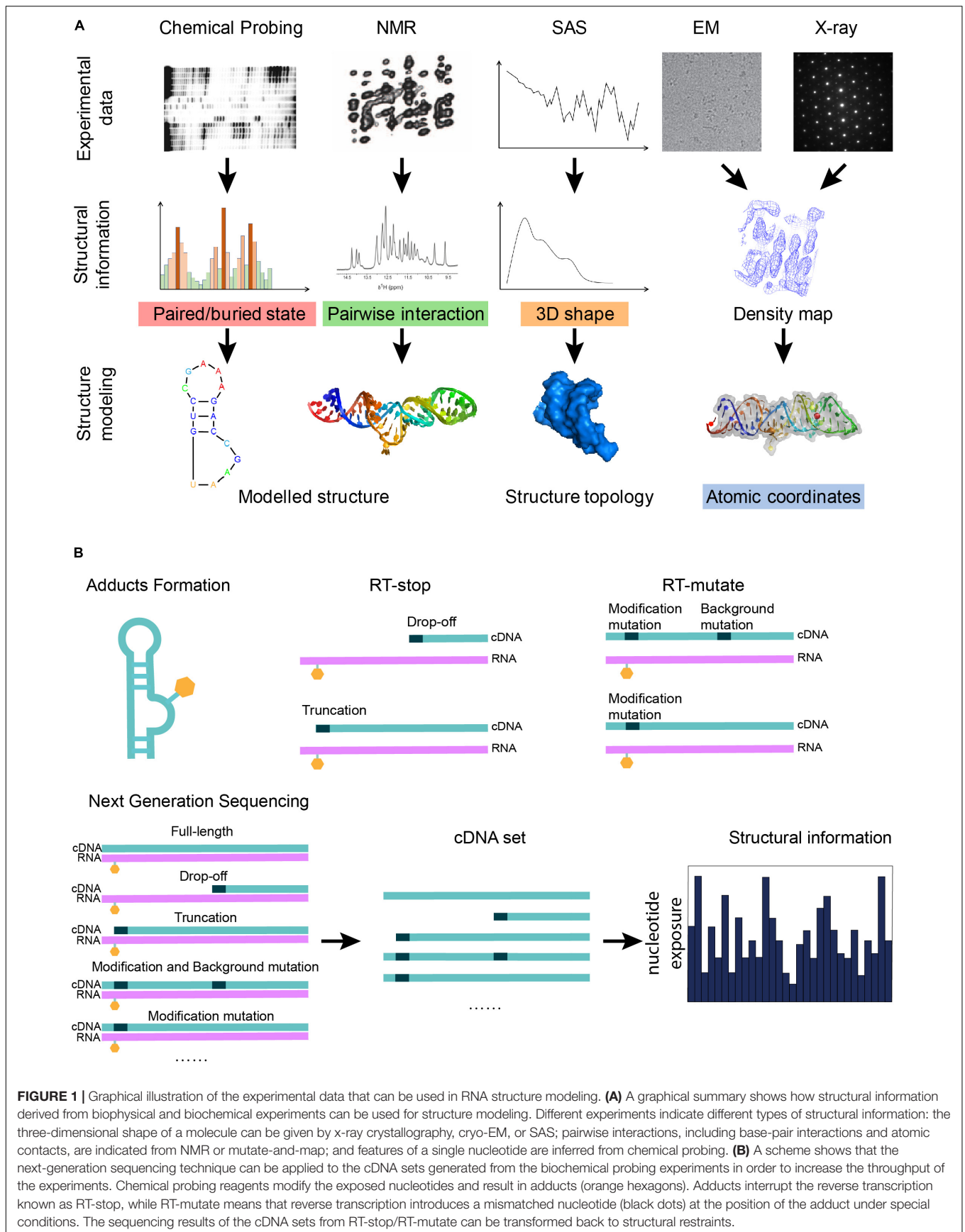
in understanding the cellular function of RNAs (Ganser et al., 2019) and computational methods with experimental data have been developed (Salmon et al., 2014). And these methods have been discussed in previous reviews (Salmon et al., 2014; Ganser et al., 2019).

Together with the advances in experiments, computational modeling or prediction methods of RNA secondary and tertiary structure (Magnus et al., 2014; Ponce-Salvatierra et al., 2019) are being developed and improved to help and complement experimental efforts. Similar to the efforts in protein structure prediction, RNA 3D structure modeling has used approaches including homology modeling (Flores and Altman, 2010; Rother et al., 2011b), fragment assembly (Das and Baker, 2007; Bida and Maher, 2012; Popena et al., 2012; Zhao et al., 2012; Xu et al., 2019), and *de novo* prediction (Sharma et al., 2008; Jonikas et al., 2009; Krokhotin et al., 2015; Boniecki et al., 2016). The direct coupling analysis approach (De Leonardis et al., 2015; Weinreb et al., 2016; Wang et al., 2017), which is based on the alignments of metagenome sequence information, also shows its ability in RNA structure prediction. Since 2011, RNA-Puzzles (Cruz et al., 2012; Miao et al., 2015, 2017, 2020), which is a community effort for evaluating these RNA 3D structure prediction methods, has reported three rounds of predictions. It revealed existing bottlenecks in RNA structure modeling, including the prediction of *non*-Watson-Crick interactions, atomic clashes in the models, and the challenges in ligand binding predictions.

The recent improvements about static structure determination, achieved by coupling high-throughput experimental data with computational modeling (Cheng et al., 2015, 2017; Kappel et al., 2019), have demonstrated great potential in RNA 3D structure characterization. In this review, we provide a concise overview of the existing experimental data in RNA structure determination and the structure modeling approaches. In particular, we highlight the recent achievements in experimental data-driven RNA structure modeling, which may revolutionize RNA structural biology in a high-throughput way in the near future.

RNA STRUCTURE CHARACTERIZATION EXPERIMENTS

Both biophysical and biochemical approaches (**Figure 1A**) have been conceived to determine RNA structures, dynamics, and interactions with other biomolecules. Biophysical experiments are normally used to generate 3D structural information in the form of shapes to describe the molecular structure, ranging from the angstrom-level methods of x-ray crystallography (Shi, 2014) and cryo-EM (Zhang et al., 2019) to the nanometer-level methods like SAS (Hura et al., 2009) and AFM (Pallesen et al., 2009). Biochemical approaches probe the RNA structures by generating the local structural features or restraints (Kubota et al., 2015) in computational modeling from early on (Westhof et al., 1989). Normally, biophysical experiments determine the structure of one RNA each time at the angstrom-level resolution, while biochemical can probe RNA structures in a high-throughput way by sacrificing the resolution. The



well-established biophysical and biochemical approaches for RNA are introduced in the section below.

Biophysics Approaches to Characterize RNA Structures

X-Ray Crystallography and Cryogenic Electron Microscopy (Cryo-EM)

X-ray crystallography and cryogenic electron microscopy (cryo-EM) approaches characterize RNA structures by 3D maps/density maps, which describe the shape of the molecule. In x-ray crystallography, crystallized samples are irradiated by x-ray from different angles to generate a group of diffraction data, which are interpreted into electron density maps after solving the phase problem by an inverse Fourier transform (Cate and Doudna, 2000). The phase problem in crystallography is normally solved by molecular replacement (Evans and McCoy, 2008), isomorphous replacement, anomalous dispersion (McCoy and Read, 2010), or their combination. The quality of a density map depends on the resolution of the structure as well as the thermodynamic mobility of the molecule. It is known that structural regions of high-temperature factors (B factors) may not have a clear electron density to infer the atomic coordinates (Magnus et al., 2020). Thus, computational modeling may optimize the crystal structures using structure knowledge learned from already solved structures (Terayama et al., 2018). Along with the recent advances in detector technology and software algorithms (Doerr, 2017), cryo-EM can be applied on purified macromolecule samples cooled to cryogenic temperatures and embedded in an environment of vitreous water, without the step of crystallization (Liao et al., 2013; Murata and Wolf, 2018; Vinayagam et al., 2018). Each image from cryo-EM shows a view of the macromolecule from a certain angle, while a massive number of particle images are used to reduce the noise and restructure the 3D shape of the molecule at an atomic resolution (Merk et al., 2016). In spite of the expensive equipment and the challenge in working with liquid samples, cryo-EM is now capable of solving a wide range of RNA structures with atomistic models. The determination of SAM-IV riboswitch demonstrates the capability of cryo-EM in solving RNA structures smaller than 40 kDa (Zhang et al., 2019), which is a breakthrough in solving small RNA molecules. Recent advances suggest that the combination of cryo-EM and x-ray crystallography can both solve the phase problem and determine the structure at high resolution (Wang and Wang, 2017). For both x-ray crystallography and cryo-EM, computational models are used to position the atoms into the experimentally obtained electron density.

Small-Angle Scattering (SAS) and Atomic Force Microscopy (AFM)

Small-angle scattering (SAS) and atomic force microscopy (AFM) are low-resolution techniques. SAS, consisting of the small-angle scattering of x-rays (SAXS) and neutrons (SANS), is capable of delivering structural information in the resolution range between 1 and 25 nm. Such information is similar to low-resolution cryo-EM data. SAS samples are dissolved in solution and are

exposed to a beam of x-rays or neutrons. Subsequently, scattering data is collected from the detector. SAXS and SANS differ in their scattering particles: SAXS shows the electron density map while SANS shows the distribution of the nuclei of atoms (Byron and Gilbert, 2000; Koch et al., 2003; Svergun and Koch, 2003), both of which are used to determine the size and shape of particles (Schnablegger and Singh, 2011). Considering the resolution of SAS, it is only possible to capture the global shape of a macromolecule rather than structural details (Skou et al., 2014; Plumridge et al., 2018). In AFM experiments, samples are immobilized on a solid base and raster-scanned by an ultrasharp tip. The tip reacts with the sample in tapping or contact mode, causing the movement of the Piezo element. The movement-resulted laser reflection change can be detected by a photodetector. And the detected signal is then transformed into surface information (Sahin et al., 2007). The 3D surface images produced by AFM can reach a nanometer scale resolution (Shahin and Barrera, 2008), while computation is required to process these images to construct 3D structure models. Possible ways to improve AFM accuracy are better sample fixation and using a sharper tip (Schön, 2018). Successful examples of AFM determined RNA structures have been reported (Husale et al., 2009; Pallesen et al., 2009; Gilmore et al., 2014, 2017), while more details have been reviewed by Schön (2018).

Nuclear Magnetic Resonance (NMR) Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy, which covers ~36% of the pure RNA structure solved in PDB, obtains RNA structural information from the chemical shift of the resonance frequencies of the nuclear spins in the sample. NMR is based on the physical observation that nuclei in a strong constant magnetic field, when perturbed by a weak oscillating magnetic field, produce an electromagnetic signal with a frequency characteristic of the magnetic field at the nucleus. Two-dimensional NMR methods (Chaloner, 1990) are used to detect couplings or connectivities between nuclei that are close to each other in space. Some local structural information can be inferred from the chemical shift spectra. Furthermore, long-range structural information can be probed by residual dipolar couplings caused by the presence of an aligning medium that interferes with the isotropic tumbling of a molecule (Marion, 2013). However, it is necessary to use structural assumptions and computational models to supplement local information when solving large structures. One significant advantage of NMR in solving RNA molecules is its ability in exploring non-canonical geometries, such as non-Watson-Crick base pairs (Hermann and Westhof, 1999), coaxial stacking (Lescoute, 2006), stem-loops, and pseudoknots (Westhof and Jaeger, 1992) which are key in RNA structure modeling. In addition, NMR may probe the ligand, such as proteins or drugs, binding to RNA by seeing which resonances are shifted upon the binding of the ligand (Wemmer, 1996). With the development of NMR technology, solid-state NMR and solution NMR become important tools to solve RNA structure. Solution-state NMR is reported to characterize RNA dynamics with atomic models, but the sample is limited in size (Bothe et al., 2011). Later, solid-state NMR has also been reported to be used in RNA structure determination in high resolution

together with matched experimental methods. Moreover, there is no limitation in size of RNA and crystallization is not necessary (Marchanka et al., 2015). Some successful examples have been reported in both solution NMR (Davis et al., 2004; Latham et al., 2005; Keane et al., 2015; Orlovsky et al., 2020) and solid-state NMR (Marchanka et al., 2015, 2018; Huang et al., 2017; Yang and Wang, 2018).

Fluorescence Resonance Energy Transfer (FRET)

Fluorescence resonance energy transfer (FRET) is used when neither crystal nor solution samples are available. In FRET experiments, the fluorescent donor and the acceptor are attached to the samples, while the evaluated energy transmission by intermolecular dipole–dipole coupling stimulates fluorescent signals. The intensity is related to the distance between the donor and the acceptor (Tuschl et al., 1994; Stephenson et al., 2016). So, FRET is capable of determining long-range contact ranging from 10 to 80 Å (Millar, 1996; Sekar and Periasamy, 2003). Compared to FRET, single-molecule FRET (smFRET) can obtain detailed information of individual molecules, which can be used to understand the folding dynamics and conformation changes (Karunatilaka and Rueda, 2009; Holmstrom and Nesbitt, 2016; Stephenson et al., 2016; Manz et al., 2017). Recently, smFRET was used to probe DNA hairpin changes under different pressures and temperatures, which indicates that we may understand structure changes in different surroundings with smFRET (Sung and Nesbitt, 2020). A brief comparison among the biophysics methods is shown in **Table 1**.

Biochemical Approaches to Probe RNA Structures

Unlike biophysical approaches, biochemical approaches establish biochemical probing as a quantitative measurement of the RNA conformation, in particular, the base pair interactions, the base exposure, and the structural flexibility. Both secondary structure information and tertiary contact information can be obtained through biochemical approaches. Such information can be transformed into structural restraints to direct structure modeling. Experiments related to chemical probing are listed below:

Chemical Probing for Secondary Structure

Chemical probing

Chemical probing, which dates back to the 1980s (Peattie and Herr, 1981), detects structural information by introducing chemical modifications and changes to RNA using chemicals (**Figure 1B**) (Weeks, 2010; Kubota et al., 2015), whose reactivity depends on local RNA structure. These chemical reagents include dimethyl sulfate (DMS) (Tijerina et al., 2007; Homan et al., 2014a), silyl derivative *N,N*-(dimethylamino) dimethyl chlorosilane (DMAS-Cl) (Mortimer et al., 2009), carbodiimides (CMCT), kethoxal (Ehresmann et al., 1987), and glyoxal and its derivatives (Mitchell et al., 2018). Different reagents can react with specific sites or structures to form covalent adducts (Busan et al., 2019) at the modification sites, while unpaired nucleotides are more exposed and more inclined to be modified.

