

Conformational and Interaction Landscape of Histone H4 Tails in Nucleosomes Probed by Paramagnetic NMR Spectroscopy

Wenjun Sun,[#] Olga O. Lebedenko,[#] Nicole Gonzalez Salguero,[#] Matthew D. Shannon, Mohamad Zandian, Michael G. Poirier, Nikolai R. Skrynnikov,^{*} and Christopher P. Jaroniec^{*}



Cite This: *J. Am. Chem. Soc.* 2023, 145, 25478–25485



Read Online

ACCESS |



Metrics & More

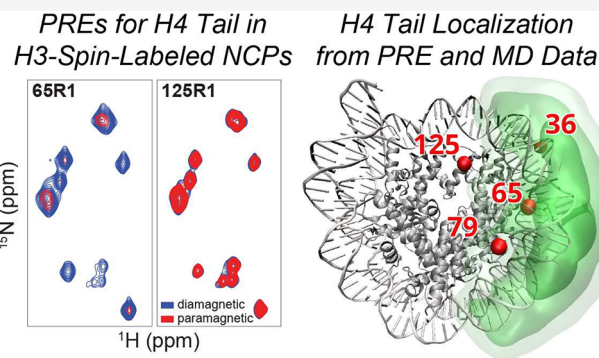


Article Recommendations



Supporting Information

ABSTRACT: The fundamental repeat unit of chromatin, the nucleosome, consists of approximately 147 base pairs of double-stranded DNA and a histone protein octamer containing two copies each of histones H2A, H2B, H3, and H4. Each histone possesses a dynamically disordered N-terminal tail domain, and it is well-established that the tails of histones H3 and H4 play key roles in chromatin compaction and regulation. Here we investigate the conformational ensemble and interactions of the H4 tail in nucleosomes by means of solution NMR measurements of paramagnetic relaxation enhancements (PREs) in recombinant samples reconstituted with ¹⁵N-enriched H4 and nitroxide spin-label tagged H3. The experimental PREs, which report on the proximities of individual H4 tail residues to the different H3 spin-label sites, are interpreted by using microsecond time-scale molecular dynamics simulations of the nucleosome core particle. Collectively, these data enable improved localization of histone H4 tails in nucleosomes and support the notion that H4 tails engage in a fuzzy complex interaction with nucleosomal DNA.



INTRODUCTION

Chromatin is a dynamic protein–DNA complex that acts to compact eukaryotic genomes and regulate DNA accessibility. The basic repeat unit of chromatin, the nucleosome core particle (NCP), contains ~147 base pairs (bps) of a DNA double helix wrapped ~1.7 times around a histone octamer protein complex composed of two copies each of histones H2A, H2B, H3, and H4. Each histone contains a structured globular core domain and a positively charged ~15–35 amino acid (aa) residue N-terminal tail domain. Early experiments using trypsin digestion and nuclear magnetic resonance (NMR) spectroscopy provided initial evidence that histone N-terminal tails are exposed on the outside of nucleosomes and highly flexible,^{1,2} and these insights were confirmed by the first high-resolution X-ray crystal structure of the NCP.³ Indeed, in this and subsequent X-ray (as well as cryoelectron microscopy) NCP structures, electron density is typically missing for extreme N-terminal residues, while for residues further downstream the electron density is usually weak and disjointed. In some cases, certain histone tail fragments can be resolved due to their interactions with neighboring nucleosomes within the crystal lattice, allowing for the tails to be partially reconstructed (e.g., Figure 1C).^{3,4}

The fact that histone tail domains are conformationally disordered and dynamic and project outward into the solvent is key to their functional role, where they provide “mooring

lines” for hundreds of chromatin-associated proteins (CAPs). The latter regulate chromatin assembly and remodeling and serve as accessory proteins to activate or repress transcription.⁵ Notably, histone tails are also processed by histone-modifying enzymes (HMEs), resulting in the addition or removal of various post-translational modifications (PTMs). Among the most common PTMs are lysine and arginine methylation as well as (charge-altering) lysine acetylation.⁶ The pattern of these and other modifications, termed the “histone code”,⁷ serves to tune the affinity of histone tails for different CAPs as well as their interactions with proximal nucleosomes in chromatin.⁸ In summary, histone tails are a focal point of an extremely vast and complex dynamic network, of which we have only a relatively limited understanding at the atomistic level.

