



OPEN ACCESS

EDITED BY
Andrzej Stasiak,
Université de Lausanne, Switzerland

REVIEWED BY
Ping Chen,
Capital Medical University, China

*CORRESPONDENCE
Grigoriy A. Armeev,
armeev@aintbio.org
Alexey K. Shaytan,
shaytan_ak@mail.bio.msu.ru

[†]These authors have contributed equally to this work and share first authorship

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

RECEIVED 14 October 2022
ACCEPTED 09 November 2022
PUBLISHED 24 November 2022

CITATION
Armeev GA, Gribkova AK and Shaytan AK (2022), Nucleosomes and their complexes in the cryoEM era: Trends and limitations. *Front. Mol. Biosci.* 9:1070489. doi: 10.3389/fmolb.2022.1070489

COPYRIGHT
© 2022 Armeev, Gribkova and Shaytan. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

Nucleosomes and their complexes in the cryoEM era: Trends and limitations

Grigoriy A. Armeev^{1*†}, Anna K. Gribkova^{1†} and Alexey K. Shaytan^{1,2*}

¹Department of Biology, Lomonosov Moscow State University, Moscow, Russia, ²Department of Computer Science, HSE University, Moscow, Russia

Twenty-five years have passed since the appearance of the first atomistic model of the nucleosome structure, and since then the number of new structures has gradually increased. With the advent of cryo-microscopy, the rate of accumulation of models has increased significantly. New structures are emerging with different histone variants and a variety of proteins that bind to nucleosomes. At the moment, there are more than four hundred structures containing nucleosomes in the Protein Data Bank. Many of these structures represent similar complexes, others differ in composition, conformation and quality. In this perspective, we investigate the diversity of known nucleosome structures, analyze data and model quality, variations in histone/DNA content of nucleosomes and spectrum of their interactors. We outline those parts of the nucleosome “structurome” that are already explored and those awaiting further exploration.

KEYWORDS

chromatin proteins, protein-protein interactions (PPI), structural datasets, nucleosome structure, chromatin structure

Introduction

Chromatin combines a high degree of compaction with the ability to access the DNA for transcription. Chromatin is subject to constant rearrangements and does not have a permanent structure. At the basic level of packaging, eukaryotic DNA forms complexes with proteins - nucleosomes. Each nucleosome carries about 200 DNA b.p. only 145–147 b.p. of which are associated with proteins constituting nucleosome core particle (NCP), the remaining 50 bp are called linkers, as they connect neighboring NCPs (Figure 1D). The protein core of the nucleosome comprises four types of histones, which in pairs form histone dimers H3-H4 and H2A-H2B. A complete nucleosome contains two dimers of each type, however, histones can form smaller subnucleosomal particles (Zlatanova et al., 2009; Armeev et al., 2019). There is a separate type of histone, the linker histone also called H1 histone (or H5 in avian species), which binds to the nucleosome to form a chromatosome. Chromatosomes are more stable than nucleosomes and allow chromatin to be packed more tightly (Fan et al., 2003). Histone genes are divided into replication-dependent (canonical histones) and replication independent