After the reaction between the probing reagents and the structured RNA, two methods can detect the modification of the RNA. For the first method, RNA is labeled before the modification can be further treated after the modification to form a strand scission, which can be directly detected by electrophoresis (Ehresmann et al., 1987). In another way, the RNA is reverse transcribed using a reverse transcriptase into a DNA copy. Reverse transcription with stop (RT-stop) or mutation (RT-mutate) is a common method to detect modifications. RT-stop (**Figure 1B**) (Brunel and Romby, 2000; Merino et al., 2005) is based on the reverse transcription interruption by the adducts, while RT-mutate (**Figure 1B**) is based on the fact that reverse transcription introduces a mismatched nucleotide at the position of the adduct under special conditions (Siegfried et al., 2014). RT-stop results in a pool of DNA truncations of different lengths, whose frequencies reflect the RNA structure profile and can be assayed on a gel, while RT-mutate results in a pool of cDNAs of different sequences which need to be profiled by high-throughput sequencing (HTS). Besides local structure blocking, nucleotide positions can also be protected by a binding protein (Smola et al., 2015a).

Several other recently proposed chemical probing techniques are introduced below. (1) Using bifunctional reagents to probe long-range contacts (Weeks, 2010); (2) Using mutational profiling (MaP) to reveal the dynamic states and reactive sites of an RNA (Homan et al., 2014a; Krokhotin et al., 2017); (3) Using 2'-Hydroxyl molecular interference (HMX) to identify tertiary interactions in the highly packed regions (Homan et al., 2014b); and (4) Using hydroxyl radical probing (HRP) to study RNA folding and direct structure refinement (Ding et al., 2012; Costa and Monachello, 2014). Simultaneously, computational methods are being developed to cooperate with the experimental developments, e.g., discrete molecular dynamics simulations (DMD) were used to produce 3D models using HMX data (Homan et al., 2014b).

In-line probing and enzymatic probing

In-line probing and enzymatic probing have similar principles to chemical probing but differ in their reactions. Enzymatic probing (Wan et al., 2013) is based on the fact that RNase enzymes, which are local structure-specific enzymes, cut different regions of an RNA molecule. Thus, the RNA structure can be probed using the enzyme cut truncations. In-line probing utilizes the feature that the 2'-hydroxyl reacts with the backbone phosphate group in some conditions (the 2'-hydroxyl and the phosphate group form an angle of 180 degrees, i.e., a 'line' structure required) to break the backbone. As flexible RNA nucleotides are more inclined to fulfill such a condition, in-line probing tests the local structure flexibility (Regulski and Breaker, 2008).

Chemical Probing for Secondary and Tertiary Information

Many chemical probing experiments are used to measure secondary structure information, while some reagents and methods may infer tertiary structure information.

TABLE 1 | A comparison of biophysics approaches.

Biophysics methods	Advantages	Limitations
X-ray crystallography	Atomic resolution structure (Shi, 2014)	Difficulty in crystal preparation; Phasing problem
Cryo-EM	Near atomic resolution structure (Murata and Wolf, 2018); Capable of solving macromolecule structures; Shows a cluster of conformations	Expensive equipment; Difficulty in dealing with liquid sample
SAS	Easy to perform experiment; Capable of dealing with large molecules; Time efficient	Nanometer-level resolution (Svergun and Koch, 2003; Rambo and Tainer, 2013); Sample aggregation
AFM	Easy to perform experiment; Surface information	Nanometer- to molecular resolution result (Schön, 2018)
NMR	Sample in solution condition; Possible to probe the structural dynamics or invisible states; Atomistic structure models (Bothe et al., 2011)	Sample aggregation; Limited size of the structure
FRET	Easy to perform experiment; Angstrom-level resolution (Sekar and Periasamy, 2003)	Can get limited structure information; Limited to small RNA structures; The signal-to-noise ratio

Experiments using bifunctional reagents (Weeks, 2010) and M2-seq (Cheng et al., 2017) were reported to reveal long-range contacts (Weeks, 2010). Moreover, HMX (Homan et al., 2014b), MOHCA (Das et al., 2008), RING-MaP (Krokhotin et al., 2017) experiments, and comparing SHAPE results achieved by different reagents (Steen et al., 2012) were also used to measure proximal tertiary interactions. Those examples show great potential in probing tertiary structure.

A widely used and further explored chemical probing method, selective hydroxyl acylation analyzed by primer extension (SHAPE) (Wilkinson et al., 2006; Spitale et al., 2013; Siegfried et al., 2014; McGinnis et al., 2015; Smola et al., 2015b; Lee et al., 2017; Mustoe et al., 2018; Smola and Weeks, 2018; Busan et al., 2019), is based on the acrylate reaction between electrophilic reagents and active RNA 2'-hydroxyl groups which always appear in the single-strand region (Wilkinson et al., 2006). It reports on RNA structure at a single nucleotide resolution, thus is used to generate highly accurate secondary structure models (Deigan et al., 2009). It can be combined with high-throughput sequencing by barcode based multiplexing (Lucks et al., 2011) and can use a combination of chemical probing reagents and experimental data (Kladwang et al., 2011b). SHAPE has been used to analyze large RNA structures, including the SARS-CoV-2 genome (Manfredonia et al., 2020) and the HIV-1 genome (Watts et al., 2009).

Recent advances in chemical probing lie in the enlarged throughput (Kwok et al., 2015) (probing many RNAs simultaneously) and the *in vivo* probing of RNA molecules in their cellular environment (Kubota et al., 2015; Nguyen et al., 2016). As for modification detection methods, some studies show that combining RT-mutate and RT-stop can mitigate bias and have a better insight into the structure since they can provide complementary information (Novoa et al., 2017; Sexton et al., 2017). A more recent study of mutational profiling (MaP) shows the advantages of RT-mutate in several aspects: RT-mutate is simpler and faster (Busan et al., 2019), while it works on long and

complex RNAs (Siegfried et al., 2014; Smola et al., 2015b; Busan and Weeks, 2018). Additionally, RT-mutate provides relative adduct frequencies inferred from read-depth, which measures the reactivities of different segments in the same experiment (Mustoe et al., 2018).

Progress has also been made in reagents exploration and the assessments of the SHAPE method (Lee et al., 2017; Busan et al., 2019). Five reagents 1M7 (1-methyl-7-nitroisatoic anhydride), 1M6 (1-methyl-6-nitroisatoic anhydride), NMIA (*N*-methylisatoic anhydride), NAI (2-methylnicotinic acid imidazolide), and 5NIA (5-nitroisatoic anhydride) have been assessed and recommended to use in experiments of different conditions [94]. Moreover, to explore the correlation between RNA structure and its SHAPE result, a computational model named 3D structure-SHAPE relationship (3DSSR) was developed to generate the SHAPE profile for a given RNA structure. This model can also indicate the inconsistency between the structure model and the SHAPE data and thus exclude the unreasonable structure models [152].

Chemical Probing Coupled With Mutagenesis

Chemical probing (or also known as chemical mapping) and related methods generally probe the features, e.g., the pair/unpair state, of a single nucleotide. To understand the base-pair connection between nucleotides, chemical mapping is coupled with mutagenesis. If a base-pair forming nucleotide is mutated, its partner may become unpaired, and thus tends to become more exposed and more detectable by chemical mapping (Kladwang and Das, 2010; Kladwang et al., 2011a; Cordero and Das, 2015). Early examples have been shown from the group I intron (Duncan and Weeks, 2008) and tetrahymena ribozyme (Pyle et al., 1992) studies. Of course, some mutations may lead to significant perturbations to the structure, even the unfolding of an entire helix. In general, the method of coupling chemical mapping and mutagenesis, known as mutate-and-map or M^2 , achieves a highly accurate detection of the canonical base pairs (Kladwang et al., 2011a;

Cheng et al., 2017), with ~2% error rate in the helix region (Kladwang et al., 2011a).

M2-seq (Cheng et al., 2017), which is an advanced version of mutate-and-map (Figure 2), uses error-prone PCR to introduce mutations and couple DMS modification with HTS (high-throughput sequencing) based on the RT-mutate mechanism. M2-net is the algorithm to normalize M2-seq data and get secondary structure information (Cheng et al., 2017). MOHCA (Das et al., 2008) is an experiment that can map the positions of nucleotides that are close together in the three-dimensional structure. Highly reactive chemicals [2'-NH₂-2'-dATP and isothio cyanobenzyl-Fe(III)EDTA] are attached to the nucleotides. And the Fenton reaction damages other nucleotides nearby the reacting Fe-modified nucleotide. It is possible to map the positions of these reacting nucleotides by detecting the damaged nucleotides. However, finding the damaged nucleotides is tedious and requires specialized equipment. MOHCA-seq (Cheng et al., 2015) is a new development of MOHCA (Figure 2), which couples MOHCA with HTS based on RT-stop. MAPseeker is the computational program used to analyze MOHCA-seq data and infer proximal tertiary interactions. Both the base-pair information from M2-seq and the tertiary interaction information from MOHCA-seq are key restraints for RNA structure modeling.

Besides *in vitro* structure probing, structure ensembles, and structure dynamics in living cells (Ganser et al., 2019) *in vivo* probing is being rapidly developed. Apart from *in vitro* structure probing, methods like DMS, SHAPE, and HRP can also be applied *in vivo* (Climie and Friesen, 1988; Moazed et al., 1988; Kwok et al., 2013; Spitale et al., 2013, 2015; Tyrrell et al., 2013; Ding et al., 2014; Smola et al., 2015a, 2016; Watters et al., 2016a; Lee et al., 2017; Feng et al., 2018; Mitchell et al., 2018). DMS demonstrates the advantages of its small size and short reaction time to capture some transient changes in structure (Waldsich, 2002; Waldsich et al., 2002; Tijerina et al., 2007). Reagent such as 1M7 (Tyrrell et al., 2013; McGinnis and Weeks, 2014; McGinnis et al., 2015) FAI, and NAI (Spitale et al., 2013) are developed to be used for *in vivo* SHAPE probing, while NAI shows the potential to reveal the structural differences between *in vivo* and *in vitro* probing (Kwok et al., 2013; Hector et al., 2014). Complementarily, HRP can probe RNA dynamics, folding (Latham and Cech, 1989) and solvent accessibility (Latham and Cech, 1989) in living cells (Kubota et al., 2015). More detailed information about *in vivo* probing and its challenges can be obtained from the review from Kubota et al. (2015). A recent study presented by Tomezsko et al. (2020) integrates DMS-MaPseq data with the DREEM (detection of RNA folding ensembles using expectation-maximization) algorithm to identify possible *in vivo* structures, suggesting a systematic method for dynamic detection and captures transient conformations *in vivo*. Another *in vivo* RNA-RNA interaction and RNA structure detection method, MARIO, utilizes an RNA linker to identify RNA-RNA interactome and finds secondary single-stranded regions as well as tertiary proximal contacts (Nguyen et al., 2016).

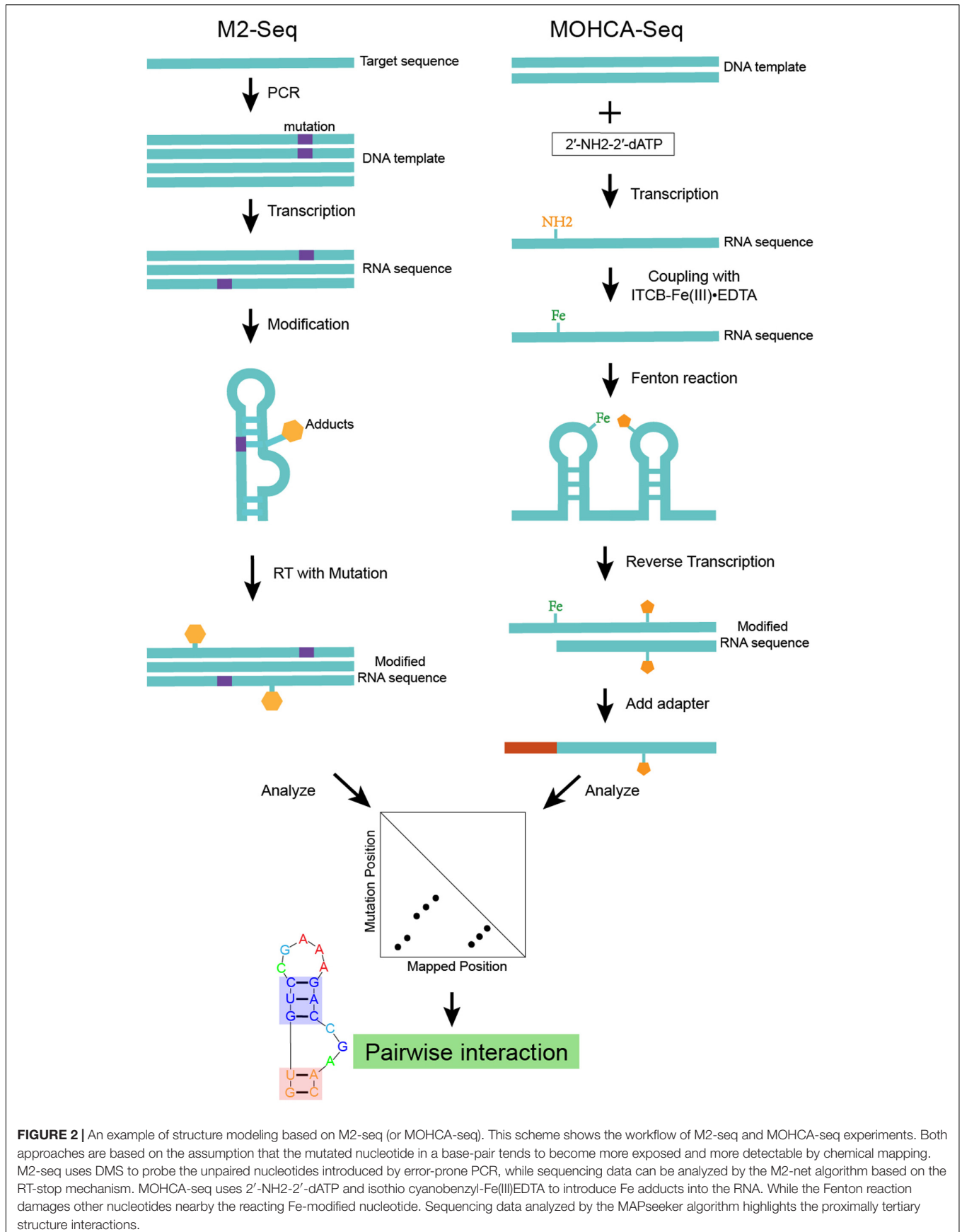
EXPERIMENTAL DATA-DRIVEN COMPUTATIONAL MODELING OF RNA 3D STRUCTURES

Types of Data-Driven Restraints

Most experimental data alone, except x-ray crystallography and high-resolution cryo-EM, is not enough to infer atomic-resolution RNA structure. Hence computational modeling is indispensable in interpreting the experimental data into an atomic description of RNA 3D structure. Such computational modeling can be based on existing structure determination software, such as Phenix (Liebschner et al., 2019) for x-ray crystallography and cryo-EM, or CANDID (Herrmann et al., 2002) for NMR, or based on structure prediction programs, e.g., Rosetta (Das and Baker, 2007), Assemble2 (Jossinet et al., 2010; Jossinet, 2015), or MC-sym (Parisien and Major, 2008). The type of data used in structure modeling can be classified into five types:

(1) Sequence and sequence alignments are also determined from sequencing experiments. Sequence alignments can indicate certain co-evolution between base pairs. With the human genome project (Bentley, 2000) and advances in sequencing technologies (Schuster, 2008), a massive number of metagenome sequences are being deposited in databases (Sayers et al., 2011). Sequence covariation gives an alternative to understanding base-pair interactions and even RNA-protein interactions (Weinreb et al., 2016). The secondary structures of many significant RNA structures, such as the 16S ribosomal RNA structure (Noller and Woese, 1981), were derived from sequence covariations. Besides, sequence covariation also predicts some long-range Watson-Crick base-pairs in pseudoknots: a successful case was illustrated in the group I intron structure (Lehnert et al., 1996). DIRECT is a new algorithm reported to improve the prediction in long-range contacts and captures more tertiary structural information (Jian et al., 2019). A recent study demonstrated some success in RNA secondary structure prediction through the integration of sequence alignment based on direct coupling analysis and minimum free energy (MFE) (He et al., 2020). A machine learning approach was also reported to achieve better secondary structure prediction through the integration of direct coupling analysis and SHAPE data (Calonaci et al., 2020). However, some functional long non-coding RNAs may not necessarily have structures and may not have many homologous sequences or some sequence alignments could be too conserved to infer covariations. This severely limits the prevalence of these covariation-based RNA structure prediction methods. Details of covariation bases methods have been discussed by Rivas et al. (2017, 2020), while the resulting method has been used to curate Rfam RNA sequence families.

(2) The 3D shape of the molecule ranges from high-resolution x-ray crystallography to SAS, which only describes the global topology of a structure. In the case of x-ray crystallography and high-resolution cryo-EM, atomic coordinates may directly fit into the density maps. As for low-resolution



density maps, computational modeling is required to generate structural models to fit in the global topology, while additional information (possibly the other three types of information, elaborated in the integrative modeling section below) may help the modeling.

(3) Features of a single nucleotide, such as paired/unpaired state and buried/exposed state, can be inferred from chemical probing data. These features are normally used to validate the secondary structure predicted by minimum free energy (MFE), e.g., using SHAPE to probe RNA secondary structures or to test probing profile change upon protein/ligand binding. Some of these HTS-based experiments, including DMS-seq (Rouskin et al., 2014) and SHAPE-seq (Loughrey et al., 2014; Watters et al., 2016b), have been commercialized. Together with these experiments, computational methods have been developed to derive more accurate structural information. An RNAsc method is proposed according to SHAPE data, which includes pseudo-energy terms and base stacking positions, making it more accurate to predict the secondary structure (Zarringhalam et al., 2012). Another method also utilizes pseudo-energy information obtained from chemical probing and integrates with thermodynamic folding algorithms to reach a good secondary structure prediction (Washietl et al., 2012). The Kalman approach is used to filter SHAPE data from noise (Vaziri et al., 2018) and the RING-MaP method reveals diverse interactions at both secondary and tertiary levels (Krokhotin et al., 2017). FoldAtlas was developed for high-throughput chemical probing data processing (Norris et al., 2017) and a statistical modeling was developed to improve the sensitivity of high-throughput probing data (Selega et al., 2017). An SNPfold algorithm was designed to recognize SNP-induced conformational changes in genome-wide analysis (Halvorsen et al., 2010). The above mentioned DREEM algorithm can be used to characterize different conformations from DMS-MaPseq data (Tomezsko et al., 2020). These examples show that efforts made in data processing modeling and methods may give us more useful and accurate information.

(4) Base pair interaction cannot only be inferred from sequence covariations (De Leonardis et al., 2015; Weinreb et al., 2016) but can also be determined from biochemical experiments (e.g., M2-seq) or secondary structure prediction (Zuker, 2003). Considering the base-pairing nature of RNA structures, this type of information has particular importance in determining the structure. In terms of secondary structure prediction, the MFE (minimum of free energy) structure simplifies the RNA structure as Watson–Crick base-pair interactions. Thus, it can be predicted by the combination of a loop-based energy model and the dynamic programming algorithm introduced by Zuker (2003). However, it ignores the contribution of pseudoknots and non-Watson–Crick interactions. When covariation or experimental evidence is available, the Watson–Crick base pairs in pseudoknots can still be correctly identified. It is most difficult to determine non-Watson–Crick base-pairs. The best way besides x-ray

crystallography and cryo-EM is to use a combination of chemical probing experiments and sequence covariation. And some successful examples have been reported (Walczak et al., 1996). As discussed above, MOHCA-seq has demonstrated its ability in determining both Watson–Crick and non-Watson–Crick long-range base-pairs. In 3D structure modeling, the base-pair information is often used together with the features of a single nucleotide. For example, M2-seq is coupled with DMS to probe RNA structure and has been successful (Cheng et al., 2017).

(5) Non-canonical interaction-based RNA modules are RNA structural modules that are formed by non-canonical interactions and can be predicted by module prediction methods (Cruz and Westhof, 2011; Zirbel et al., 2015). Some recent determined RNA structures have demonstrated the importance of these non-canonical interactions (Butcher and Pyle, 2011): the recently crystallized MALAT1_th11 RNA (Ruszkowska et al., 2020) shows a UA-U-rich RNA triple helix with 11 consecutive base triples (**Figure 3A**); the two *trans*-sugar-Hoogsteen G:A base-pairs in the kink-turn module (Huang et al., 2019a) enables its folding in 3D (**Figure 3B**); and the triple interactions in the pseudoknot structure of the glutamine-II riboswitch (Huang et al., 2019b) is known to be crucial for ligand binding (**Figure 3C**). A structural module in RNA is a set of ordered *non*-Watson–Crick base-pairs embedded between Watson–Crick pairs, which are recurrent in the RNA structure (**Figure 4B**). With a good sequence alignment, such modules can often be predicted because of their restricted sequence variations. Automatic prediction programs based on this pragmatic approach have been developed (Cruz and Westhof, 2011; Theis et al., 2013). Without *non*-Watson–Crick interactions, the structural description is still at the secondary structure level. According to RNA-Puzzles, the prediction accuracy of *non*-Watson–Crick base-pairs is less accurate than the prediction of Watson–Crick base-pairs (**Figure 4A**), hindering the accurate RNA structure prediction. RNA structural modules are important for the folding of 3D RNA structures as well as for their functionality. For example, the kink-turn module, known to bind to the L7Ae protein (Turner et al., 2005), encodes a certain signature in its sequence (**Figure 4C**) and folds into a fixed structural fold (**Figure 4B**). Such a signature can be summarized by a probabilistic model (Cruz and Westhof, 2011; Sarrazin-Gendron et al., 2019), such as the *Basepairing* program (**Figure 4D**), or by machine learning (Zirbel et al., 2015).

Types of Structure Prediction Approaches

All of these five types of structural information complement each other. Therefore, the integration of several types of the available information may improve the prediction of RNA structures. When neither a crystalline structure nor high-resolution cryo-EM data is available, the modeling of RNA structure in 3D may rely on RNA 3D structure prediction, which uses some existing knowledge about RNA structure derived from

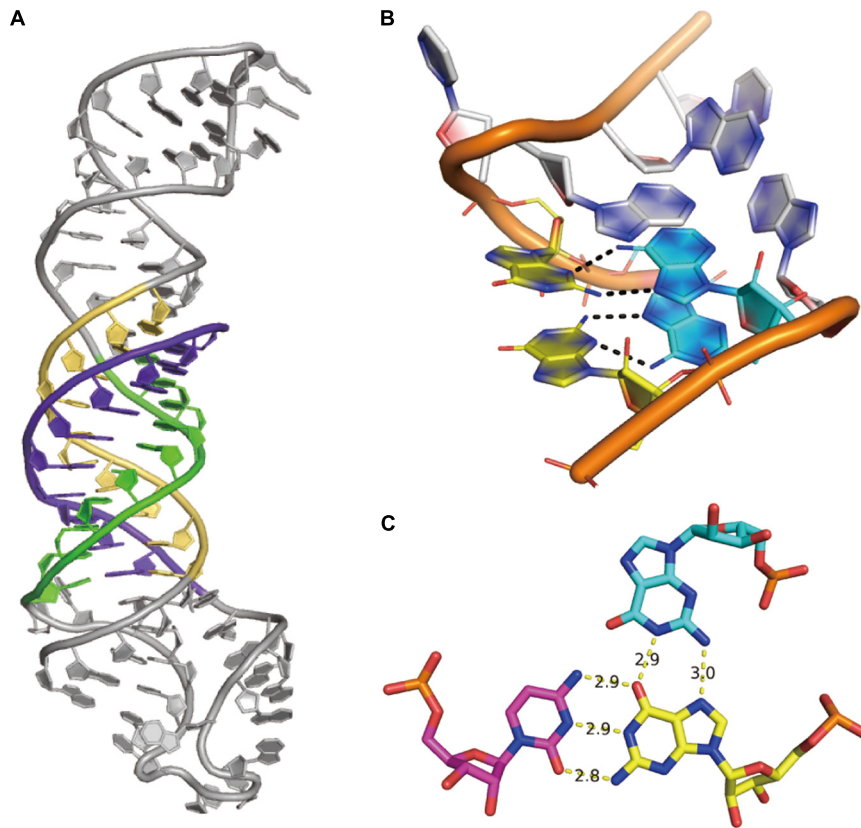


FIGURE 3 | Examples of non-canonical interactions. **(A)** The UA-U-rich RNA triple helix in MALAT1_th11 RNA, PDB id 6SVS. **(B)** The simple k-turn with two *trans*-sugar-Hoogsteen G:A base pairs, PDB id 6HCT. **(C)** The structure of the G18 > G2:C39 triple interaction in glutamine-III riboswitch, PDB id 6QN3.

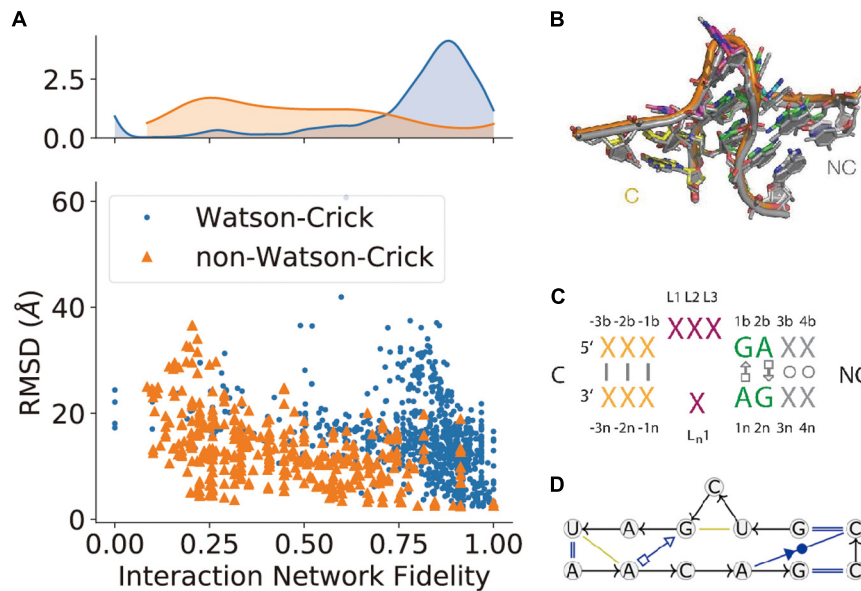


FIGURE 4 | The prediction of non-canonical base-pairs. **(A)** The comparison between Watson-Crick and non-Watson-Crick base-pairs prediction in terms of interaction network fidelity (INF; Parisien et al., 2009) in RNA-Puzzles. **(B)** The superimposition of recurrent kink-turn modules. **(C)** The graphical abstraction of the kink-turn module. **(D)** The module abstraction in the *Basepairing* program.

the existing databases. RNA 3D structure prediction has used some computational approaches, such as comparative modeling, fragment assembly, and *de novo* modeling (Figure 5) details can be referred to in the review paper by Rother et al. (2011a).

Comparative modeling is based on the assumption that homologous RNA molecules present similar structures in the relatively conservative regions. An RNA can be predicted using the known homologous structure in the database as a template. A prediction may either use the pairwise sequence alignment to model based on the template structure (Rother et al., 2011b) (ModeRNA) or translate the template structure into structural restraints in prediction (Flores et al., 2010) (RNABuilder). However, comparative modeling highly depends on the availability of a homologous template, thus it cannot be applied to explore a novel RNA structure.