In the present study, we pursue detailed characterization of the structural ensemble and interactions of the dynamically disordered histone H4 N-terminal tails in nucleosomes. Previous studies employing solution- and solid-state NMR

Received: September 19, 2023

Revised: October 21, 2023

Accepted: October 23, 2023

Published: November 9, 2023



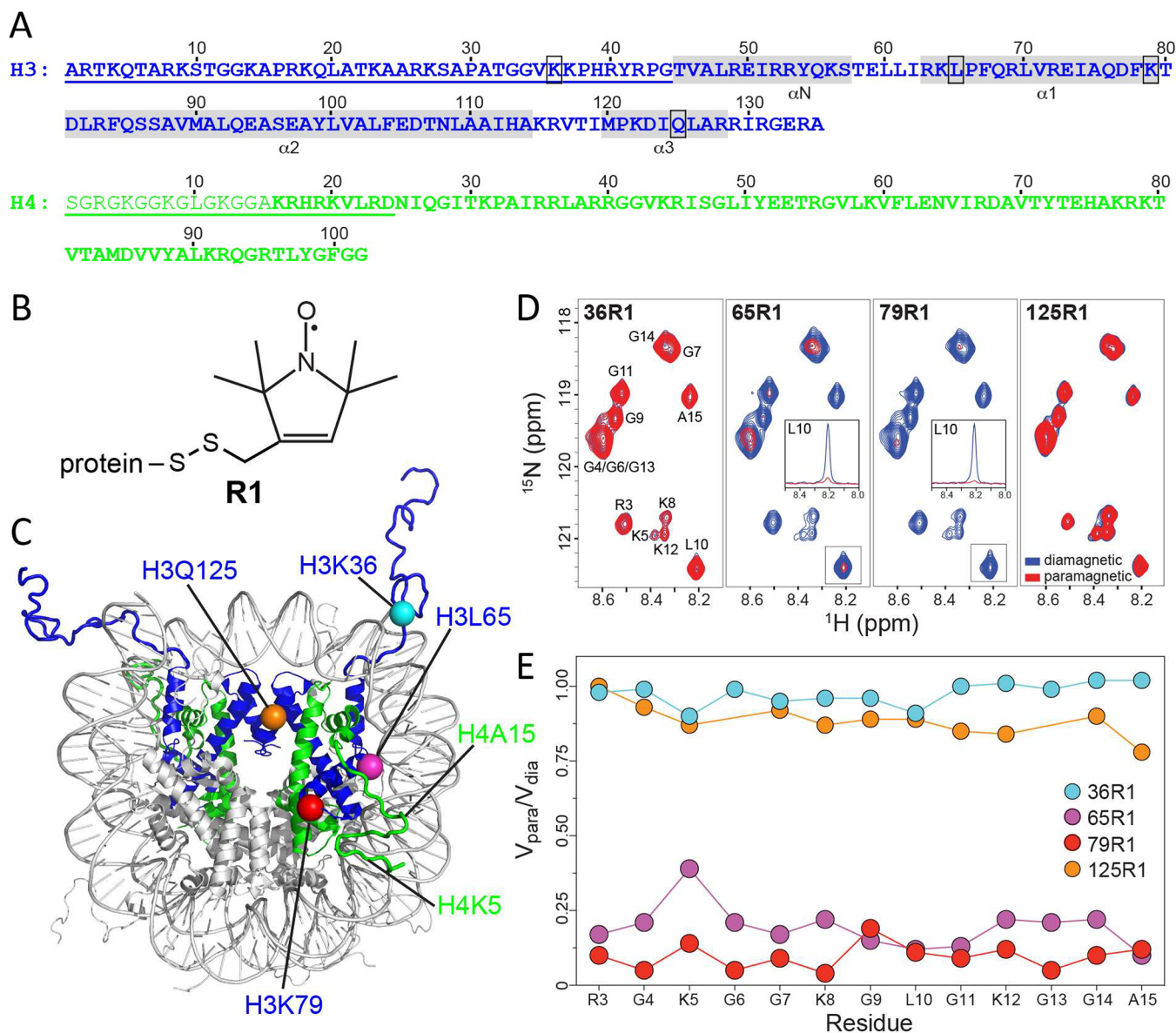


Figure 1. Measurement of paramagnetic relaxation enhancements for H4 N-terminal tail residues in nucleosomes. (A) Amino acid sequences of *Xenopus laevis* histones H3(C110A) and H4. Underlined residues are relatively unstructured and associated with low electron densities and/or high B factors in the NCP X-ray crystal structure (PDB entry 1KX5).⁴ The H3 core helices are indicated by gray rectangles and labeled underneath; the spin-label incorporation sites are denoted by black boxes around the relevant residues. The most flexible H4 tail residues, up to and including A15, are typeset in thin font. (B) Schematic representation of a paramagnetic nitroxide spin-label (MTSL) conjugated to a cysteine residue in a protein. The resulting side-chain is termed R1.^{54,55} (C) Crystal structure of the nucleosome core particle (PDB entry 1KX5) with histone H3 colored blue and histone H4 colored green. The spin-labeling sites in histone H3 (viz., C α atoms of the relevant residues) are indicated by colored spheres and labeled for one of the H3 copies. Locations of representative H4 tail residues are also indicated. (D) ^{15}N - ^1H HSQC NMR spectra of the H4 N-terminal tail in nucleosomes reconstituted with spin-labeled histone H3 and ^{15}N -labeled H4. Shown in each panel is a superposition of paramagnetic (red) and reference diamagnetic (blue) spectra, plotted at the same contour level, for each NCP sample (as identified at the top of the panel); the reference diamagnetic samples were generated by addition of sodium ascorbate to the corresponding paramagnetic nitroxide spin-labeled sample. Resonances for A15 and all glycine residues are aliased in the ^{15}N dimension. Representative one-dimensional ^1H traces are shown in the insets for residue L10 in the 65R1 and 79R1 samples. All NMR data were recorded at 25 °C at ^1H frequency of 850 MHz on NCP samples with concentration of ~ 40 – $45\ \mu\text{M}$ in aqueous buffer solution containing 5 mM Tris, 0.5 mM EDTA, 100 mM NaCl, 0.1 mM MgCl_2 , and 5% D_2O at pH 7.0. (E) Ratios of peak volumes for H4 N-terminal tail residues in HSQC NMR spectra recorded for paramagnetic and diamagnetic NCP samples, V_{para}/V_{dia} used to determine residue-specific PREs (cf. Table S1). Given the high sensitivity of the HSQC NMR spectra, the uncertainties in V_{para}/V_{dia} ratios were 0.01 or smaller for virtually all residues in the different samples studied and within the size of the symbols used to depict the data. For clarity, the error bars were therefore omitted from the plot. The color coding that identifies the four spin-labeled NCP samples is shown in the figure legend and corresponds to that used for spin-labeling sites in panel C.

methods have found that ~ 15 – 20 N-terminal H4 residues are largely dynamically disordered in soluble nucleosomes⁹ and that this dynamic disorder persists in compacted nucleosomes¹⁰ and in highly condensed large nucleosome arrays.¹¹

Due to its high positive charge, the H4 N-terminal tail is expected to generally localize near and interact with negatively