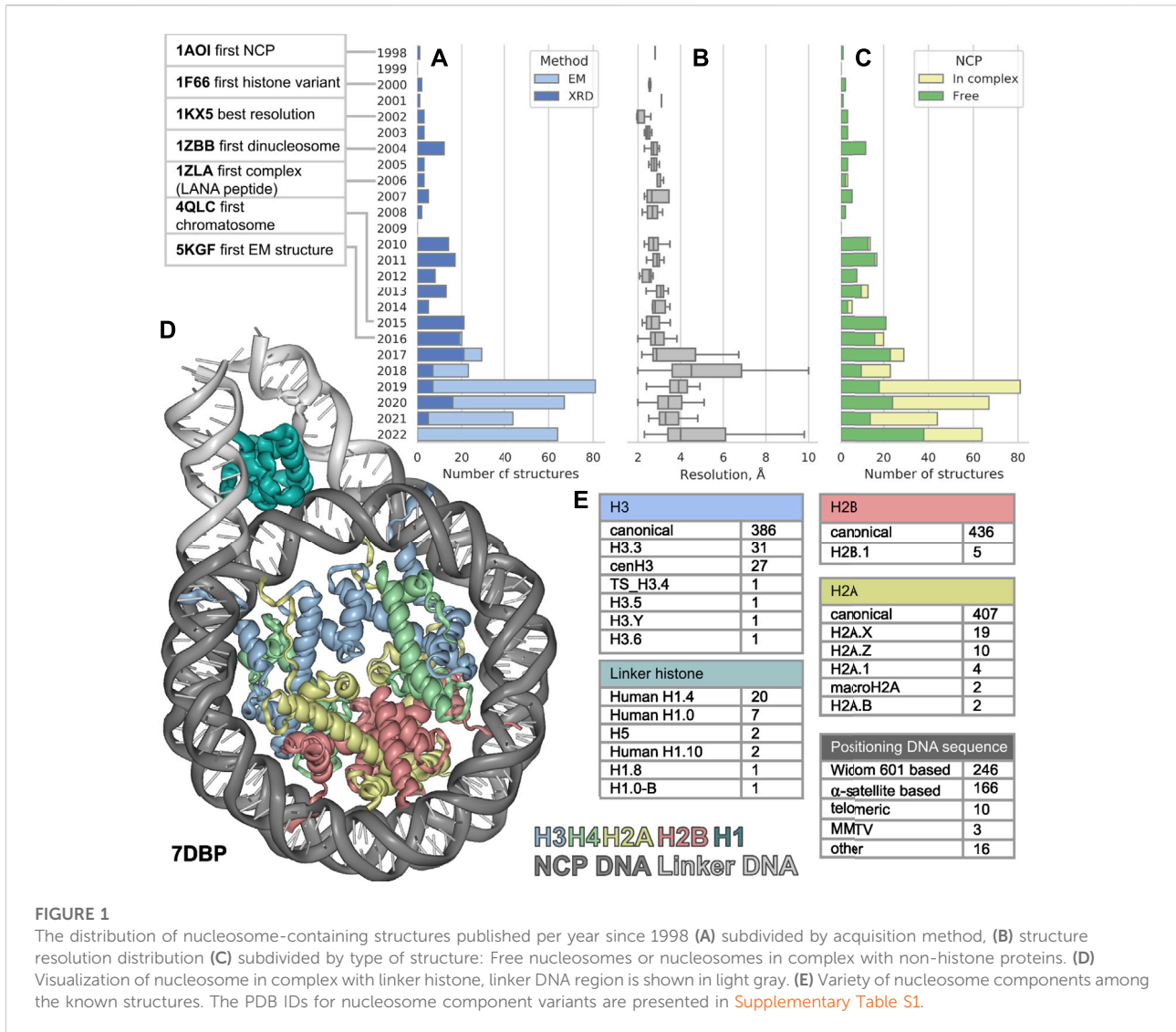


FIGURE 1

The distribution of nucleosome-containing structures published per year since 1998 (A) subdivided by acquisition method, (B) structure resolution distribution (C) subdivided by type of structure: Free nucleosomes or nucleosomes in complex with non-histone proteins. (D) Visualization of nucleosome in complex with linker histone, linker DNA region is shown in light gray. (E) Variety of nucleosome components among the known structures. The PDB IDs for nucleosome component variants are presented in [Supplementary Table S1](#).

(histone variants), which may be involved in specific nuclear processes (Seal et al., 2022). Nucleosomes are also the target for an extensive family of non-histone chromatin proteins. Such proteins specifically interact with nucleosomes to perform a wide range of functions, from kinetochore formation during DNA replication to regulation of transcription and DNA repair.

The first atomic NCP structure was obtained in 1997 (deposited to PDB in 1998) by X-ray diffraction (XRD) (Luger, 1997). In the past 25 years after the first one, 441 new nucleosome containing structures were released (as of 20 September 2022), nearly half of them were released in the last 5 years. Such burst could be explained by recent rapid progress in cryo-electron microscopy (cryoEM). Numerous structures of nucleosomes with histone variants have appeared, as well as structures of complexes with other chromatin proteins. However, this flow of structures makes it

difficult to keep pace and see all available information about chromatin structure at its basis. In this work, we examine a set of all currently available structures of nucleosomes and their complexes. We evaluate the diversity of nucleosome components and proteins associated with them.

Resolution and quality of known nucleosome structures

Since the first structure released in 1997 until 2015 all 113 structures of nucleosomes in PDB were obtained by XRD. The yearly median resolution of those structures was better than 3 Å, and is sufficient for protein backbone and large side chains position determination. The best structure in terms of resolution was released in 2002 (Davey et al., 2002). The structure presented

a complete atomic picture of a nucleosome assembled on a modified human α -satellite DNA sequence, together with bound water molecules and partially resolved histone tails. In 2016 the first structure of nucleosome obtained by cryoEM was published with resolution of 4.5 Å (Wilson et al., 2016, 1) and ever since 253 structures were obtained with cryoEM and 188—with XRD (Figure 1A). As a result, the per-year median resolution of nucleosome structures has decreased significantly, sometimes worse than 4 Å. (Figure 1B). However, the criteria for determining resolution may be different between cryoEM structures. The resolution in cryoEM structures is unevenly distributed and could be significantly worse in the regions of interest. Despite a general decrease in the resolution of structures, the quality of the models is growing, judging by the clash-score, rotamer-outliers (Supplementary Figures S1A,B). Such a trend suggests that the models might be biased towards optimal molecular geometry rather than experimental structure. The local resolution is usually lower in highly mobile regions, such as DNA in nucleosomes, flexible histone tails and interacting proteins. CryoEM resolution of nucleosome-containing structures is steadily improving (Supplementary Figure S1C), reaching the XRD level of detail for some structures. We expect an improvement of nucleosome structure resolutions in the future by establishing consistent protocols as the method matures. However, for EM structures it is always important to assess the local resolution of electron density maps to separate the confident model areas from vaguely defined zones.