Fragment assembly first decomposes the known RNA structures into fragments and makes a fragment library, and predicts an RNA structure by searching fragments that are similar to the target sequence and assembling them together. Several approaches in this type use secondary structure as a predefined restraint [e.g., RNAComposer (Popenda et al., 2012), 3dRNA (Zhao et al., 2012), and VfoldLA (Xu et al., 2019)], while experimental data may help in the determination of the secondary structure.

de novo modeling is a collective name for those predictions without templates. *de novo* modeling is normally based on force-field simulation and searches for the ideal structure by sampling the conformational space [e.g., NAST (Jonikas et al., 2009), iFoldRNA (Sharma et al., 2008) (Krokhotin et al., 2015), and SimRNA (Boniecki et al., 2016)]. As the whole conformational space can be too large to explore, using predefined restraints such as the secondary structure to narrow down the search space may effectively reduce the unnecessary search and achieve a better prediction.

Integrative Modeling

When no homologous template is available, fragment assembly and *de novo* modeling may integrate various types of information to help with the prediction. A widely-used approach is to predict the secondary structure first and predict the 3D structure according to the secondary structure. One simple case is the RNAComposer: one can use RNAfold from ViennaRNA (Lorenz et al., 2011) to predict the secondary structure and use RNAComposer to assemble the 3D modeling according to the predicted secondary structure using the fragment library RNA FRABASE (Popenda et al., 2010). A more complicated version could use DMS-seq and M2-seq to constrain the secondary structure, while use MOHCA-seq to probe the tertiary interactions. Then, use fragment assembly to build an initial 3D model and optimize it through conformational search using Rosetta or other force-field based methods. One may even use covariations from the multiple sequence alignment to confirm the base-pairs/interactions, and use module prediction methods to assign the non-canonical base-pairs. Such modeling integrates several types of information (features of a single nucleotide, base-pair interaction, and non-canonical interactions) in several prediction methods (fragment assembly

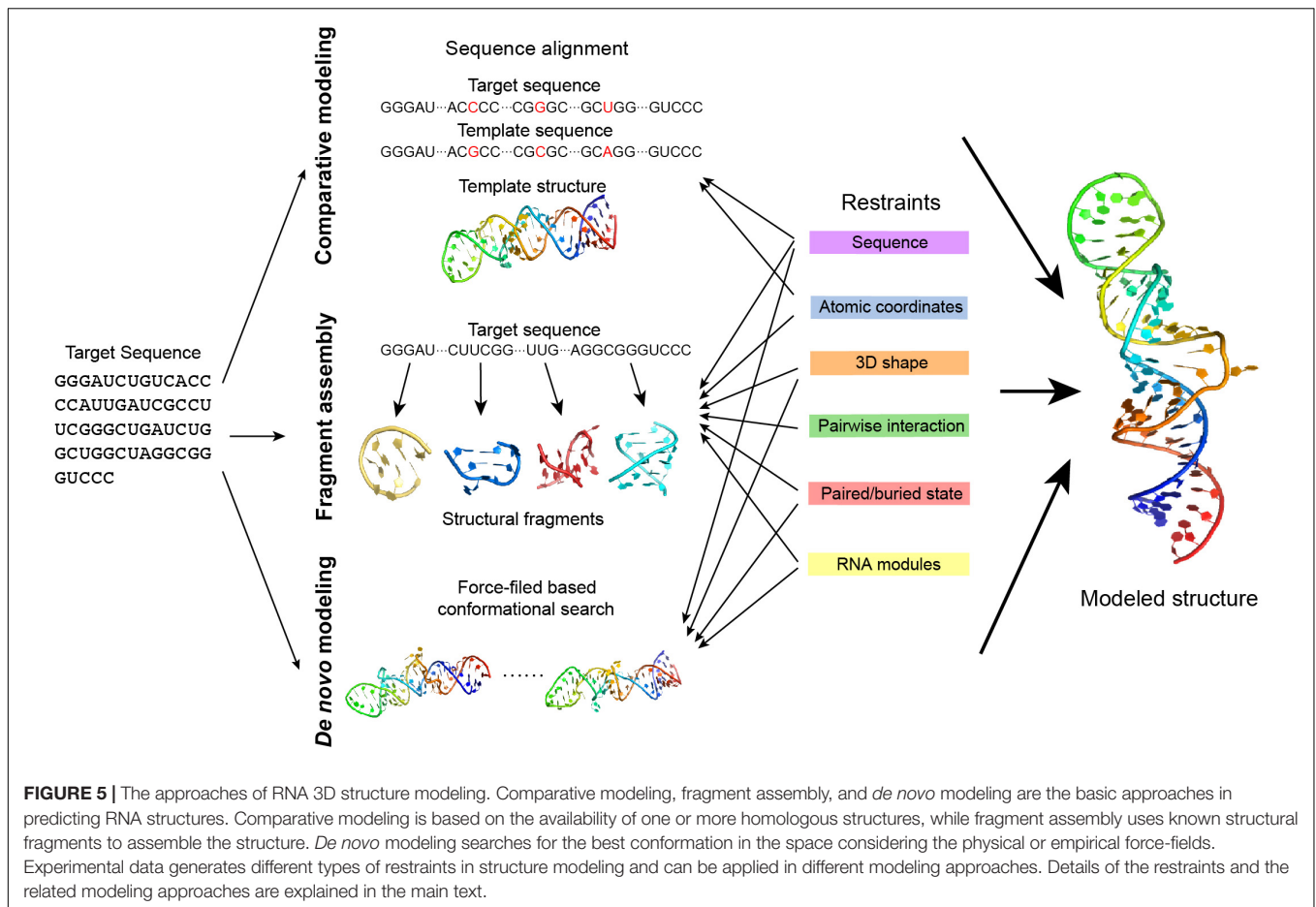
and *de novo* modeling). Normally, integrative modeling is defined as modeling that uses more than one modeling approach. To extend, it may also integrate different types of information in different types of prediction methods to achieve a high-quality prediction. As experiments, especially the HTS based techniques, are becoming better established and cheaper, integrative modeling using experimental data is becoming easier and makes modeling more reliable.

Various integrative approaches have been reported in recent years: A fully computational workflow, EvoClustRNA (Magnus et al., 2019), effectively integrates *de novo* modeling and fragment assembly and statistical potential to achieve accurate predictions. EvoClustRNA first selects a set of sequences homologous to the target sequence to be modeled and uses both FARFAR (fragment assembly) and SimRNA (*de novo* modeling) to model all the sequences. The conserved regions of all these sequences are extracted and clustered. And the predicted structure is selected from the most commonly preserved structural arrangements of the homologous structures based on a statistical potential. Although EvoClustRNA is a fully computational workflow, it demonstrates good predictions in several targets of RNA-Puzzles. The integrative modeling platform (Russel et al., 2012) is a computational platform to integrate SAS, EM, x-ray crystallography, or NMR data by comparative modeling. It was designed for protein structures but may also extend to other macromolecules like RNA. Another example is PLUMED-ISDB, which is based on the molecular dynamics library PLUMED (Tribello et al., 2014). It may integrate experimental data from NMR, FRET, SAXS, or cryo-EM to model the structure and dynamics of RNA as well as other biomacromolecules (Bonomi and Camilloni, 2017). Apart from the above mentioned approaches, there are also other computational modeling methods, such as molecular dynamics (MD) and quantum mechanics (QM). Discussion about those methods have been introduced in other reviews (Dawson and Bujnicki, 2016; Miao and Westhof, 2017; Schlick and Pyle, 2017).

However, not all experimental data are accurate and informative enough. Integrating noisy or bad data into structure modeling may result in a worse prediction. A recent study (Wang et al., 2019) performing structure prediction with excessive restraints from experimental data demonstrated this conclusion. Therefore, more strict assessment and calibration of the experimental data for modeling is needed and the related computational methods need to be developed.

RECENT SUCCESSES IN COMPUTATION-AIDED RNA STRUCTURE DETERMINATION

Massive progress is being made through the integration of different methods, both computational and experimental. Not only experimental data can be used as restraints to help computational modeling, but also computational approaches benefit experimental research.

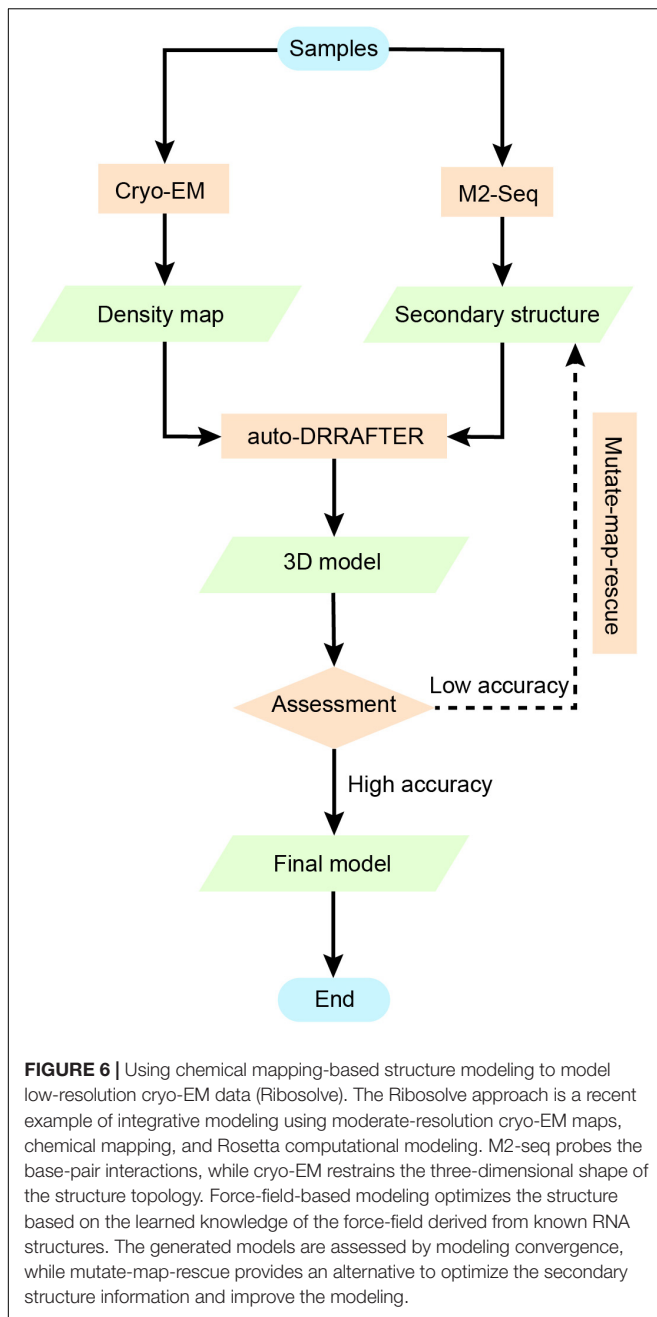


One such advance was to use the *de novo* predicted structure to determine the x-ray crystal phase (Huang et al., 2019b). Huang et al. (2019b) reported *de novo* predictions, which are within 3 Å root mean square deviation (RMSD) from the crystal structure of glutamine-II riboswitch. And two of the predicted models were able to achieve the molecular replacement to determine the phase of the crystal. This work reports a useful potential that computational modeling may take up a more important role in interpreting experimental data.

With the rapid development of cryo-EM technology, computational studies are accentuated to help in structure determination. Deep learning models have been proposed to address the laborious particle picking problem (Norouzi et al., 2013; Wang et al., 2016; Al-Azzawi et al., 2019; Yao et al., 2019). For high-resolution cryo-EM density maps, a fully automatic computational method, *map_to_model* (Terwilliger et al., 2018) from Phenix (Liebschner et al., 2019), is capable of yielding high-accuracy initial models. However, it is still not easy to deal with low-resolution data. And more efforts are being made. DRRAFTER (Kappel et al., 2018) allows for the tracing of RNA atomic coordinates in the biologically important but low-resolution regions using *de novo* computational modeling. cryo-EM maps offer the information of the 3D shape of a molecule, while

computational modeling infers the atomic coordinates from this information.

However, as DRRAFTER was developed specifically for modeling large RNA-protein complexes, which require initial manual setup, it may not be effective and may include bias when modeling smaller RNAs without protein partners. auto-DRRAFTER was designed to address this problem in an automatic way. auto-DRRAFTER constitutes the computational part of Ribosolve (Kappel et al., 2020), which is a hybrid workflow integrating moderate-resolution cryo-EM maps, chemical mapping, and Rosetta (Das and Baker, 2007) computational modeling. As shown in Figure 6, Ribosolve first uses native gel to check for the formation of sharp bands in the RNA samples. RNA secondary structures can be determined by M2-seq, while the 3D shape of the RNA global architecture can be obtained from cryo-EM. In addition, mutate-map-rescue experiments can confirm or refine novel secondary structure rearrangements. Based on these two types of information, auto-DRRAFTER is able to build all-atom models. Model accuracy is predicted from the overall modeling convergence, while uncertain regions are identified by comparing the per-residue-convergence and real-space correlation between the map and model. As a result, Ribosolve dramatically captures the structure of full-length *tetrahymena* ribozyme and builds 13 RNA structures without



protein binding ranging from 119 to 338 nt. Moreover, the structure of several other important RNA, like hc16 ligase and full-length *Vibrio cholerae* and *Fusobacterium nucleatum* glycine riboswitch aptamers, are determined. The SAM-IV riboswitch structure illustrates Ribosolve's ability in solving small RNA structures within 40 kDa (Zhang et al., 2019). These results from Ribosolve demonstrate a rapid and routine determination of RNA-only 3D structures.

There are also other successful examples in both computational-aided structure determination and in new computational modeling methods. Zhang et al. (2018) integrated NMR and cryo-EM data using MD simulation to solve the

30 kDa HIV-1 RNA. NMR data and SAXS data were used jointly to determine the HIV-1 intron splicing silencer structure (Jain et al., 2016). Ab initio 3D structure modeling helped in determining the HIV-1 Rev response element using SAXS data (Fang et al., 2013). The RS3D method can utilize secondary structure information, SAXS data, and any tertiary contacts information to determine the RNA structure topology (Bhandari et al., 2017). FARFAR (Das et al., 2010; Yesselman and Das, 2016) and FARFAR2 (Watkins et al., 2020) were developed to model non-canonical RNA structure at near-atomic accuracy. Swellix explores RNA conformational space when integrating data from crystallography, cryo-EM, or *in vivo* crosslinking and chemical probing methods (Sloat et al., 2017). Statistical modeling helped to construct RNA structure landscapes from structure profiling data, which can facilitate the studies of RNA dynamics and function (Li and Aviran, 2018).