charged nucleosomal DNA (nDNA). However, quantitative measurements of residue-specific NMR spin relaxation rates for H4 tails in nucleosomes coupled with microsecond time-scale molecular dynamics (MD) simulations revealed that electrostatic and hydrogen-bonding interactions between H4 tail residues and nDNA are highly transient and dynamic even at low ionic strength,¹² following the “fuzzy complex” paradigm.^{12,13} While this previous study of H4 tails in nucleosomes yielded valuable insights about the nature of nDNA–H4 tail interactions, the experimental NMR methods used were not suited to directly probe proximities of H4 tail residues to different regions of the nucleosome or the extent to which the H4 N-terminal tail is sequestered by nDNA. These aspects, therefore, remain largely unexplored, although they are potentially important for understanding H4 tail accessibility to CAPs (and HMEs) and its availability to form internucleosomal connections⁸ as well as the recently reported crosstalk between H4 and H3 tails.^{14,15}

As noted above, high-resolution structural techniques, such as X-ray crystallography or cryo-EM, are generally unable to localize the disordered histone tail domains in nucleosomes. Instead, this problem can be uniquely addressed by multidimensional NMR spectroscopy^{9–14,16–27}—in particular, by measurements of residue-specific proton paramagnetic relaxation enhancements (PREs), which report on distances between individual backbone amides and covalent paramagnetic tags attached at specific protein sites and are significant over distances up to ~20–30 Å. For structured biomolecules in solution, the PRE-based restraints can be used to build or refine structural models,^{28–30} and indeed, studies along these lines have been reported for nucleosome particles and their complexes with CAPs.^{31–33} This approach relies on the well-known Solomon–Bloembergen equation³⁴ which represents PRE rates as a product of distance- and dynamics-dependent terms, where the latter can be determined from heteronuclear relaxation measurements³⁵ and the former can be converted into structural restraints. Additionally, methods have been developed to account for the conformational flexibility of paramagnetic tags.^{30,36} In contrast, for dynamically disordered protein domains, such as histone N-terminal tails in nucleosomes, the distance- and dynamics-related variables cannot be readily deconvoluted. In these cases, the measured PRE rates are often interpreted in a qualitative manner using one of the empirical random-coil models as a point of reference.^{37–39} Otherwise, it is common to make certain simplifying assumptions about the dynamics of the disordered peptide chain.^{40–42} Using these assumptions, one can employ PRE data to construct structural models of disordered proteins, typically in a form of multimer conformational ensembles.^{43–45} In addition, specialized methods have been developed to facilitate the extraction of distance information and circumvent the dependence of PREs on dynamics.^{46,47}

In our earlier studies of disordered proteins, we have pointed out that PREs in such systems can be rigorously calculated from MD trajectories.^{48,49} Historically, the use of MD simulations to analyze PRE data has been hampered by two factors: the short length of MD trajectories and poor performance of MD force fields in the modeling of disordered polypeptides. Both of these factors are no longer as limiting. With the advent of GPU computing it has become possible to routinely record microsecond trajectories of fully or partially disordered systems in large simulation cells.⁵⁰ At the same time, a new generation of force fields has been developed

(notably including new water models) with the intent to model both structured and disordered proteins.^{51,52} Here we take advantage of these recent advances to show that long MD simulations of the nucleosome core particle can broadly reproduce the experimental PREs measured for histone H4 N-terminal tails in nucleosome samples carrying paramagnetic nitroxide tags attached to different histone H3 sites. Altogether, our results provide a realistic model for the conformational ensemble and interactions of the dynamically disordered H4 tails in nucleosomes.