Variety of DNA and histones in nucleosome structures

It is known that the structure of nucleosome depends on the DNA sequence. There are three main groups of nucleosomes containing such sequences present in PDB. The largest group contains “Widom 601” sequence (Lowary and Widom, 1998) and its derivatives—246 structures (this is an artificially created high affinity nucleosome positioning sequence). The second sequence group is α -satellite sequences and their derivatives—166 structures (α -satellite sequence is a natural sequence from human centromeric chromosome region). Most of the nucleosome structures in this group contain 146 bp of DNA (90 structures), 26 structures contain shortened 145 bp DNA α -satellite sequence and 26 structures contain longer 147 bp DNA sequence. The third sequence group contains 10 structures with telomeric sequence and several other sequences (cen3 DNA, MMTV, DNA1, sat2 and Human-D02, see Supplementary Table S1). Despite the fact that hundreds of different nucleosome structures are stored in PDB, their DNA sequences belong to a rather limited set. It is clear that increase in the variety of different DNA types in nucleosome structures is vital for understanding the nucleosome positioning. The median resolution across all nucleosomal structures is 3.3 Å, since DNA in nucleosomes is more mobile than histones the effective resolution

in DNA areas is lower. Thus, half of nucleosome structures likely inherit DNA geometry from other high-resolution structures used as templates.

Histone variants provide some specific functions, for example participate in various biological processes such as gene expression, DNA replication and DNA damage repair. Even a few amino acid substitutions in histone variants may alter nucleosome dynamics and stability (Dai et al., 2021). Therefore, the NCP structures with different histone sequences may shed light on the physical nature of the specific biological processes in which histone variants are involved.

We employed histone variant annotation from HistoneDB 2.0 (Draizen et al., 2016) to assess the distribution of histone variants in NCPs across the PDB (Figure 1E). H3 is the most represented histone type in nucleosome structures; PDB contains the following variants: H3.3, cenH3, TS H3.4, H3.5, H3.6, H3.Y. There are structures with H2A variants H2A.1, H2A.B, H2A.X, H2A.Z and macroH2A. However, macroH2A is present in truncated form (around 107 out of 370 amino acids). Variants H2A.L, H2A.P, H2A.W are still missing in PDB. The latter variant H2A.W is a plant specific histone with a unique SPKK motif in the C-terminal tail, whose implications on nucleosome structure are yet to be characterized. There is only one H2B variant present in PDB-deposited nucleosome structures - H2B.1, while other histone variants, which are involved in spermiogenesis (H2B.W, sperm H2B and subH2B) are not available.

The effect of histone variants on the nucleosome structure and dynamics is one of the major challenges in structural studies. However, even canonical histones vary between species. Most of the nucleosomes contain human and model organisms histones: *Homo sapiens* (219 structures), *Xenopus laevis* (172), *Saccharomyces cerevisiae* (12), *Drosophila melanogaster* (8), *Mus musculus* (7). But there are structures of nucleosomes with histones of methylotrophic yeast *Komagataella pastoris* which are used for protein production (PDB ID: 7WLR) (Fukushima et al., 2022), nucleosome of pathogenic unicellular eukaryotic parasite *Giardia lamblia* (PDB ID: 7D69) (Sato et al., 2021) and human nucleosome with incorporation of protozoan parasite *Leishmania* histone H3 (PDB ID: 6KXV) (Dacher et al., 2019). The first studies of archaea histones date to 1990 (Sandman et al., 1990), however the structure of histone-based archaeal chromatin was resolved only recently. There is only one structure of archaeal so-called “hypernucleosomes” from *Methanothermobacter feravidus* (PDB ID: 5T5K) (Mattioli et al., 2017), where DNA is wrapped around an “endless” histone-protein core. Interestingly, several viruses encode their own histones, while other type viruses can even incorporate host eukaryotic histones into their nucleosomes. The first examples of viral nucleosomes of Marseilleviridae (giant viruses that infect amoebae) were resolved in 2021 (PDB IDs: 7LV8, 7LV9, 7N8N) (Liu et al., 2021a; Valencia-Sánchez et al., 2021). Marseilleviridae histones are present in doublets (H4 is fused to H3 and H2B is

fused to H2A) and form particles similar to eukaryotic “canonical” nucleosomes, though less stable and contain only 120 bp of DNA.

Higher order chromatin structures

The linker histone H1 binds to the nucleosome near the DNA entry-exit point, thus forming a chromatosome. Currently, there are 33 structures containing linker histone, predominantly human H1.4 variant (20 structures). Also, nucleosome structures contain human H1.0 and H1.10 variants, chicken H.5 and one structure from *Xenopus* with H1.0 and H1.8. Variants scH1, TS H1.6, TS H1.7, TS H1.9 are not present in structure databases now (Figure 1C).

The geometry of nucleosome fibers changes in the presence of linker histone (Routh et al., 2008). The PDB contains 24 models containing parts of such fibers, both with and without a linker histone. Here, we assessed the available models (bioassemblies) and counted nucleosomes located on continuous double-stranded DNA segments. The majority of supranucleosomal structures contain two nucleosomes (11 structures), 8 structures contain three nucleosomes and 5 structures contain four nucleosomes and one structure contain six nucleosomes (PDB ID: 6HKT). Even larger fibers containing 12 and 24 nucleosomes were studied (Song et al., 2014), unfortunately the models were not deposited to the PDB. The studies of larger fibers are of great importance, though much more complex due to high flexibility of their structure.