CHALLENGES AND PERSPECTIVES

In spite of the good number of RNA structures being determined, it is still urgent to probe the structural basis of RNA functions. For example, with the break out of virus infection [2015–2016 Zika virus (Akiyama et al., 2016), 2019 coronavirus (Huang et al., 2020)], rapid determination of the structures of functional elements in virus RNA may provide invaluable insights into human health. Computational modeling is playing a more critical role in biological studies (Markowitz, 2017), while RNA structure determination is a particular field where computational modeling can exert its potential. Although RNA structure determination is now becoming rapid and routine, there exist several major challenges.

A major challenge is to improve experimental accuracy for high-throughput biochemical experiments. For this purpose, it is necessary to systematically benchmark these high-throughput techniques with high-resolution RNA structures. However, when high-resolution RNA structures are not available, such benchmarking of experimental accuracy is indirect and confounded with computational modeling. Thanks to the increasing throughput of experimental data being generated, a more realistic evaluation of experimental noise will become possible. Considering that the high-throughput sequencing data tends to reflect a mixture of the structural conformations, computational modeling methods are expected to deconvolve this structural information and dissect the conformations. High-throughput experiments will also help in determining more unknown structures in Rfam families.

Another significant challenge lies in probing RNA structures *in vivo* to understand the structural dynamics. Although DMS, SHAPE, HRP as well as other recent methods, such as DMS-MaPseq, provide good tools for structure ensembles and dynamics detection, the folding process to the functional state remains unsolved. Compared to biochemical experiments, x-ray crystallography and cryo-EM normally need to prepare RNA samples in extreme conditions which are relatively difficult. And high resolution methods with easier sample preparation may

be developed in the future (Stagno et al., 2017). Computational modeling is the interface between experimental data and the RNA structure coordinates. It is always a central problem to improve computational modeling according to the experiments and also help to correct experimental biases.

In terms of structure modeling, it is known that the prediction of non-canonical interactions (non-Watson–Crick base-pairs) is still difficult. A good description of these non-canonical interactions both benefit the RNA structure modeling and the comprehension of functional RNA modules. To improve the RNA structure prediction methods, optimizing non-canonical interaction is still inevitable. As non-Watson–Crick base-pairs could include additional chemical probing information, e.g., the N7G methylation by DMS, new computational methods dealing with these different chemical probing datasets may effectively improve the determination of non-Watson–Crick interactions and achieve better structure predictions. Automatic modeling workflow is another challenge in RNA modeling. Most currently available automatic web servers depend on the sole input of the RNA sequence and optionally the secondary structure information. And the structure modeling accuracy is not ideal because of the lack of experimental data. However, it will be key to transform various experiment data into structural restraints in a standard manner. Automatic web servers should also promote the input of experimental data as restraints to improve prediction accuracy.

REFERENCES

- Akiyama, B. M., Laurence, H. M., Massey, A. R., Costantino, D. A., Xie, X., Yang, Y., et al. (2016). Zika virus produces noncoding RNAs using a multi-pseudoknot structure that confounds a cellular exonuclease. *Science* 354, 1148–1152. doi: 10.1126/science.aah3963
- Al-Azzawi, A., Ouadou, A., Tanner, J. J., and Cheng, J. (2019). AutoCryoPicker: an unsupervised learning approach for fully automated single particle picking in Cryo-EM images. *BMC Bioinform.* 20:326. doi: 10.1186/s12859-019-2926-y
- Al-Hashimi, H. M., and Walter, N. G. (2008). RNA dynamics: it is about time. *Curr. Opin. Struct. Biol.* 18, 321–329. doi: 10.1016/j.sbi.2008.04.004
- Bentley, D. R. (2000). The human genome project—an overview. *Med. Res. Rev.* 20, 189–196. doi: 10.1002/(sici)1098-1128(200005)20:3<189::aid-med2>3.0.co;2-#
- Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., et al. (1992). The nucleic acid database. A comprehensive relational database of three-dimensional structures of nucleic acids. *Biophys. J.* 63, 751–759. doi: 10.1016/s0006-3495(92)81649-1
- 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.
- Bhandari, Y. R., Fan, L., Fang, X., Zaki, G. F., Stahlberg, E. A., Jiang, W., et al. (2017). Topological Structure determination of RNA using small-angle X-Ray scattering. *J. Mol. Biol.* 429, 3635–3649.
- Bida, J. P., and Maher, L. J. III (2012). Improved prediction of RNA tertiary structure with insights into native state dynamics. *RNA* 18, 385–393. doi: 10.1261/rna.027201.111
- Bonicki, M. J., Lach, G., Dawson, W. K., Tomala, K., Lukasz, P., Soltysinski, T., et al. (2016). SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction. *Nucleic Acids Res.* 44:e63. doi: 10.1093/nar/gkv1479
- Bonomi, M., and Camilloni, C. (2017). Integrative structural and dynamical biology with PLUMED-ISDB. *Bioinformatics* 33, 3999–4000. doi: 10.1093/bioinformatics/btx529
- Bothe, J. R., Nikolova, E. N., Eichhorn, C. D., Chugh, J., Hansen, A. L., and Al-Hashimi, H. M. (2011). Characterizing RNA dynamics at atomic resolution using solution-state NMR spectroscopy. *Nat. Methods* 8, 919–931. doi: 10.1038/nmeth.1735
- Brunel, C., and Romby, P. (2000). Probing RNA structure and RNA-ligand complexes with chemical probes. *Methods Enzymol.* 318, 3–21. doi: 10.1016/s0076-6879(00)18040-1
- Buratti, E., and Baralle, F. E. (2004). Influence of RNA secondary structure on the pre-mRNA splicing process. *Mol. Cell. Biol.* 24, 10505–10514. doi: 10.1128/mcb.24.24.10505-10514.2004
- Busan, S., and Weeks, K. M. (2018). Accurate detection of chemical modifications in RNA by mutational profiling (MaP) with ShapeMapper 2. *RNA* 24, 143–148. doi: 10.1261/rna.061945.117
- Busan, S., Weidmann, C. A., Sengupta, A., and Weeks, K. M. (2019). Guidelines for SHAPE reagent choice and detection strategy for RNA structure probing studies. *Biochemistry* 58, 2655–2664. doi: 10.1021/acs.biochem.8b01218
- Butcher, S. E., and Pyle, A. M. (2011). The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks. *Acc. Chem. Res.* 44, 1302–1311. doi: 10.1021/ar200098t
- Byron, O., and Gilbert, R. J. (2000). Neutron scattering: good news for biotechnology. *Curr. Opin. Biotechnol.* 11, 72–80. doi: 10.1016/s0958-1669(99)00057-9
- Calonaci, N., Jones, A., Cuterello, F., Sattler, M., and Bussi, G. (2020). Machine learning a model for RNA structure prediction. *arXiv* [Preprint]. Available online at: <http://arxiv.org/abs/2004.00351> (accessed August 7, 2020).
- Cammass, A., and Millevoi, S. (2017). RNA G-quadruplexes: emerging mechanisms in disease. *Nucleic Acids Res.* 45, 1584–1595.
- Cate, J. H., and Doudna, J. A. (2000). Solving large RNA structures by X-ray crystallography. *Methods Enzymol.* 317, 169–180. doi: 10.1016/s0076-6879(00)17014-4
- Cech, T. R., and Steitz, J. A. (2014). The noncoding RNA revolution—trashing old rules to forge new ones. *Cell* 157, 77–94. doi: 10.1016/j.cell.2014.03.008

AUTHOR CONTRIBUTIONS

ZM designed the project. BL drafted the manuscript. All authors contributed to the manuscript.

FUNDING

ZM was supported by a Single Cell Gene Expression Atlas grant from the Wellcome Trust (nr. 108437/Z/15/Z). YC was supported by the National Natural Science Foundation of China under Grant (number 81973243, 81830108, and 31401130). This work was funded by the Shanghai Fourth People's Hospital.

- Chaloner, P. A. (1990). Two-dimensional NMR methods for establishing molecular connectivity; a chemist's guide to experiment selection, performance and interpretation. *J. Organomet. Chem.* 386, C16–C17. doi: 10.1016/0022-328x(90)85255-w
- Cheng, C. Y., Chou, F.-C., Kladwang, W., Tian, S., Cordero, P., and Das, R. (2015). Consistent global structures of complex RNA states through multidimensional chemical mapping. *eLife* 4:e07600.
- Cheng, C. Y., Kladwang, W., Yesselman, J. D., and Das, R. (2017). RNA structure inference through chemical mapping after accidental or intentional mutations. *Proc. Natl. Acad. Sci. U.S.A.* 114, 9876–9881. doi: 10.1073/pnas.1619897114
- Climie, S. C., and Friesen, J. D. (1988). In vivo and in vitro structural analysis of the rplJ mRNA leader of *Escherichia coli*. Protection by bound L10-L7/L12. *J. Biol. Chem.* 263, 15166–15175.
- Coimbatore Narayanan, B., Westbrook, J., Ghosh, S., Petrov, A. I., Sweeney, B., Zirbel, C. L., et al. (2014). The nucleic acid database: new features and capabilities. *Nucleic Acids Res.* 42, D114–D122.
- Cordero, P., and Das, R. (2015). Rich RNA structure landscapes revealed by mutate-and-map analysis. *PLoS Comput. Biol.* 11:e1004473. doi: 10.1371/journal.pcbi.1004473
- Costa, M., and Monachello, D. (2014). Probing RNA folding by hydroxyl radical footprinting. *Methods Mol. Biol.* 1086, 119–142. doi: 10.1007/978-1-62703-667-2_7
- Cruz, J. A., Blanchet, M.-F., Boniecki, M., Bujnicki, J. M., Chen, S.-J., Cao, S., et al. (2012). RNA-Puzzles: a CASP-like evaluation of RNA three-dimensional structure prediction. *RNA* 18, 610–625. doi: 10.1261/rna.031054.111
- Cruz, J. A., and Westhof, E. (2011). Sequence-based identification of 3D structural modules in RNA with RMDetect. *Nat. Methods* 8, 513–521. doi: 10.1038/nmeth.1603
- Das, R., and Baker, D. (2007). Automated de novo prediction of native-like RNA tertiary structures. *Proc. Natl. Acad. Sci. U.S.A.* 104, 14664–14669. doi: 10.1073/pnas.0703836104
- Das, R., Karanicolas, J., and Baker, D. (2010). Atomic accuracy in predicting and designing noncanonical RNA structure. *Nat. Methods* 7, 291–294. doi: 10.1038/nmeth.1433
- Das, R., Kudaravalli, M., Jonikas, M., Laederach, A., Fong, R., Schwans, J. P., et al. (2008). Structural inference of native and partially folded RNA by high-throughput contact mapping. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4144–4149. doi: 10.1073/pnas.0709032105
- Davis, B., Afshar, M., Varani, G., Murchie, A. I. H., Karn, J., Lentzen, G., et al. (2004). Rational design of inhibitors of HIV-1 TAR RNA through the stabilisation of electrostatic “hot spots”. *J. Mol. Biol.* 336, 343–356. doi: 10.1016/j.jmb.2003.12.046
- Dawson, W. K., and Bujnicki, J. M. (2016). Computational modeling of RNA 3D structures and interactions. *Curr. Opin. Struct. Biol.* 37, 22–28. doi: 10.1016/j.sbi.2015.11.007
- De Leonardi, E., Lutz, B., Ratz, S., Cocco, S., Monasson, R., Schug, A., et al. (2015). Direct-Coupling Analysis of nucleotide coevolution facilitates RNA secondary and tertiary structure prediction. *Nucleic Acids Res.* 43, 10444–10455.
- Deigan, K. E., Li, T. W., Mathews, D. H., and Weeks, K. M. (2009). Accurate SHAPE-directed RNA structure determination. *Proc. Natl. Acad. Sci. U.S.A.* 106, 97–102. doi: 10.1073/pnas.0806929106
- Ding, F., Lavender, C. A., Weeks, K. M., and Dokholyan, N. V. (2012). Three-dimensional RNA structure refinement by hydroxyl radical probing. *Nat. Methods* 9, 603–608. doi: 10.1038/nmeth.1976
- Ding, Y., Tang, Y., Kwok, C. K., Zhang, Y., Bevilacqua, P. C., and Assmann, S. M. (2014). In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* 505, 696–700. doi: 10.1038/nature12756
- Doerr, A. (2017). Cryo-electron tomography. *Nat. Methods* 14:34. doi: 10.1038/nmeth.4115
- Duncan, C. D. S., and Weeks, K. M. (2008). SHAPE analysis of long-range interactions reveals extensive and thermodynamically preferred misfolding in a fragile group I intron RNA[†]. *Biochemistry* 47, 8504–8513. doi: 10.1021/bi800207b
- Ehresmann, C., Baudin, F., Mougél, M., Romby, P., Ebel, J.-P., and Ehresmann, B. (1987). Probing the structure of RNAs in solution. *Nucleic Acids Res.* 15, 9109–9128. doi: 10.1093/nar/15.22.9109
- Evans, P., and McCoy, A. (2008). An introduction to molecular replacement. *Acta Crystallogr. D Biol. Crystallogr.* 64, 1–10. doi: 10.1053/jarr.2001.21705
- Fang, X., Wang, J., O'Carroll, I. P., Mitchell, M., Zuo, X., Wang, Y., et al. (2013). An unusual topological structure of the HIV-1 Rev response element. *Cell* 155, 594–605. doi: 10.1016/j.cell.2013.10.008
- Feng, C., Chan, D., Joseph, J., Muuronen, M., Coldren, W. H., Dai, N., et al. (2018). Light-activated chemical probing of nucleobase solvent accessibility inside cells. *Nat. Chem. Biol.* 14, 276–283. doi: 10.1038/nchembio.2548
- Flores, S. C., and Altman, R. B. (2010). Turning limited experimental information into 3D models of RNA. *RNA* 16, 1769–1778. doi: 10.1261/rna.2112110
- Flores, S. C., Wan, Y., Russell, R., and Altman, R. B. (2010). Predicting RNA structure by multiple template homology modeling. *Pac. Symp. Biocomput.* 2020, 216–227. doi: 10.1142/9789814295291_0024
- Ganser, L. R., Kelly, M. L., Herschlag, D., and Al-Hashimi, H. M. (2019). The roles of structural dynamics in the cellular functions of RNAs. *Nat. Rev. Mol. Cell Biol.* 20, 474–489. doi: 10.1038/s41580-019-0136-0
- Gilmore, J. L., Aizaki, H., Yoshida, A., Deguchi, K., Kumeta, M., Junghof, J., et al. (2014). Nanoimaging of ssRNA: genome architecture of the hepatitis C virus revealed by atomic force microscopy. *J. Nanomed. Nanotechnol.* 55, 1–7.
- Gilmore, J. L., Yoshida, A., Hejna, J. A., and Takeyasu, K. (2017). Visualization of conformational variability in the domains of long single-stranded RNA molecules. *Nucleic Acids Res.* 45, 8493–8507. doi: 10.1093/nar/gkx502
- Halvorsen, M., Martin, J. S., Broadaway, S., and Laederach, A. (2010). Disease-associated mutations that alter the RNA structural ensemble. *PLoS Genet.* 6:e1001074. doi: 10.1371/journal.pcbi.1001074
- He, X., Wang, J., Wang, J., and Xiao, Y. (2020). Improving RNA secondary structure prediction using direct coupling analysis. *Chin. Phys. B.* 29:078702. doi: 10.1088/1674-1056/ab889d
- Hector, R. D., Burlacu, E., Aitken, S., Le Bihan, T., Tuijtel, M., Zaplatina, A., et al. (2014). Snapshots of pre-rRNA structural flexibility reveal eukaryotic 40S assembly dynamics at nucleotide resolution. *Nucleic Acids Res.* 42, 12138–12154. doi: 10.1093/nar/gku815
- Hermann, T., and Westhof, E. (1999). Non-watson-crick base pairs in RNA-protein recognition. *Chem. Biol.* 6, R335–R343.
- Herrmann, T., Güntert, P., and Wüthrich, K. (2002). Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* 319, 209–227. doi: 10.1016/s0022-2836(02)00241-3
- Holmstrom, E. D., and Nesbitt, D. J. (2016). Biophysical Insights from temperature-dependent single-molecule Förster resonance energy transfer. *Annu. Rev. Phys. Chem.* 67, 441–465. doi: 10.1146/annurev-physchem-040215-112544
- Homan, P. J., Favorov, O. V., Lavender, C. A., Kursun, O., Ge, X., Busan, S., et al. (2014a). Single-molecule correlated chemical probing of RNA. *Proc. Natl. Acad. Sci. U.S.A.* 111, 13858–13863. doi: 10.1073/pnas.1407306111
- Homan, P. J., Tandon, A., Rice, G. M., Ding, F., Dokholyan, N. V., and Weeks, K. M. (2014b). RNA tertiary structure analysis by 2'-hydroxyl molecular interference. *Biochemistry* 53, 6825–6833. doi: 10.1021/bi501218g
- Horos, R., Büscher, M., Kleinendorst, R., Alleaume, A.-M., Tarafder, A. K., Schwarzl, T., et al. (2019). The small non-coding vault RNA1-1 Acts as a riboregulator of autophagy. *Cell* 176, 1054–1067.e12.
- Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., et al. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395, 497–506.
- Huang, L., Ashraf, S., and Lilley, D. M. J. (2019a). The role of RNA structure in translational regulation by L7Ae protein in archaea. *RNA* 25, 60–69. doi: 10.1261/rna.068510.118
- Huang, L., Wang, J., Watkins, A. M., Das, R., and Lilley, D. M. J. (2019b). Structure and ligand binding of the glutamine-II riboswitch. *Nucleic Acids Res.* 47, 7666–7675. doi: 10.1093/nar/gkz539
- Huang, W., Emani, P. S., Varani, G., and Drobny, G. P. (2017). Ultraslow domain motions in HIV-1 TAR RNA revealed by solid-state deuterium NMR. *J. Phys. Chem. B* 121, 110–117. doi: 10.1021/acs.jpcc.6b11041
- Hura, G. L., Menon, A. L., Hammel, M., Rambo, R. P., Poole, F. L. II, Tsutakawa, S. E., et al. (2009). Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). *Nat. Methods* 6, 606–612. doi: 10.1038/nmeth.1353
- Husale, S., Persson, H. H. J., and Sahin, O. (2009). DNA nanomechanics allows direct digital detection of complementary DNA and microRNA targets. *Nature* 462, 1075–1078. doi: 10.1038/nature08626