RESULTS AND DISCUSSION

Spin-Labeled Nucleosome Samples. Four different paramagnetically labeled NCP samples reconstituted with recombinant *Xenopus laevis* histones and the 147-bp Widom 601 DNA motif⁶³ were used in this study. Specifically, as described in detail in the [Supporting Information](#), we employed a histone H3 construct containing the well-studied C110A mutation^{9,11} which removes the lone native cysteine residue while preserving the native histone structure and, starting with this construct, generated the following single-cysteine mutants of H3: K36C, L65C, K79C, and Q125C ([Figure 1A](#)). Each H3 mutant was combined with ¹⁵N-labeled histone H4 along with unlabeled histones H2A and H2B to produce the histone octamer complex, which was then incubated with excess MTSL nitroxide spin-label reagent resulting in the tagging of histone H3 with the non-native paramagnetic side-chain R1 ([Figure 1B](#)) at the unique cysteine site^{54,55} and subsequently reconstituted with the Widom 601 DNA to generate the NCP sample used for the NMR studies. For brevity, the spin-labeled NCP samples are termed 36R1, 65R1, 79R1, and 125R1.

Note that the specific H3 mutation sites selected for this study all correspond to solvent-exposed residues located at the nucleosome surface near nDNA in the coil region preceding helix α N (K36), the ends of helix α 1 (L65 and K79), and the center of helix α 3 (Q125) to avoid any significant perturbations of the NCP structure and assembly. Potential structural perturbations were further minimized by the use of a tagging protocol that involves spin-labeling of the properly assembled histone octamer complex as opposed to individual H3 molecules, with sample fidelity confirmed by mass spectrometry and gel electrophoresis (see the SI text and [Figures S1 and S2 in the Supporting Information](#)). Finally, the locations of paramagnetic tag sites were selected to be within or near the structured globular domain of H3 in order to explore the H4 tail conformational dynamics relative to multiple well-defined reference points within the NCP structure ([Figure 1C](#)). These sites were also chosen to be ~20–60 Å from one another to provide nonredundant data reporting on the spatial distribution of the H4 tail.

Measurements of PREs in Spin-Labeled Nucleosomes. The PRE effects for histone H4 tail residues induced by nitroxide tags attached to histone H3 were quantified by recording for each spin-labeled NCP sample a pair of ¹⁵N–¹H^N heteronuclear single quantum coherence (HSQC) NMR spectra: one for the paramagnetic nitroxide spin-labeled sample and the other for the same sample but with the nitroxide moiety reduced to diamagnetic hydroxylamine by the addition of excess sodium ascorbate ([SI text](#)).

The HSQC NMR spectra of nucleosomes reconstituted with ¹⁵N–H4 ([Figure 1D](#)) feature a limited number of resonances corresponding to the most flexible H4 tail residues (aa ~1–15;