Nucleosome complexes

A plethora of different non-histone proteins interact with nucleosomes to regulate nuclear processes. It is important to study the mechanism of such interactions, as they provide an insight into the fundamental basis of genome regulation. There are 193 structures that contain NCPs in complex with other proteins. The first six structures were resolved by XRD in the period from 2006 to 2013. The interactors resolved in the first nucleosome complexes were relatively small peptides or proteins - a fragment of Kaposi’s sarcoma herpesvirus LANA protein (PDB ID: 1ZLA), regulator of chromosome condensation factor (RCC1, PDB ID: 3MVD), regulatory protein SIR3 (PDB IDs: 3TU4, 4JJN, 4KUD, 4LD9). Since 2016 the number of complexes has rapidly increased, most of them are obtained with cryoEM but have relatively low resolution (Supplementary Figure S1D). The years 2019 and 2020 were the most fruitful in terms of published nucleosome complexes - 60 and 38, respectively (versus 21 and 29 for single nucleosomes). Molecular weight of NCP interactors spans from 0.4 to 687 kDa with 53 kDa median. The heaviest complex is the histone acetyltransferase NuA4 complex, shown in Figure 2 (PDB ID: 7VVZ).

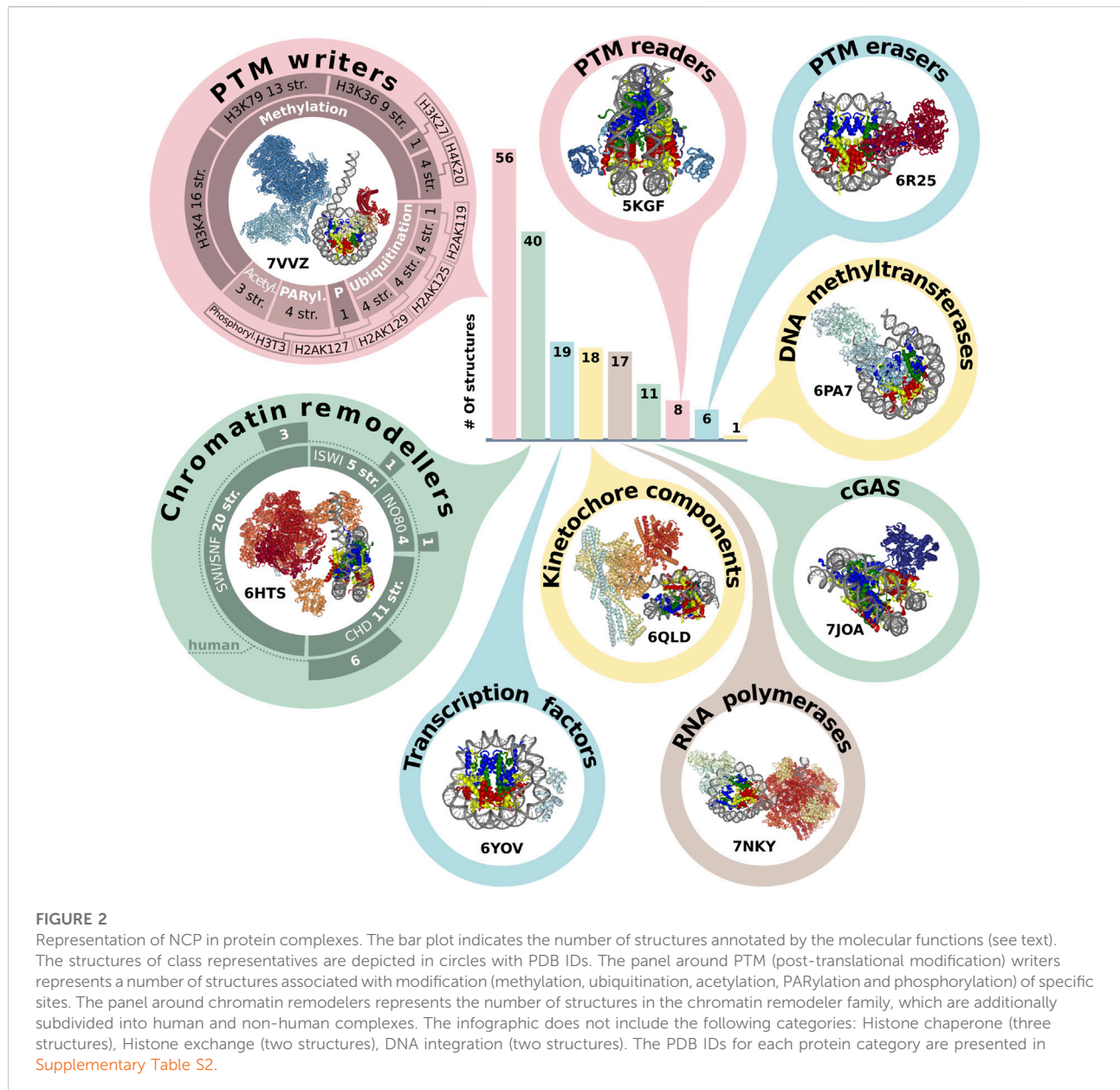
There are 215 structural non-histone nucleosome interacting partners found in PDB. The majority of proteins belong to *Saccharomyces cerevisiae* proteins (96 proteins in 54 structures) and human proteins (78 proteins in 102 structures). Interestingly, there are only 15 nucleosomes containing *Saccharomyces* histones. The half of interactor proteins (94 out of 215) are present only in a single structure, and 89 are present in two or more structures. Not all protein sequences are resolved in the structures, but the length of the resolved protein fragments steadily increases (Supplementary Figure S2A).