- Jain, N., Morgan, C. E., Rife, B. D., Salemi, M., and Tolbert, B. S. (2016). Solution structure of the HIV-1 intron splicing silencer and its interactions with the UPI domain of heterogeneous nuclear ribonucleoprotein (hnRNP) A1. *J. Biol. Chem.* 291, 2331–2344. doi: 10.1074/jbc.m115.674564
- Jian, Y., Wang, X., Qiu, J., Wang, H., Liu, Z., Zhao, Y., et al. (2019). DIRECT: RNA contact predictions by integrating structural patterns. *BMC Bioinform.* 20:497. doi: 10.1186/s12859-019-3099-4
- Jonikas, M. A., Radmer, R. J., Laederach, A., Das, R., Pearlman, S., Herschlag, D., et al. (2009). Coarse-grained modeling of large RNA molecules with knowledge-based potentials and structural filters. *RNA* 15, 189–199. doi: 10.1261/rna.1270809
- Jossinet, F. (2015). “Assemble2: an interactive graphical environment dedicated to the study and construction of RNA architectures,” in *Proceedings of the 2015 IEEE 1st International Workshop on Virtual and Augmented Reality for Molecular Science (VARMS@IEEEVR)*, Arles.
- Jossinet, F., Ludwig, T. E., and Westhof, E. (2010). Assemble: an interactive graphical tool to analyze and build RNA architectures at the 2D and 3D levels. *Bioinformatics* 26, 2057–2059. doi: 10.1093/bioinformatics/btq321
- Kalvari, I., Argasinska, J., Quinones-Olvera, N., Nawrocki, E. P., Rivas, E., Eddy, S. R., et al. (2018). Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res.* 46, D335–D342.
- Kappel, K., Liu, S., Larsen, K. P., Skiniotis, G., Puglisi, E. V., Puglisi, J. D., et al. (2018). De novo computational RNA modeling into cryo-EM maps of large ribonucleoprotein complexes. *Nat. Methods* 15, 947–954. doi: 10.1038/s41592-018-0172-2
- Kappel, K., Zhang, K., Su, Z., Kladwang, W., Li, S., Pintilie, G., et al. (2019). Ribosolve: rapid determination of three-dimensional RNA-only structures. *bioRxiv* [Preprint]. doi: 10.1101/717801
- Kappel, K., Zhang, K., Su, Z., Watkins, A. M., Kladwang, W., Li, S., et al. (2020). Accelerated cryo-EM-guided determination of three-dimensional RNA-only structures. *Nat. Methods* 17, 699–707. doi: 10.1038/s41592-020-0878-9
- Karunatilaka, K. S., and Rueda, D. (2009). Single-molecule fluorescence studies of RNA: a Decade's progress. *Chem. Phys. Lett.* 476, 1–10. doi: 10.1016/j.cplett.2009.06.001
- Keane, S. C., Heng, X., Lu, K., Kharytonchyk, S., Ramakrishnan, V., Carter, G., et al. (2015). RNA structure. Structure of the HIV-1 RNA packaging signal. *Science* 348, 917–921.
- Kladwang, W., and Das, R. (2010). A mutate-and-map strategy for inferring base pairs in structured nucleic acids: proof of concept on a DNA/RNA helix. *Biochemistry* 49, 7414–7416.
- Kladwang, W., VanLang, C. C., Cordero, P., and Das, R. (2011a). A two-dimensional mutate-and-map strategy for non-coding RNA structure. *Nat. Chem.* 3, 954–962. doi: 10.1038/nchem.1176
- Kladwang, W., VanLang, C. C., Cordero, P., and Das, R. (2011b). Understanding the errors of SHAPE-directed RNA structure modeling. *Biochemistry* 50, 8049–8056. doi: 10.1021/bi200524n
- Koch, M. H., Vachette, P., and Svergun, D. I. (2003). Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q. Rev. Biophys.* 36, 147–227. doi: 10.1017/s0033583503003871
- Krokhutin, A., Houlihan, K., and Dokholyan, N. V. (2015). iFoldRNA v2: folding RNA with constraints. *Bioinformatics* 31, 2891–2893. doi: 10.1093/bioinformatics/btv221
- Krokhutin, A., Mustoe, A. M., Weeks, K. M., and Dokholyan, N. V. (2017). Direct identification of base-paired RNA nucleotides by correlated chemical probing. *RNA* 23, 6–13. doi: 10.1261/rna.058586.116
- Kubota, M., Tran, C., and Spitale, R. C. (2015). Progress and challenges for chemical probing of RNA structure inside living cells. *Nat. Chem. Biol.* 11, 933–941. doi: 10.1038/nchembio.1958
- Kwok, C. K., Ding, Y., Tang, Y., Assmann, S. M., and Bevilacqua, P. C. (2013). Determination of in vivo RNA structure in low-abundance transcripts. *Nat. Commun.* 4:2971.
- Kwok, C. K., Tang, Y., Assmann, S. M., and Bevilacqua, P. C. (2015). The RNA structurome: transcriptome-wide structure probing with next-generation sequencing. *Trends Biochem. Sci.* 40, 221–232. doi: 10.1016/j.tibs.2015.02.005
- Latham, J. A., and Cech, T. R. (1989). Defining the inside and outside of a catalytic RNA molecule. *Science* 245, 276–282. doi: 10.1126/science.2501870
- Latham, M. P., Brown, D. J., McCallum, S. A., and Pardi, A. (2005). NMR methods for studying the structure and dynamics of RNA. *ChemBiochem* 6, 1492–1505. doi: 10.1002/cbic.200500123
- Lee, B., Flynn, R. A., Kadina, A., Guo, J. K., Kool, E. T., and Chang, H. Y. (2017). Comparison of SHAPE reagents for mapping RNA structures inside living cells. *RNA* 23, 169–174. doi: 10.1261/rna.058784.116
- Lehnert, V., Jaeger, L., Michele, F., and Westhof, E. (1996). New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the Tetrahymena thermophila ribozyme. *Chem. Biol.* 3, 993–1009. doi: 10.1016/s1074-5521(96)90166-90160
- Lescoute, A. (2006). Topology of three-way junctions in folded RNAs. *RNA* 12, 83–93. doi: 10.1261/rna.2208106
- Levitt, M. (1969). Detailed molecular model for transfer ribonucleic acid. *Nature* 224, 759–763. doi: 10.1038/224759a0
- Li, H., and Aviran, S. (2018). Statistical modeling of RNA structure profiling experiments enables parsimonious reconstruction of structure landscapes. *Nat. Commun.* 9:606.
- Liao, M., Cao, E., Julius, D., and Cheng, Y. (2013). Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* 504, 107–112. doi: 10.1038/nature12822
- Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* 75, 861–877.
- Long, Y., Wang, X., Youmans, D. T., and Cech, T. R. (2017). How do lncRNAs regulate transcription? *Sci. Adv.* 3:eaa02110. doi: 10.1126/sciadv.aao2110
- Lorenz, R., Bernhart, S. H., Höner Zu Siederdisen, C., Tafer, H., Flamm, C., Stadler, P. F., et al. (2011). ViennaRNA Package 2.0. *Algorith. Mol. Biol.* 6:26.
- Loughrey, D., Watters, K. E., Settle, A. H., and Lucks, J. B. (2014). SHAPE-Seq 2.0: systematic optimization and extension of high-throughput chemical probing of RNA secondary structure with next generation sequencing. *Nucleic Acids Res.* 42:gku909. doi: 10.1093/nar/gku909
- Lucks, J. B., Mortimer, S. A., Trapnell, C., Luo, S., Aviran, S., Schroth, G. P., et al. (2011). Multiplexed RNA structure characterization with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). *Proc. Natl. Acad. Sci. U.S.A.* 108, 11063–11068. doi: 10.1073/pnas.1106501108
- Luco, R. F., and Misteli, T. (2011). More than a splicing code: integrating the role of RNA, chromatin and non-coding RNA in alternative splicing regulation. *Curr. Opin. Genet. Dev.* 21, 366–372. doi: 10.1016/j.gde.2011.03.004
- Magnus, M., Antczak, M., Zok, T., Wiedemann, J., Lukasiak, P., Cao, Y., et al. (2020). RNA-Puzzles toolkit: a computational resource of RNA 3D structure benchmark datasets, structure manipulation, and evaluation tools. *Nucleic Acids Res.* 48, 576–588.
- Magnus, M., Kappel, K., Das, R., and Bujnicki, J. M. (2019). RNA 3D structure prediction guided by independent folding of homologous sequences. *BMC Bioinform.* 20:512. doi: 10.1186/s12859-019-3120-y
- Magnus, M., Matelska, D., Lach, G., Chojnowski, G., Boniecki, M. J., Purta, E., et al. (2014). Computational modeling of RNA 3D structures, with the aid of experimental restraints. *RNA Biol.* 11, 522–536. doi: 10.4161/rna.28826
- Manfredonia, I., Nithin, C., Ponce-Salvatierra, A., Ghosh, P., Wirecki, T. K., Marinus, T., et al. (2020). Genome-wide mapping of therapeutically-relevant SARS-CoV-2 RNA structures. *bioRxiv* [Preprint]. doi: 10.1101/2020.06.15.151647
- Manz, C., Kobitski, A. Y., Samanta, A., Keller, B. G., Jäschke, A., and Nienhaus, G. U. (2017). Single-molecule FRET reveals the energy landscape of the full-length SAM-I riboswitch. *Nat. Chem. Biol.* 13, 1172–1178. doi: 10.1038/nchembio.2476
- Marchanka, A., Ahmed, M., and Carlomagno, T. (2018). A solid view on RNA: solid-state NMR of RNA and RNP complexes. *Biophys. J.* 114:366.
- Marchanka, A., Simon, B., Althoff-Ospelt, G., and Carlomagno, T. (2015). RNA structure determination by solid-state NMR spectroscopy. *Nat. Commun.* 6:7024.
- Marion, D. (2013). An introduction to biological NMR spectroscopy. *Mol. Cell. Proteom.* 12, 3006–3025.
- Markowitz, F. (2017). All biology is computational biology. *PLoS Biol.* 15:e2002050. doi: 10.1371/journal.pcbi.2002050