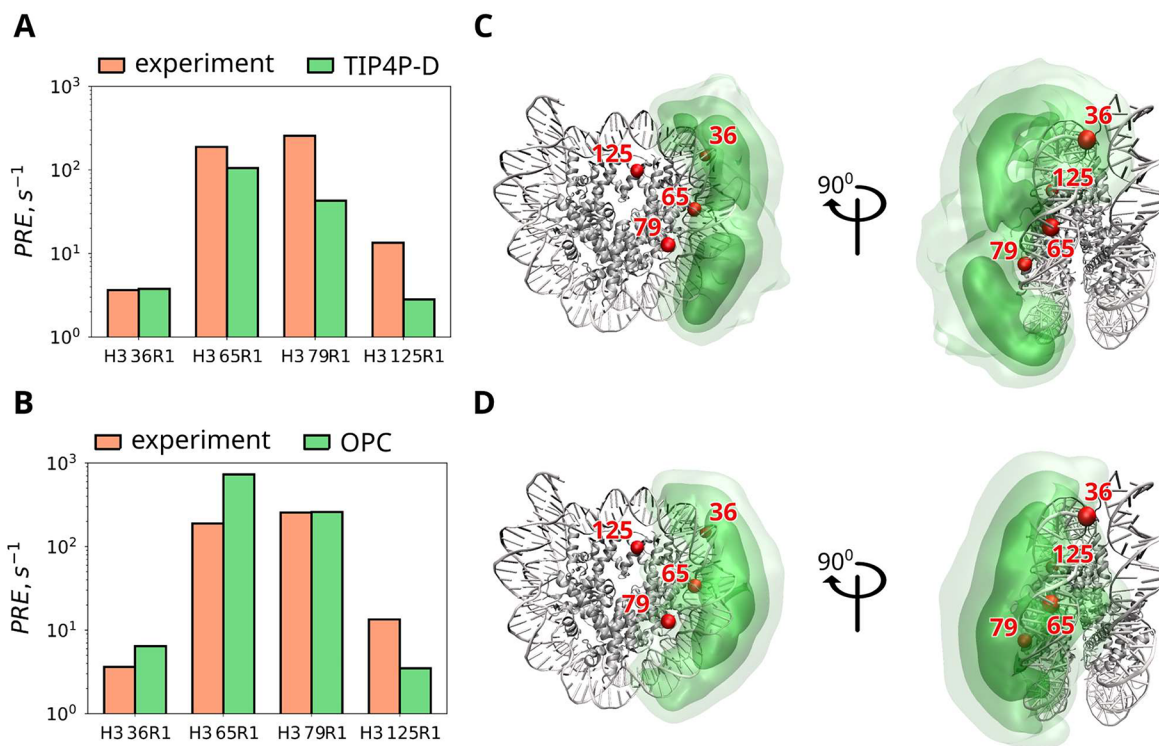


Figure 2. Comparison of experimental and MD-derived paramagnetic relaxation enhancements. (A, B) Experimental PREs averaged over histone H4 residues 3–15 and over two copies of H4 in nucleosomes (coral bars) and the corresponding MD predictions (green bars). The MD results are from TIP4P-D and OPC simulations (panels A and B, respectively). (C, D) Localization of histone H4 tail residues 1–15 in TIP4P-D and OPC simulations (panels C and D, respectively). The images are based on the coordinate set 3LZ0.⁶¹ In constructing these images, we took advantage of the 2-fold pseudosymmetry of the NCP such as to combine the data from the two H4 tails. Shown is the density of C^α atoms for H4 residues 1–15; the isosurfaces containing 50%, 90%, and 99% of the integral density are colored dark, medium, and light green, respectively. The nitroxide tag attachment sites in histone H3 are indicated by spheres and labeled (cf. Figure 1C).

note that NMR signals for the two N-terminal residues are not detected due to amide proton exchange⁹), with only one set of peaks observed in the spectra, which indicates that there are no major structural and dynamic differences between the two H4 N-terminal tails within the NCP. The remaining histone H4 residues reorient relatively slowly—including those in the histone core with dynamics dominated by the slow overall tumbling of the nucleosome particle ($\tau_{\text{rot}} \sim 160 \text{ ns}$ ¹²)—which leads to severe spectral line broadening and prevents their observation.

Superposition of the paramagnetic and corresponding reference diamagnetic NCP spectra (Figure 1D) shows that the presence of paramagnetic tags in H3 positions 65 and 79 causes dramatic attenuation of spectral resonances for most of the H4 tail residues. On the other hand, spin-labels in H3 positions 36 and 125 cause only relatively minor reduction in H4 tail peak intensities. Qualitatively, this can be readily understood given that the former H3 sites are located in relatively close proximity to the attachment point of the H4 N-terminal tail, whereas the latter two are located some distance away, including residue 36, which is separated from the H4 tail origin by the nDNA helix. For each NCP sample, the ratio of peak volumes in the paramagnetic and diamagnetic HSQC NMR spectra, $V_{\text{para}}/V_{\text{dia}}$ is plotted in Figure 1E as a function of H4 residue number and was used to evaluate the residue-specific PREs (SI text and Table S1) as described in detail previously.⁴⁸

Calculation of PREs from MD Trajectories. The experimental PRE data were used to validate two different

MD models of the nucleosome core particle. The first one is our own 2 μs NCP trajectory recorded under Amber ff14SB force field⁵⁶ using the TIP4P-D water model.⁵⁷ Previously, we have shown that this trajectory successfully reproduces ¹⁵N relaxation data for the H4 N-terminal tail in nucleosomes.¹² The second one is a collection of NCP trajectories with the net length of 41 μs ,⁵⁸ which has been recorded in the Panchenko laboratory using the same Amber ff14SB force field together with the OPC water model.⁵⁹ While TIP4P-D was originally designed to improve the modeling of disordered polypeptides (such as H4 tails), OPC was initially devised as a general-purpose model and later shown to perform well for disordered systems.⁶⁰

The methodology used to calculate PREs from MD data is described in detail in the Supporting Information. Importantly, the algorithm employs rigorous Redfield-theory expressions,⁴⁸ which fully account for the relative motion of ¹H^N spins (residing on the H4 tail) and the paramagnetic center (residing on the core histone H3). This motion mainly stems from the conformational dynamics of the histone H4 tail and can be thought of as restricted diffusion; it efficiently modulates both the orientation of the interspin vector and its length (i.e., the distance between the two spins). The distinctive correlation functions resulting from these dynamics are illustrated in Figure S3.