To assess the functional diversity of NCP complexes with non-histone interactors we designed a simplified protein classification scheme [based on UniProt protein annotation (The UniProt Consortium, 2021) and Gene ontology (Ashburner et al., 2000; The Gene Ontology Consortium, 2021)]. This scheme describes Biological Process, Molecular function and association of structure with histone post-translational modifications (PTM). According to this scheme, the majority of NCP complexes with non-histone interactors belong to the following biological processes: Transcription associated (112 structures), DNA repair (34) and Cell division (24), (Supplementary Figure S2B).

There are five most represented molecular functional protein groups in PDB by our classification: PTM writers (56 structures), Chromatin remodelers (40), Transcription factors (19), RNA polymerases (17) and complexes with Cyclic GMP-AMP synthase (guanosine monophosphate-adenosine monophosphate synthase) (noted as cGAS) (11) (Figure 2). There are 17 structures of NCP with RNA polymerase. However, there are no human RNA polymerases, even in complexes with human histones. All interactors in this group are from *Saccharomyces cerevisiae* or *Komagataella phaffii*.

We also analyzed the representation of human chromatin proteins of different functional classes in the structures of the complexes, Supplementary Figure S2C. The most structurally studied class is kinetochore components (70%). In other classes (Chromatin remodelers, PTM erasers, PTM readers, PTM writers) the range of coverage is 2%–18%.

One of the epigenetics regulation layers is histone post translational modifications, which can alter nucleosome structural dynamics and regulate DNA accessibility (Bernier et al., 2015; Brehove et al., 2015; Kim et al., 2015). We analyzed NCP complexes’ structures for association with histone PTMs and revealed that one-third of NCP complexes is associated with methylation and one-seventh with ubiquitination, Figure 2. Human most studied methylation sites include H3K4, H3K9, H3K27, H3K36, H3K79, H4K20, H3R8, H4R3 [annotated at HISTome 2 (Shah et al., 2020)]. The largest number of structures is associated with methylation of H3K4 (16 structures) predominantly by COMPASS complex, this mark indicates active gene promoters. H3K79 methylation—one of marks of transcription



regulation and DNA damage response is presented in 13 structures and the mark of active chromatin H3K36—in 9 structures. Currently, 8 out of 30 human histone lysine methyltransferases are in the structures of NCP complexes and there are no structures of NCP in complex with human methyltransferases for H3K9 and H4R3 sites. First structures with acetylation machinery appeared in 2022 (PDBs: 7W9V, 7VVZ, 7VVU). Less than 10 structures are linked with writing or reading PTMs from the following sites: phosphorylation of H3T3; methylation of H3K27, H4K20; demethylation of H3K4, H3K9 and ubiquitination of H2AK15, H2A119, H2AK127, H2AK125, H2AK129, H2BK120 [Figure 2](#), [Supplementary Table S2](#).

Another key functional class of chromatin complexes is chromatin remodellers, which modify the composition or location of nucleosomes using the energy from ATP hydrolysis. The resolution revolution in cryoEM contributed to the emergence of chromatin remodeler typo complexes in 2017. Chromatin remodellers are divided into four families, the most represented in complexes with nucleosomes is SWI/SNF family (29 structures), followed by CHD (13), INO80 (6) and ISWI (5) ([Supplementary Table S2](#)). Thus, the discovered structures cover all known remodeler families, and even if the resolution is low, it is possible to obtain some mechanistic insights about protein-DNA interactions. The next stage of nucleosome complexes research might be expanding the range of proteins interacting

with nucleosomes to other DNA interacting proteins. For example, there are currently no structural details about the interplay between nucleosome and CRISPR-Cas systems, which are potential candidates for novel therapy approaches and detection systems (Yarrington et al., 2018; Novikov et al., 2021).

Nucleosome complexes in human pathologies

Nucleosomes play a pivotal role in chromatin maintenance and gene expression regulation. There are well-studied histone mutations which lead to cancers (Nacev et al., 2019; Espiritu et al., 2021). The part of these histone mutations is represented in NCP structures: 23 out of 434 nucleosome structures contain the following histone mutations: Lysine to Glutamine (KQ) Substitutions in the H3 and H4 Histone-Fold Domains (Iwasaki et al., 2011), oncohistone mutations H3.3K36M (yeast and human nucleosomes) (Liu et al., 2021b), H3 and H4 mutated residues located at the protein–DNA interfaces flanking the nucleosomal dyad (Muthurajan et al., 2004).

We analyzed non-histone nucleosome structural interactors that are involved in oncogenesis by OncoKB annotation (Zdobnov et al., 2021). To increase the number of proteins, we also considered human ortholog proteins for non-human proteins from OrthoDB (Zdobnov et al., 2021). As a result, 24 out of 105 human nucleosome interacting genes are annotated in OncoKB (Chakravarty et al., 2017): five oncogenes (EZH2, NSD2, SOX2, DOT1L, SMARCE1) and 18 Tumor Suppressor Gene (ARID1A, BRCA1, DNMT3A, EP300, EZH2, PBRM1, SETD2, SMARCA4, SMARCB1, ARID2, BARD1, KMT2A, KMT2C, SUZ12, PARP1, DNMT3B, SMARCE1, TP53BP1).