- McCoy, A. J., and Read, R. J. (2010). Experimental phasing: best practice and pitfalls. *Acta Crystallogr. D Biol. Crystallogr.* 66, 458–469. doi: 10.1107/s0907444910006335
- McGinnis, J. L., Liu, Q., Lavender, C. A., Devaraj, A., McClory, S. P., Fredrick, K., et al. (2015). In-cell SHAPE reveals that free 30S ribosome subunits are in the inactive state. *Proc. Natl. Acad. Sci. U.S.A.* 112, 2425–2430. doi: 10.1073/pnas.1411514112
- McGinnis, J. L., and Weeks, K. M. (2014). Ribosome RNA assembly intermediates visualized in living cells. *Biochemistry* 53, 3237–3247. doi: 10.1021/bi500198b
- Merino, E. J., Wilkinson, K. A., Coughlan, J. L., and Weeks, K. M. (2005). RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc.* 127, 4223–4231.
- Merk, A., Bartesaghi, A., Banerjee, S., Falconieri, V., Rao, P., Davis, M. I., et al. (2016). Breaking Cryo-EM resolution barriers to facilitate drug discovery. *Cell* 165, 1698–1707. doi: 10.1016/j.cell.2016.05.040
- Miao, Z., Adamiak, R. W., Antczak, M., Batey, R. T., Becka, A. J., Biesiada, M., et al. (2017). RNA-Puzzles Round III: 3D RNA structure prediction of five riboswitches and one ribozyme. *RNA* 23, 655–672. doi: 10.1261/rna.060368.116
- Miao, Z., Adamiak, R. W., Antczak, M., Boniecki, M. J., Bujnicki, J., Chen, S.-J., et al. (2020). RNA-Puzzles Round IV: 3D structure predictions of four ribozymes and two aptamers. *RNA* 26, 982–995. doi: 10.1261/rna.075341.120
- Miao, Z., Adamiak, R. W., Blanchet, M.-F., Boniecki, M., Bujnicki, J. M., Chen, S.-J., et al. (2015). RNA-Puzzles round II: assessment of RNA structure prediction programs applied to three large RNA structures. *RNA* 21, 1066–1084.
- Miao, Z., and Westhof, E. (2017). RNA structure: advances and assessment of 3D structure prediction. *Annu. Rev. Biophys.* 46, 483–503. doi: 10.1146/annurev-biophys-070816-034125
- Michel, F., and Westhof, E. (1990). Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.* 216, 585–610. doi: 10.1016/0022-2836(90)90386-z
- Millar, D. P. (1996). Fluorescence studies of DNA and RNA structure and dynamics. *Curr. Opin. Struct. Biol.* 6, 322–326. doi: 10.1016/s0959-440x(96)80050-80059
- Mitchell, D., Ritchey, L. E., Park, H., Babitzke, P., Assmann, S. M., and Bevilacqua, P. C. (2018). Glyoxals as in vivo RNA structural probes of guanine base-pairing. *RNA* 24, 114–124. doi: 10.1261/rna.064014.117
- Moazed, D., Robertson, J. M., and Noller, H. F. (1988). Interaction of elongation factors EF-G and EF-Tu with a conserved loop in 23S RNA. *Nature* 334, 362–364. doi: 10.1038/334362a0
- Mortimer, S. A., Johnson, J. S., and Weeks, K. M. (2009). Quantitative analysis of RNA solvent accessibility by N-silylation of guanosine. *Biochemistry* 48, 2109–2114. doi: 10.1021/bi801939g
- Murata, K., and Wolf, M. (2018). Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. *Biochim. Biophys. Acta Gen. Subj.* 1862, 324–334. doi: 10.1016/j.bbagen.2017.07.020
- Mustoe, A. M., Brooks, C. L., and Al-Hashimi, H. M. (2014). Hierarchy of RNA functional dynamics. *Annu. Rev. Biochem.* 83, 441–466. doi: 10.1146/annurev-biochem-060713-035524
- Mustoe, A. M., Busan, S., Rice, G. M., Hajdin, C. E., Peterson, B. K., Ruda, V. M., et al. (2018). Pervasive regulatory functions of mRNA structure revealed by high-resolution SHAPE probing. *Cell* 173, 181–195.e18.
- Nguyen, T. C., Cao, X., Yu, P., Xiao, S., Lu, J., Biase, F. H., et al. (2016). Mapping RNA-RNA interactome and RNA structure in vivo by MARIO. *Nat. Commun.* 7:12023.
- Noller, H., and Woese, C. (1981). Secondary structure of 16S ribosomal RNA. *Science* 212, 403–411. doi: 10.1126/science.6163215
- Norouzi, R., Wickles, S., Leidig, C., Becker, T., Schmid, V. J., Beckmann, R., et al. (2013). Automatic post-picking using MAPPOS improves particle image detection from cryo-EM micrographs. *J. Struct. Biol.* 182, 59–66. doi: 10.1016/j.jsb.2013.02.008
- Norris, M., Kwok, C. K., Cheema, J., Hartley, M., Morris, R. J., Aviran, S., et al. (2017). FoldAtlas: a repository for genome-wide RNA structure probing data. *Bioinformatics* 33, 306–308. doi: 10.1093/bioinformatics/btw611
- Novoa, E. M., Beaudoin, J. D., Giraldez, A. J., and Mattick, J. S. (2017). Best practices for genome-wide RNA structure analysis: combination of mutational profiles and drop-off information. *bioRxiv* [Preprint]. Available online at: <https://www.biorxiv.org/content/10.1101/176883v1.abstract> (accessed March 31, 2020).
- Orlovsky, N. I., Al-Hashimi, H. M., and Oas, T. G. (2020). Exposing hidden high-affinity RNA conformational states. *J. Am. Chem. Soc.* 142, 907–921. doi: 10.1021/jacs.9b10535
- Pallesen, J., Dong, M., Besenbacher, F., and Kjems, J. (2009). Structure of the HIV-1 Rev response element alone and in complex with regulator of virion (Rev) studied by atomic force microscopy. *FEBS J.* 276, 4223–4232. doi: 10.1111/j.1742-4658.2009.07130.x
- Parisien, M., Cruz, J. A., Westhof, F., and Major, F. (2009). New metrics for comparing and assessing discrepancies between RNA 3D structures and models. *RNA* 15, 1875–1885. doi: 10.1261/rna.1700409
- Parisien, M., and Major, F. (2008). The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* 452, 51–55. doi: 10.1038/nature06684
- Peattie, D. A., and Herr, W. (1981). Chemical probing of the tRNA-ribosome complex. *Proc. Natl. Acad. Sci. U.S.A.* 78, 2273–2277. doi: 10.1073/pnas.78.4.2273
- Plumridge, A., Katz, A. M., Calvey, G. D., Elber, R., Kirmizialtin, S., and Pollack, L. (2018). Revealing the distinct folding phases of an RNA three-helix junction. *Nucleic Acids Res.* 46, 7354–7365. doi: 10.1093/nar/gky363
- Ponce-Salvatierra, A., Astha Merdas, K., Nithin, C., Ghosh, P., and Mukherjee, S. (2019). Computational modeling of RNA 3D structure based on experimental data. *Biosci. Rep.* 39:430. doi: 10.1042/BSR20180430
- Popenda, M., Szachniuk, M., Antczak, M., Purzycka, K. J., Lukasiak, P., Bartol, N., et al. (2012). Automated 3D structure composition for large RNAs. *Nucleic Acids Res.* 40:e112. doi: 10.1093/nar/gks339
- Popenda, M., Szachniuk, M., Blazewicz, M., Wasik, S., Burke, E. K., Blazewicz, J., et al. (2010). RNA FRABASE 2.0: an advanced web-accessible database with the capacity to search the three-dimensional fragments within RNA structures. *BMC Bioinform.* 11:231. doi: 10.1186/1471-2105-11-231
- Pyle, A. M., Murphy, F. L., and Cech, T. R. (1992). RNA substrate binding site in the catalytic core of the *Tetrahymena* ribozyme. *Nature* 358, 123–128. doi: 10.1038/358123a0
- Rambo, R. P., and Tainer, J. A. (2013). Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* 496, 477–481. doi: 10.1038/nature12070
- Regulski, E. E., and Breaker, R. R. (2008). In-line probing analysis of riboswitches. *Methods Mol. Biol.* 419, 53–67. doi: 10.1007/978-1-59745-033-1_4
- Rivas, E., Clements, J., and Eddy, S. R. (2017). A statistical test for conserved RNA structure shows lack of evidence for structure in lncRNAs. *Nat. Methods* 14, 45–48. doi: 10.1038/nmeth.4066
- Rivas, E., Clements, J., and Eddy, S. R. (2020). Estimating the power of sequence covariation for detecting conserved RNA structure. *Bioinformatics* 36, 3072–3076. doi: 10.1093/bioinformatics/btaa080
- Rother, K., Rother, M., Boniecki, M., Puton, T., and Bujnicki, J. M. (2011a). RNA and protein 3D structure modeling: similarities and differences. *J. Mol. Model.* 17, 2325–2336. doi: 10.1007/s00894-010-0951-x
- Rother, M., Rother, K., Puton, T., and Bujnicki, J. M. (2011b). ModeRNA: a tool for comparative modeling of RNA 3D structure. *Nucleic Acids Res.* 39, 4007–4022. doi: 10.1093/nar/gkq1320
- Rouskin, S., Zubradt, M., Washietl, S., Kellis, M., and Weissman, J. S. (2014). Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature* 505, 701–705. doi: 10.1038/nature12894
- Russel, D., Lasker, K., Webb, B., and Velázquez-Muriel, J. (2012). Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies. *PLoS Biol* 10:e1001244. doi: 10.1371/journal.pbio.1001244
- Ruszkowska, A., Ruszkowski, M., Hulewicz, J. P., Dauter, Z., and Brown, J. A. (2020). Molecular structure of a UA-U-rich RNA triple helix with 11 consecutive base triples. *Nucleic Acids Res.* 48, 3304–3314. doi: 10.1093/nar/gkz1222
- Sahin, O., Magonov, S., Su, C., Quate, C. F., and Solgaard, O. (2007). An atomic force microscope tip designed to measure time-varying nanomechanical forces. *Nat. Nanotechnol.* 2, 507–514. doi: 10.1038/nnano.2007.226
- Salmon, L., Yang, S., and Al-Hashimi, H. M. (2014). Advances in the determination of nucleic acid conformational ensembles. *Annu. Rev. Phys. Chem.* 65, 293–316. doi: 10.1146/annurev-physchem-040412-110059