Before we compare the MD-based predictions to the experimental results, we note that for a given spin-labeled sample there is relatively little residue-to-residue variation in the measured PREs over the entire length of the N-terminal

tail (cf. Figure 1E and Table S1). Considering limited convergence of the MD-based residue-specific PRE calculations (discussed in more detail below), it is unlikely that these small residue-to-residue PRE variations can be reproduced. Therefore, we have elected instead to focus on the tail-averaged experimental PREs for each of the four spin-labeled NCP samples and compare these with similarly averaged calculated PREs, which are less susceptible to the convergence issues.

Comparison of Experimental and Calculated PREs.

Figures 2A and 2B show the comparison of experimental and MD-derived PREs for simulations employing TIP4P-D and OPC water, respectively. In these plots, the coral and green bars indicate the experimentally measured and calculated values, respectively. Note that the calculated PREs shown in these graphs constitute strictly a prediction with no tunable parameters of any kind. Clearly, both sets of calculated PREs, based on TIP4P-D and OPC trajectories, capture the qualitative behavior as observed in the experiment, with much higher paramagnetic rates predicted for H3 79R1 and 65R1 samples than for 125R1 and 36R1 samples. While the MD simulations fail to predict the PREs with quantitative accuracy, it is encouraging that TIP4P-D results and OPC results tend to “bracket” the experimentally determined PRE values. For example, the TIP4P-D simulation correctly recovers our experimental 36R1 result, while the OPC simulation produces an overestimated value. Further, the TIP4P-D simulation somewhat underestimates the 65R1 result, while the OPC simulation significantly overestimates it. Finally, the TIP4P-D simulation substantially underestimates the average PRE rate in the 79R1 sample, whereas the OPC simulation provides a good match. Based on these observations, we conclude that the two MD simulations taken together offer a broadly correct model for the PRE effects in spin-labeled nucleosomes.

It is worthwhile to consider the possible reasons for the observed lack of quantitative agreement between experimental and predicted PREs in Figure 2. Almost certainly this has to do with the lack of convergence for the MD-based PRE calculations in the system, which involves a 25-residue-long disordered H4 tail probing a large swath of the nucleosome surface. Indeed, the calculated PREs show an exceedingly steep dependence on the interspin distance (see Supporting Information eq S3). As a consequence, even a brief close encounter of the H4 tail with the paramagnetic label can have a dramatic effect on the calculated PREs. In other words, MD-based calculations of residue-specific PREs are extremely demanding in terms of convergence.⁴⁵ This idea is further illustrated in Figure S4, where the comparison is drawn to four different subsets of the OPC trajectories.

Acknowledging this limitation, we nevertheless note that the results in Figures 2A and 2B provide a broad semiquantitative validation for the MD models at hand. Of special interest to us, these models encode the information about the localization of the histone H4 tails in nucleosomes. This is visualized in Figures 2C and 2D in a form of spatial density distribution for H4 tail residues 1–15 (volumes shaded with different shades of green). Inspection of these plots confirms that these H4 residues are significantly delocalized, as expected for the dynamically disordered tail. At the same time, the H4 tail clearly gravitates toward the nucleosomal DNA, in agreement with the notion previously discussed by us and others that histone tails engage in fuzzy interactions with nDNA.^{12,13}

The results in Figures 2C and 2D show that H3 labels 65R1 and 79R1 are indeed centrally located with respect to the H4 tail “cloud”. On the other hand, 36R1 and 125R1 are located peripherally at a distance from the H4 tail as seen in the side and front views of the NCP, respectively (cf. Figures 2C and 2D). Of interest, different subsets of the OPC trajectories show some variability in terms of the H4 tail density distribution (Figure S4). Some of these models show the density distribution that is somewhat more extended toward the 125R1 label and, accordingly, predict higher PRE rates in this sample (more in line with the experimental PRE data), suggesting that the details of the density distribution can be further elucidated based on longer trajectories and additional experimental data.

While the MD-based interpretation of PRE data is currently faced with certain technical challenges, including subpar convergence in calculation of residue-specific PREs, it constitutes a general and powerful strategy the accuracy of which is only expected to improve in the coming years with continuing increases in simulation length and advances in force field development. In contrast, the more traditional static ensemble-based approaches are faced with several inherent limitations in dealing with PRE data as discussed below.⁴⁵

First, it is noteworthy that ensemble-based approaches have no knowledge of the residue-specific correlation times, τ_c , which are necessary to calculate PRE rates but are challenging to determine experimentally. For histone H4 tails in the nucleosome, MD simulations indicate that these τ_c values vary from one residue to another and can be significantly shorter than the tumbling time of the entire nucleosome particle, τ_{rot} ; furthermore, τ_c values cannot be reliably reconstructed from ¹⁵N relaxation data (see Figure S3). Therefore, when constructing a conformational ensemble, one either needs to make assumptions about τ_c or otherwise use it as a tuning parameter.