Design of novel compounds for epigenetic targets requires detailed structural information of the interaction interface. Meanwhile, there are only 4 non-histone nucleosome interacting proteins which are approved as targets for epigenetic drugs: Poly [ADP-ribose] polymerase 1 (P09874), Poly [ADP-ribose] polymerase 2 (Q9UGN5), DNA (cytosine-5)-methyltransferase 3A (Q9Y6K1), Histone-lysine N-methyltransferase EZH2 (Q15910). To date, one molecule is in clinical trials targeting Lysine-specific histone demethylase 1A (O60341) as provided by ChEMBL database.

Discussion

Histone proteins are among the most conservative proteins, and indeed from the first glance all nucleosomes in PDB look alike. However, nucleosomes differ in histone variants, PTMs DNA sequences and there are chromatin proteins that alter their structure.

Sequence-dependence of nucleosome assembly still remains a subject of debate. Unfortunately, the variety of PDB-deposited nucleosomal DNA sequences is limited. Moreover, the majority of nucleosome models do not have sufficient resolution and the

structure of DNA in such models is usually derived from other high-resolution structures. There is a strong need for systematic research of DNA sequence influence on nucleosome structure and positioning. We believe that such studies currently are limited by cryoEM data analysis pipelines, which limit the ability to reconstruct relatively flexible DNA.

Another under-researched field of nucleosome structural biology is that dealing with the effects of PTMs on nucleosome structure. Given the amount of different PTMs and their combinations, such effects could be feasibly studied *in silico via* full atom molecular dynamics simulation. Besides, the majority of PTMs are located in intrinsically disordered histone tails, thus making computational methods even more suitable for this task.

Large complexes of nucleosomes and parts of chromatin fibers appear in PDB. The size of structures grows together with their conformational flexibility and heterogeneity, making it difficult to apply structural biology methods. XRD structures are influenced by crystal packing effects; cryoEM is not yet capable of reconstructing the structure ensembles of highly flexible complexes. Nevertheless, there are already structures of large complexes that reveal functional roles of nucleosomes, such as nucleosome rearrangements during the RNA polymerase II passage (Kujirai et al., 2018). In order to explore such structures, it is necessary to combine structural biology methods and integrative modeling approaches based on different heterogeneous experimental datasets (Webb et al., 2018).

Despite a certain number of resolved structures of NCP in complexes with other proteins (193 structures), the set of known interactor structures is incomplete. Nonetheless, there are resolved complexes of nucleosomes with all main chromatin protein categories and even all chromatin remodeler families. The latter structures provided crucial molecular details of remodelers action and an interplay between DNA and histone octamer configuration during nucleosome remodeling (Farnung et al., 2017; Liu et al., 2017). However, the number of structures varies considerably between chromatin protein categories and the representation of functional categories is very heterogeneous. Among the PTM writers the majority of resolved nucleosome complex structures are associated with methylation, while acetylation plays an equally important role in the regulation of nuclear processes. The first structures with acetyltransferases were resolved this year. There are NCPs in complex with methyltransferases for extensively studied lysine methylation, however there are no structures with arginine methyltransferases. There is a plethora of possible PTMs including some emerging noncanonical histone modifications, which also need to be studied to reveal a specific mechanism of their functioning.

In this perspective, we considered the diversity of nucleosome structures and their complexes. Nucleosome structural studies are useful not only for fundamental science, but also for a number of applications. For therapeutic applications, non-histone interactors of nucleosomes can serve as specific drug targets, particularly for treatment of cancer. Histone PTMs (Chervona and Costa, 2012; Zhao and Shilatfard, 2019; Lee and Kim, 2022), as well as free

nucleosomes in tissues (Fedyuk et al., 2022) can be used as prognostic and predictive biomarkers in oncology diagnostics. On the other hand, it is necessary to study plant nucleosomes, especially those of cultivated plant species. To sum up, there is a need to expand our knowledge of molecular epigenetic processes by building a comprehensive map of nucleosome structural interactions with non-histone partners for fundamental science, medicine and agriculture.

Data availability statement

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

Author contributions

GA, AG, and AS contributed to conception of the work. GA performed data extraction, structural analysis and DNA sequences analysis, AG performed complexes' classification and analysis of histone variants. AS contributed to analysis and conceptualization. GA and AG wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This research was funded by Russian Science Foundation grant #21-74-00033 (Data harvesting and structural annotation).