- Sarrazin-Gendron, R., Reinharz, V., Oliver, C. G., Moitessier, N., and Waldspühl, J. (2019). Automated, customizable and efficient identification of 3D base pair modules with BayesPairing. *Nucleic Acids Res.* 47, 3321–3332. doi: 10.1093/nar/gkz102
- Sayers, E. W., Barrett, T., Benson, D. A., Bolton, E., Bryant, S. H., Canese, K., et al. (2011). Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 39, D38–D51.
- Schlick, T., and Pyle, A. M. (2017). Opportunities and challenges in RNA structural modeling and design. *Biophys. J.* 113, 225–234. doi: 10.1016/j.bpj.2016.12.037
- Schnablegger, H., and Singh, Y. (2011). *The SAXS Guide: Getting Acquainted With the Principles*. Austria: Anton Paar GmbH.
- Schön, P. (2018). Atomic force microscopy of RNA: state of the art and recent advancements. *Semin. Cell Dev. Biol.* 73, 209–219. doi: 10.1016/j.semcdb.2017.08.040
- Schuster, S. C. (2008). Next-generation sequencing transforms today's biology. *Nat. Methods* 5, 16–18. doi: 10.1038/nmeth1156
- Sekar, R. B., and Periasamy, A. (2003). Fluorescence resonance energy transfer (FRET) microscopy imaging of live cell protein localizations. *J. Cell Biol.* 160, 629–633. doi: 10.1083/jcb.200210140
- Selega, A., Sirocchi, C., Iosub, I., Granneman, S., and Sanguinetti, G. (2017). Robust statistical modeling improves sensitivity of high-throughput RNA structure probing experiments. *Nat. Methods* 14, 83–89. doi: 10.1038/nmeth.4068
- Sexton, A. N., Wang, P. Y., Rutenberg-Schoenberg, M., and Simon, M. D. (2017). Interpreting reverse transcriptase termination and mutation events for greater insight into the chemical probing of RNA. *Biochemistry* 56, 4713–4721. doi: 10.1021/acs.biochem.7b00323
- Shahin, V., and Barrera, N. P. (2008). Providing unique insight into cell biology via atomic force microscopy. *Int. Rev. Cytol.* 265, 227–252. doi: 10.1016/s0074-7696(07)65006-2
- Sharma, S., Ding, F., and Dokholyan, N. V. (2008). iFoldRNA: three-dimensional RNA structure prediction and folding. *Bioinformatics* 24, 1951–1952. doi: 10.1093/bioinformatics/btn328
- Sharp, P. A. (2009). The centrality of RNA. *Cell* 136, 577–580. doi: 10.1016/j.cell.2009.02.007
- Shi, Y. (2014). A glimpse of structural biology through X-ray crystallography. *Cell* 159, 995–1014. doi: 10.1016/j.cell.2014.10.051
- Siegfried, N. A., Busan, S., Rice, G. M., Nelson, J. A. E., and Weeks, K. M. (2014). RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods* 11, 959–965. doi: 10.1038/nmeth.3029
- Skou, S., Gillilan, R. E., and Ando, N. (2014). Synchrotron-based small-angle X-ray scattering of proteins in solution. *Nat. Protoc.* 9, 1727–1739. doi: 10.1038/nprot.2014.116
- Sloat, N., Liu, J.-W., and Schroeder, S. J. (2017). Swellix: a computational tool to explore RNA conformational space. *BMC Bioinform.* 18:504. doi: 10.1186/s12859-017-1910-7
- Smola, M. J., Calabrese, J. M., and Weeks, K. M. (2015a). Detection of RNA-protein interactions in living cells with SHAPE. *Biochemistry* 54, 6867–6875. doi: 10.1021/acs.biochem.5b00977
- Smola, M. J., Rice, G. M., Busan, S., Siegfried, N. A., and Weeks, K. M. (2015b). Selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis. *Nat. Protoc.* 10, 1643–1669. doi: 10.1038/nprot.2015.103
- Smola, M. J., Christy, T. W., Inoue, K., Nicholson, C. O., Friedersdorf, M., Keene, J. D., et al. (2016). SHAPE reveals transcript-wide interactions, complex structural domains, and protein interactions across the Xist lncRNA in living cells. *Proc. Natl. Acad. Sci. U.S.A.* 113, 10322–10327. doi: 10.1073/pnas.1600081113
- Smola, M. J., and Weeks, K. M. (2018). In-cell RNA structure probing with SHAPE-MaP. *Nat. Protoc.* 13, 1181–1195. doi: 10.1038/nprot.2018.010
- Spitale, R. C., Crisalli, P., Flynn, R. A., Torre, E. A., Kool, E. T., and Chang, H. Y. (2013). RNA SHAPE analysis in living cells. *Nat. Chem. Biol.* 9, 18–20. doi: 10.1038/nchembio.1131
- Spitale, R. C., Flynn, R. A., Zhang, Q. C., Crisalli, P., Lee, B., Jung, J.-W., et al. (2015). Structural imprints in vivo decode RNA regulatory mechanisms. *Nature* 519, 486–490. doi: 10.1038/nature14263
- Stagno, J. R., Liu, Y., Bhandari, Y. R., Conrad, C. E., Panja, S., Swain, M., et al. (2017). Structures of riboswitch RNA reaction states by mix-and-inject XFEL serial crystallography. *Nature* 541, 242–246. doi: 10.1038/nature20599
- Steen, K.-A., Rice, G. M., and Weeks, K. M. (2012). Fingerprinting noncanonical and tertiary RNA structures by differential SHAPE reactivity. *J. Am. Chem. Soc.* 134, 13160–13163. doi: 10.1021/ja304027m
- Stephenson, J. D., Kenyon, J. C., Symmons, M. F., and Lever, A. M. L. (2016). Characterizing 3D RNA structure by single molecule FRET. *Methods* 103, 57–67. doi: 10.1016/j.ymeth.2016.02.004
- Strobel, E. J., Watters, K. E., Loughrey, D., and Lucks, J. B. (2016). RNA systems biology: uniting functional discoveries and structural tools to understand global roles of RNAs. *Curr. Opin. Biotechnol.* 39, 182–191. doi: 10.1016/j.copbio.2016.03.019
- Strobel, E. J., Yu, A. M., and Lucks, J. B. (2018). High-throughput determination of RNA structures. *Nat. Rev. Genet.* 19, 615–634. doi: 10.1038/s41576-018-0034-x
- Suddala, K. C., and Zhang, J. (2019). High-affinity recognition of specific tRNAs by an mRNA anticodon-binding groove. *Nat. Struct. Mol. Biol.* 26, 1114–1122. doi: 10.1038/s41594-019-0335-6
- Sung, H.-L., and Nesbitt, D. J. (2020). DNA hairpin hybridization under extreme pressures: a single-molecule FRET study. *J. Phys. Chem. B* 124, 110–120. doi: 10.1021/acs.jpcc.9b10131
- Svergun, D. I., and Koch, M. H. J. (2003). Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys.* 66:1735. doi: 10.1088/0034-4885/66/10/r05
- Terayama, K., Yamashita, T., Oguchi, T., and Tsuda, K. (2018). Fine-grained optimization method for crystal structure prediction. *NPJ Comput. Mater.* 4:32. doi: 10.1038/s41524-018-0090-y
- Terwilliger, T. C., Adams, P. D., Afonine, P. V., and Sobolev, O. V. (2018). A fully automatic method yielding initial models from high-resolution cryo-electron microscopy maps. *Nat. Methods* 15, 905–908. doi: 10.1038/s41592-018-0173-171
- Theis, C., Höner Zu Siederdisen, C., Hofacker, I. L., and Gorodkin, J. (2013). Automated identification of RNA 3D modules with discriminative power in RNA structural alignments. *Nucleic Acids Res.* 41, 9999–10009. doi: 10.1093/nar/gkt795
- Tijerina, P., Mohr, S., and Russell, R. (2007). DMS footprinting of structured RNAs and RNA-protein complexes. *Nat. Protoc.* 2, 2608–2623. doi: 10.1038/nprot.2007.380
- Tomezsko, P. J., Corbin, V. D. A., Gupta, P., Swaminathan, H., Glasgow, M., Persad, S., et al. (2020). Determination of RNA structural diversity and its role in HIV-1 RNA splicing. *Nature* 582, 438–442. doi: 10.1038/s41586-020-2253-5
- Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C., and Bussi, G. (2014). PLUMED 2: new feathers for an old bird. *Comput. Phys. Commun.* 185, 604–613. doi: 10.1016/j.cpc.2013.09.018
- Turner, B., Melcher, S. E., Wilson, T. J., Norman, D. G., and Lilley, D. M. J. (2005). Induced fit of RNA on binding the L7Ae protein to the kink-turn motif. *RNA* 11, 1192–1200. doi: 10.1261/rna.2680605
- Tuschl, T., Gohlke, C., Jovin, T. M., Westhof, E., and Eckstein, F. (1994). A three-dimensional model for the hammerhead ribozyme based on fluorescence measurements. *Science* 266, 785–789. doi: 10.1126/science.7973630
- Tyrrell, J., McGinnis, J. L., Weeks, K. M., and Pielak, G. J. (2013). The cellular environment stabilizes adenine riboswitch RNA structure. *Biochemistry* 52, 8777–8785. doi: 10.1021/bi401207q
- Valencia-Sanchez, M. A., Liu, J., Hannon, G. J., and Parker, R. (2006). Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev.* 20, 515–524. doi: 10.1101/gad.1399806
- Vaziri, S., Koehl, P., and Aviran, S. (2018). Extracting information from RNA SHAPE data: kalman filtering approach. *PLoS One* 13:e0207029. doi: 10.1371/journal.pbio.1207029
- Vinayagam, D., Mager, T., Apelbaum, A., Bothe, A., Merino, F., Hofnagel, O., et al. (2018). Electron cryo-microscopy structure of the canonical TRPC4 ion channel. *eLife* 7:e36615. doi: 10.7554/eLife.36615
- Walczak, R., Westhof, E., Carbon, P., and Krol, A. (1996). A novel RNA structural motif in the selenocysteine insertion element of eukaryotic selenoprotein mRNAs. *RNA* 2, 367–379.
- Waldsich, C. (2002). RNA chaperone StpA loosens interactions of the tertiary structure in the td group I intron in vivo. *Genes Dev.* 16, 2300–2312. doi: 10.1101/gad.231302

- Waldsich, C., Masquida, B., Westhof, E., and Schroeder, R. (2002). Monitoring intermediate folding states of the td group I intron in vivo. *EMBO J.* 21, 5281–5291. doi: 10.1093/emboj/cdf504
- Wan, Y., Qu, K., Ouyang, Z., and Chang, H. Y. (2013). Genome-wide mapping of RNA structure using nuclease digestion and high-throughput sequencing. *Nat. Protoc.* 8, 849–869. doi: 10.1038/nprot.2013.045
- Wang, F., Gong, H., Liu, G., Li, M., Yan, C., Xia, T., et al. (2016). DeepPicker: a deep learning approach for fully automated particle picking in cryo-EM. *J. Struct. Biol.* 195, 325–336. doi: 10.1016/j.jsb.2016.07.006
- Wang, H. W., and Wang, J. W. (2017). How cryo-electron microscopy and X-ray crystallography complement each other. *Protein Sci.* 26, 32–39. doi: 10.1002/pro.3022
- Wang, J., Mao, K., Zhao, Y., Zeng, C., Xiang, J., Zhang, Y., et al. (2017). Optimization of RNA 3D structure prediction using evolutionary restraints of nucleotide-nucleotide interactions from direct coupling analysis. *Nucleic Acids Res.* 45, 6299–6309. doi: 10.1093/nar/gkx386
- Wang, J., Williams, B., Chirasi, V. R., Krokhotin, A., Das, R., and Dokholyan, N. V. (2019). Limits in accuracy and a strategy of RNA structure prediction using experimental information. *Nucleic Acids Res.* 47, 5563–5572. doi: 10.1093/nar/gkz427
- Washietl, S., Hofacker, I. L., Stadler, P. F., and Kellis, M. (2012). RNA folding with soft constraints: reconciliation of probing data and thermodynamic secondary structure prediction. *Nucleic Acids Res.* 40, 4261–4272. doi: 10.1093/nar/gks009
- Watkins, A. M., Rangan, R., and Das, R. (2020). FARFAR2: Improved De Novo Rosetta Prediction of Complex Global RNA Folds. *Structure* 28, 963–976.e6.
- Watters, K. E., Abbott, T. R., and Lucks, J. B. (2016a). Simultaneous characterization of cellular RNA structure and function with in-cell SHAPE-Seq. *Nucleic Acids Res.* 44:e12. doi: 10.1093/nar/gkv879
- Watters, K. E., Yu, A. M., Strobel, E. J., Settle, A. H., and Lucks, J. B. (2016b). Characterizing RNA structures in vitro and in vivo with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). *Methods* 103, 34–48. doi: 10.1016/j.ymeth.2016.04.002
- Watts, J. M., Dang, K. K., Gorelick, R. J., Leonard, C. W., Bess, J. W. Jr., Swanson, R., et al. (2009). Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature* 460, 711–716. doi: 10.1038/nature08237
- Weeks, K. M. (2010). Advances in RNA structure analysis by chemical probing. *Curr. Opin. Struct. Biol.* 20, 295–304. doi: 10.1016/j.sbi.2010.04.001
- Weinberg, Z., Barrick, J. E., Yao, Z., Roth, A., Kim, J. N., Gore, J., et al. (2007). Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. *Nucleic Acids Res.* 35, 4809–4819. doi: 10.1093/nar/gkm487
- Weinberg, Z., Perreault, J., Meyer, M. M., and Breaker, R. R. (2009). Exceptional structured noncoding RNAs revealed by bacterial metagenome analysis. *Nature* 462, 656–659. doi: 10.1038/nature08586
- Weinberg, Z., Wang, J. X., Bogue, J., Yang, J., Corbino, K., Moy, R. H., et al. (2010). Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol.* 11:R31.
- Weinreb, C., Riesselman, A. J., Ingraham, J. B., Gross, T., Sander, C., and Marks, D. S. (2016). 3D RNA and functional interactions from evolutionary couplings. *Cell* 165, 963–975. doi: 10.1016/j.cell.2016.03.030
- Wemmer, D. E. (1996). "Nucleic acid structure and dynamics from NMR," in *NMR Spectroscopy and its Application to Biomedical Research*, ed. S. K. Sarkar (Amsterdam: Elsevier), 281–312. doi: 10.1016/b978-044489410-6/50008-8
- Westhof, E., and Jaeger, L. (1992). RNA pseudoknots. *Curr. Biol.* 2:300.
- Westhof, E., Romby, P., Romaniuk, P. J., Ebel, J. P., Ehresmann, C., and Ehresmann, B. (1989). Computer modeling from solution data of spinach chloroplast and of *Xenopus laevis* somatic and oocyte 5 S rRNAs. *J. Mol. Biol.* 207, 417–431. doi: 10.1016/0022-2836(89)90264-7
- Wilkinson, K. A., Merino, E. J., and Weeks, K. M. (2006). Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat. Protoc.* 1, 1610–1616. doi: 10.1038/nprot.2006.249
- Xu, X., Zhao, C., and Chen, S.-J. (2019). VfoldLA: a web server for loop assembly-based prediction of putative 3D RNA structures. *J. Struct. Biol.* 207, 235–240. doi: 10.1016/j.jsb.2019.06.002
- Yang, Y., and Wang, S. (2018). RNA characterization by solid-state NMR spectroscopy. *Chemistry* 24, 8698–8707. doi: 10.1002/chem.201705583
- Yaniv, M., Favre, A., and Barrell, B. G. (1969). Structure of transfer RNA: evidence for interaction between two non-adjacent nucleotide residues in tRNA^{Val} from *Escherichia coli*. *Nature* 223, 1331–1333. doi: 10.1038/2231331a0
- Yao, R., Qian, J., and Huang, Q. (2019). Deep-learning with synthetic data enables automated picking of cryo-EM particle images of biological macromolecules. *Bioinformatics* 36, 1252–1259. doi: 10.1093/bioinformatics/btz728
- Yesselman, J. D., and Das, R. (2016). Modeling small noncanonical RNA motifs with the rosetta FARFAR server. *Methods Mol. Biol.* 1490, 187–198. doi: 10.1007/978-1-4939-6433-8_12
- Zarringhalam, K., Meyer, M. M., Dotu, I., Chuang, J. H., and Clote, P. (2012). Integrating chemical footprinting data into RNA secondary structure prediction. *PLoS One* 7:e45160. doi: 10.1371/journal.pbio.045160
- Zhang, K., Keane, S. C., Su, Z., Irobalieva, R. N., Chen, M., Van, V., et al. (2018). Structure of the 30 kDa HIV-1 RNA dimerization signal by a hybrid Cryo-EM, NMR, and molecular dynamics approach. *Structure* 26, 490–498.e3.
- Zhang, K., Li, S., Kappel, K., Pintilie, G., Su, Z., Mou, T.-C., et al. (2019). Cryo-EM structure of a 40 kDa SAM-IV riboswitch RNA at 3.7 Å resolution. *Nat. Commun.* 10:5511.
- Zhao, Y., Huang, Y., Gong, Z., Wang, Y., Man, J., and Xiao, Y. (2012). Automated and fast building of three-dimensional RNA structures. *Sci. Rep.* 2:734.
- Zirbel, C. L., Roll, J., Sweeney, B. A., Petrov, A. I., Pirrung, M., and Leontis, N. B. (2015). Identifying novel sequence variants of RNA 3D motifs. *Nucleic Acids Res.* 43, 7504–7520. doi: 10.1093/nar/gkv651
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31, 3406–3415.

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.

Copyright © 2020 Li, Cao, Westhof and Miao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.