Second, an earlier study by Ganguly and Chen⁴⁵ concluded that PRE-restrained ensemble models of disordered proteins tend to be ill-defined. Specifically, due to the r^{-6} weighting associated with PREs, only one or a few ensemble members are typically needed to satisfy the experimental PREs, and the rest of the ensemble could be essentially unrestrained. As such, the resulting ensemble may not necessarily be a realistic representation of the underlying disordered state.

These observations also apply to the system in our study, as illustrated in Figures S5 and S6. In Figure S5, we present a minimal ensemble consisting of just two H4 tail conformers, which clearly does not represent the conformational diversity of the H4 tail domain in the nucleosome core particle yet successfully reproduces the experimental PRE data. Additionally, in Figure S6 we show a larger ensemble of 42 H4 tail conformers, where the PRE rates are largely determined by two lightly populated conformers, while the remaining 40 conformers are effectively unrestrained and largely arbitrary.

Finally, we note that validation of the two MD models of the NCP discussed in this paper does not need to be limited to the PRE data. For instance, both TIP4P-D and OPC simulations were successful in reproducing secondary chemical shifts for H4 tail residues 1–15, correctly predicting the random-coil character of this glycine-rich fragment. The TIP4P-D simulation was also successful in predicting ¹⁵N relaxation rates for H4 tail residues.¹² In contrast, the OPC simulations seem to overemphasize the attraction between the histone tails and nDNA and thus produce biased predictions for ¹⁵N

relaxation rates. Several factors may contribute to this situation, which will be the subject of a separate study. At the same time, the TIP4P-D simulation shows some early signs of nDNA unwrapping which affects the two ends of the DNA superhelix. This behavior likely reflects the known stability issues with this water model.⁵⁷ In contrast, the OPC simulations appear to be stable in this regard, although some of the OPC trajectories show evidence of dynamic nDNA fluctuations (Figure S7).

CONCLUDING REMARKS

Characterization of conformational ensembles for the dynamic histone tail domains in the context of nucleosomes and nucleosome arrays is important for developing an atomic level understanding of their interactions and ultimately function. Of particular importance are changes in tail conformation and interactions that arise in response to different PTMs and involvement of chromatin-associated proteins. These problems can be uniquely addressed by probing a range of NMR observables in both solution and solid state, as appropriate. In this study, we have demonstrated that measurements of residue-specific PREs in nucleosome samples employing multiple paramagnetic tags, which provide unique long-range structural information, should be one of the central elements of such multiexperiment NMR data sets. We have also shown that the interpretation of these PRE data based on long MD simulations of nucleosome particles yields valuable insights about the conformational and interaction landscapes of histone tails in these systems. In the future, we envision that this strategy can be extended to nucleosome complexes with CAPs and even larger nucleosome assemblies. Finally, it is noteworthy that PREs and other types of experimental NMR data can be used to validate MD models and, by extension, inform efforts in the area of force field development. Collectively, these efforts are expected to contribute to an improved understanding of the functional mechanisms in chromatin and other large biomolecular complexes containing functionally important flexible domains.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacs.3c10340>.

Experimental procedures and computational details; table of experimental PRE rates; figures illustrating characterization of nucleosome samples by mass spectrometry and gel electrophoresis, MD-extracted time correlation functions underlying the PRE effect, PREs derived from different subsets of the MD trajectories in OPC water, localization of histone H4 tails as observed in these subsets, examples of PRE interpretation using conformational ensembles, root-mean-square deviation traces for the MD simulations in TIP4P-D and OPC water (PDF)

AUTHOR INFORMATION

Corresponding Authors

Nikolai R. Skrynnikov – Laboratory of Biomolecular NMR, St. Petersburg State University, St. Petersburg 199034, Russia; Department of Chemistry, Purdue University, West Lafayette 47907, United States; orcid.org/0000-0002-7097-9192; Email: n.skrynnikov@spbu.ru

Christopher P. Jaroniec – Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States; orcid.org/0000-0003-0364-2888; Email: jaroniec.1@osu.edu

Authors

Wenjun Sun – Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States; orcid.org/0009-0007-4295-7475

Olga O. Lebedenko – Laboratory of Biomolecular NMR, St. Petersburg State University, St. Petersburg 199034, Russia

Nicole Gonzalez Salguero – Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States

Matthew D. Shannon – Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States

Mohamad Zandian – Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States

Michael G. Poirier – Department of Physics, The Ohio State University, Columbus, Ohio 43210, United States; orcid.org/0000-0002-1563-5792

Complete contact information is available at: <https://pubs.acs.org/10.1021/jacs.3c10340>

Author Contributions

#W.S., O.O.L., and N.G.S. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by grants from the NIH (R01GM123743 to C.P.J. and M.G.P. and R35GM139564 to M.G.P.) and the NSF (MCB-2303862 to C.P.J.). N.R.S. acknowledges Grant 104236506 from SPbU and the SPbU Computer Center for providing some of the computational resources. We thank Prof. Catherine Musselman for sharing the pJ201 plasmid. All the authors express their strong wishes for peace in Ukraine.