References

- Armeev, G. A., Gribkova, A. K., Pospelova, I., Komarova, G. A., and Shaytan, A. K. (2019). Linking chromatin composition and structural dynamics at the nucleosome level. *Curr. Opin. Struct. Biol.* 56, 46–55. doi:10.1016/j.sbi.2018.11.006
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: Tool for the unification of biology. The gene ontology Consortium. *Nat. Genet.* 25, 25–29. doi:10.1038/75556
- Bernier, M., Luo, Y., Nwokelo, K. C., Goodwin, M., Dreher, S. J., Zhang, P., et al. (2015). Linker histone H1 and H3K56 acetylation are antagonistic regulators of nucleosome dynamics. *Nat. Commun.* 6, 10152. doi:10.1038/ncomms10152
- Brehove, M., Wang, T., North, J., Luo, Y., Dreher, S. J., Shimko, J. C., et al. (2015). Histone core phosphorylation regulates DNA accessibility. *J. Biol. Chem.* 290, 22612–22621. doi:10.1074/jbc.M115.661363
- Chakravarty, D., Gao, J., Phillips, S., Kundra, R., Zhang, H., Wang, J., et al. (2017). OncoKB: A precision oncology knowledge base. *JCO Precis. Oncol.* 2017, 1–16. doi:10.1200/PO.17.00011
- Chervona, Y., and Costa, M. (2012). Histone modifications and cancer: Biomarkers of prognosis? *Am. J. Cancer Res.* 2, 589–597.
- Dacher, M., Tachiwana, H., Horikoshi, N., Kujirai, T., Taguchi, H., Kimura, H., et al. (2019). Incorporation and influence of Leishmania histone H3 in chromatin. *Nucleic Acids Res.* 47, 11637–11648. doi:10.1093/nar/gkz1040
- Dai, L., Xiao, X., Pan, L., Shi, L., Xu, N., Zhang, Z., et al. (2021). Recognition of the inherently unstable H2A nucleosome by Swc2 is a major determinant for unidirectional H2A.Z exchange. *Cell Rep.* 35, 109183. doi:10.1016/j.celrep.2021.109183
- Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W., and Richmond, T. J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* 319, 1097–1113. doi:10.1016/S0022-2836(02)00386-8
- Drazen, E. J., Shaytan, A. K., Mariño-Ramírez, L., Talbert, P. B., Landsman, D., and Panchenko, A. R. (2016). HistoneDB 2.0: A histone database with variants—an integrated resource to explore histones and their variants. *Database.* 2016, baw014. doi:10.1093/database/baw014
- Espirito, D., Gribkova, A. K., Gupta, S., Shaytan, A. K., and Panchenko, A. R. (2021). Molecular mechanisms of oncogenesis through the lens of nucleosomes and histones. *J. Phys. Chem. B* 125, 3963–3976. doi:10.1021/acs.jpbc.1c00694
- Fan, Y., Nikitina, T., Morin-Kensicki, E. M., Zhao, J., Magnuson, T. R., Woodcock, C. L., et al. (2003). H1 linker histones are essential for mouse development and affect nucleosome spacing *in vivo*. *Mol. Cell. Biol.* 23, 4559–4572. doi:10.1128/MCB.23.13.4559-4572.2003
- Farnung, L., Vos, S. M., Wigge, C., and Cramer, P. (2017). Nucleosome–Chd1 structure and implications for chromatin remodelling. *Nature* 550, 539–542. doi:10.1038/nature24046

Acknowledgments

This work used the infrastructure provided by the Interdisciplinary Scientific and Educational School of Moscow University “Molecular Technologies of the Living Systems and Synthetic Biology.” The work of AG was supported by Non-commercial Foundation for Support of Science and Education “INTELLECT”.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2022.1070489/full#supplementary-material>

- Fedyuk, V., Erez, N., Furth, N., Beresh, O., Andreishcheva, E., Shinde, A., et al. (2022). Multiplexed, single-molecule, epigenetic analysis of plasma-isolated nucleosomes for cancer diagnostics. *Nat. Biotechnol.* 1. doi:10.1038/s41587-022-01447-3
- Fukushima, Y., Hatazawa, S., Hirai, S., Kujirai, T., Ehara, H., Sekine, S., et al. (2022). Structural and biochemical analyses of the nucleosome containing *Komagataella pastoris* histones. *J. Biochem. (Tokyo)* 172, 79–88. doi:10.1093/jb/mvac043
- Iwasaki, W., Tachiwana, H., Kawaguchi, K., Shibata, T., Kagawa, W., and Kurumizaka, H. (2011). Comprehensive structural analysis of mutant nucleosomes containing lysine to glutamine (KQ) substitutions in the H3 and H4 Histone-Fold Domains. *Biochemistry* 50, 7822–7832. doi:10.1021/bi201021h
- Kim, J., Lee, J., and Lee, T.-H. (2015). Lysine acetylation facilitates spontaneous DNA dynamics in the nucleosome. *J. Phys. Chem. B* 119, 15001–15005. doi:10.1021/acs.jpcc.5b09734
- Kujirai, T., Ehara, H., Fujino, Y., Shirouzu, M., Sekine, S., and Kurumizaka, H. (2018). Structural basis of the nucleosome transition during RNA polymerase II passage. *Science* 362, 595–598. doi:10.1126/science.aau9904
- Lee, J. E., and Kim, M.-Y. (2022). Cancer epigenetics: Past, present and future. *Semin. Cancer Biol.* 83, 4–14. doi:10.1016/j.semcancer.2021.03.025
- Liu, X., Li, M., Xia, X., Li, X., and Chen, Z. (2017). Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. *Nature* 544, 440–445. doi:10.1038/nature22036
- Liu, Y., Bisio, H., Toner, C. M., Jeudy, S., Philippe, N., Zhou, K., et al. (2021a). Virus-encoded histone doublets are essential and form nucleosome-like structures. *Cell* 184, 4237–4250.e19. doi:10.1016/j.cell.2021.06.032
- Liu, Y., Zhang, Y., Xue, H., Cao, M., Bai, G., Mu, Z., et al. (2021b). Cryo-EM structure of SETD2/Set2 methyltransferase bound to a nucleosome containing oncohistone mutations. *Cell Discov.* 7, 32–12. doi:10.1038/s41421-021-00261-6
- Lowary, P. T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* 276, 19–42. doi:10.1006/jmbi.1997.1494
- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251–260. doi:10.1038/38444
- Mattiroli, F., Bhattacharyya, S., Dyer, P. N., White, A. E., Sandman, K., Burkhart, B. W., et al. (2017). Structure of histone-based chromatin in archaea. *Science* 357, 609–612. doi:10.1126/science.aaj1849
- Muthurajan, U. M., Bao, Y., Forsberg, L. J., Edayathumangalam, R. S., Dyer, P. N., White, C. L., et al. (2004). Crystal structures of histone Sin mutant nucleosomes reveal altered protein–DNA interactions. *EMBO J.* 23, 260–271. doi:10.1038/sj.emboj.7600046
- Nacev, B. A., Feng, L., Bagert, J. D., Lemiesz, A. E., Gao, J., Soshnev, A. A., et al. (2019). The expanding landscape of “oncohistone” mutations in human cancers. *Nature* 567, 473–478. doi:10.1038/s41586-019-1038-1
- Novikov, R. V., Gribkova, A. K., Kacher, J. G., Zaytsev, P. A., Armeev, G. A., Gluhov, G. S., et al. (2021). Design of nucleic acid biosensors based on CRISPR/cas systems and reporter split proteins. *Mosc. Univ. Biol. Sci. Bull.* 76, 52–58. doi:10.3103/S0096392521020036
- Routh, A., Sandin, S., and Rhodes, D. (2008). Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proc. Natl. Acad. Sci. U. S. A.* 105, 8872–8877. doi:10.1073/pnas.0802336105
- Sandman, K., Krzycki, J. A., Dobrinski, B., Lurz, R., and Reeve, J. N. (1990). HMF, a DNA-binding protein isolated from the hyperthermophilic archaeon *Methanothermobacter fervidus*, is most closely related to histones. *Proc. Natl. Acad. Sci. U. S. A.* 87, 5788–5791. doi:10.1073/pnas.87.15.5788
- Sato, S., Takizawa, Y., Hoshikawa, F., Dacher, M., Tanaka, H., Tachiwana, H., et al. (2021). Cryo-EM structure of the nucleosome core particle containing *Giardia lamblia* histones. *Nucleic Acids Res.* 49, 8934–8946. doi:10.1093/nar/gkab644
- Seal, R. L., Denny, P., Bruford, E. A., Gribkova, A. K., Landsman, D., Marzluft, W. F., et al. (2022). A standardized nomenclature for mammalian histone genes. *Epigenetics Chromatin* 15, 34. doi:10.1186/s13072-022-00467-2
- Shah, S. G., Mandloi, T., Kunte, P., Natu, A., Rashid, M., Reddy, D., et al. (2020). HISTome2: A database of histone proteins, modifiers for multiple organisms and epidrugs. *Epigenetics Chromatin* 13, 31. doi:10.1186/s13072-020-00354-8
- Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., et al. (2014). Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science* 344, 376–380. doi:10.1126/science.1251413
- The Gene Ontology Consortium (2021). The gene ontology resource: Enriching a GOld mine. *Nucleic Acids Res.* 49, D325–D334. doi:10.1093/nar/gkaa1113
- The UniProt Consortium (2021). UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* 49, D480–D489. doi:10.1093/nar/gkaa1100
- Valencia-Sánchez, M. I., Abini-Agbomson, S., Wang, M., Lee, R., Vasilyev, N., Zhang, J., et al. (2021). The structure of a virus-encoded nucleosome. *Nat. Struct. Mol. Biol.* 28, 413–417. doi:10.1038/s41594-021-00585-7
- Webb, B., Viswanath, S., Bonomi, M., Pellarin, R., Greenberg, C. H., Saltzberg, D., et al. (2018). Integrative structure modeling with the integrative modeling platform. *Protein Sci.* 27, 245–258. doi:10.1002/pro.3311
- Wilson, M. D., Benlekhir, S., Fradet-Turcotte, A., Sherker, A., Julien, J.-P., McEwan, A., et al. (2016). The structural basis of modified nucleosome recognition by 53BP1. *Nature* 536, 100–103. doi:10.1038/nature18951
- Yarrington, R. M., Verma, S., Schwartz, S., Trautman, J. K., and Carroll, D. (2018). Nucleosomes inhibit target cleavage by CRISPR-Cas9 *in vivo*. *Proc. Natl. Acad. Sci. U. S. A.* 115, 9351–9358. doi:10.1073/pnas.1810062115
- Zdobnov, E. M., Kuznetsov, D., Tegenfeldt, F., Manni, M., Berkeley, M., and Kriventseva, E. V. (2021). OrthoDB in 2020: Evolutionary and functional annotations of orthologs. *Nucleic Acids Res.* 49, D389–D393. doi:10.1093/nar/gkaa1009
- Zhao, Z., and Shilatifard, A. (2019). Epigenetic modifications of histones in cancer. *Genome Biol.* 20, 245. doi:10.1186/s13059-019-1870-5
- Zlatanova, J., Bishop, T. C., Victor, J. M., Jackson, V., and van Holde, K. (2009). The nucleosome family: Dynamic and growing. *Structure* 17, 160–171. doi:10.1016/j.str.2008.12.016