REFERENCES

- (1) Whitlock, J. P.; Stein, A. Folding of DNA by histones which lack their NH₂-terminal regions. *J. Biol. Chem.* **1978**, *253*, 3857–3861.
- (2) Cary, P. D.; Moss, T.; Bradbury, E. M. High-resolution proton-magnetic-resonance studies of chromatin core particles. *Eur. J. Biochem.* **1978**, *89*, 475–482.
- (3) Luger, K.; Mäder, A. W.; Richmond, R. K.; Sargent, D. F.; Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **1997**, *389*, 251–260.
- (4) Davey, C. A.; Sargent, D. F.; Luger, K.; Maeder, A. W.; Richmond, T. J. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* **2002**, *319*, 1097–1113.
- (5) Jenuwein, T.; Allis, C. D. Translating the histone code. *Science* **2001**, *293*, 1074–1080.
- (6) Bannister, A. J.; Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Research* **2011**, *21*, 381–395.
- (7) Strahl, B. D.; Allis, C. D. The language of covalent histone modifications. *Nature* **2000**, *403*, 41–45.
- (8) Kalashnikova, A. A.; Porter-Goff, M. E.; Muthurajan, U. M.; Luger, K.; Hansen, J. C. The role of the nucleosome acidic patch in modulating higher order chromatin structure. *J. Royal Soc. Interface* **2013**, *10*, No. 20121022.

- (9) Zhou, B. R.; Feng, H. Q.; Ghirlando, R.; Kato, H.; Gruschus, J.; Bai, Y. W. Histone H4 K16Q mutation, an acetylation mimic, causes structural disorder of its N-terminal basic patch in the nucleosome. *J. Mol. Biol.* **2012**, *421*, 30–37.
- (10) Shi, X. Y.; Prasanna, C.; Nagashima, T.; Yamazaki, T.; Pervushin, K.; Nordenskiöld, L. Structure and dynamics in the nucleosome revealed by solid-state NMR. *Angew. Chem., Int. Ed.* **2018**, *57*, 9734–9738.
- (11) Gao, M.; Nadaud, P. S.; Bernier, M. W.; North, J. A.; Hammel, P. C.; Poirier, M. G.; Jaroniec, C. P. Histone H3 and H4 N-terminal tails in nucleosome arrays at cellular concentrations probed by magic angle spinning NMR spectroscopy. *J. Am. Chem. Soc.* **2013**, *135*, 15278–15281.
- (12) Rabdano, S. O.; Shannon, M. D.; Izmailov, S. A.; Gonzalez Salguero, N.; Zandian, M.; Purusottam, R. N.; Poirier, M. G.; Skrynnikov, N. R.; Jaroniec, C. P. Histone H4 Tails in Nucleosomes: a Fuzzy Interaction with DNA. *Angew. Chem., Int. Ed.* **2021**, *60*, 6480–6487.
- (13) Ghoneim, M.; Fuchs, H. A.; Musselman, C. A. Histone Tail Conformations: A Fuzzy Affair with DNA. *Trends Biochem. Sci.* **2021**, *46*, 564–578.
- (14) Furukawa, A.; Wakamori, M.; Arimura, Y.; Ohtomo, H.; Tsunaka, Y.; Kurumizaka, H.; Umehara, T.; Nishimura, Y. Acetylated histone H4 tail enhances histone H3 tail acetylation by altering their mutual dynamics in the nucleosome. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 19661–19663.
- (15) Hammonds, E. F.; Harwig, M. C.; Paintsil, E. A.; Tillison, E. A.; Hill, R. B.; Morrison, E. A. Histone H3 and H4 tails play an important role in nucleosome phase separation. *Biophys. Chem.* **2022**, *283*, No. 106767.
- (16) Stützer, A.; Liokatis, S.; Kiesel, A.; Schwarzer, D.; Sprangers, R.; Soding, J.; Selenko, P.; Fischle, W. Modulations of DNA contacts by linker histones and post-translational modifications determine the mobility and modifiability of nucleosomal H3 tails. *Mol. Cell* **2016**, *61*, 247–259.
- (17) Morrison, E. A.; Bowerman, S.; Sylvers, K. L.; Wereszczynski, J.; Musselman, C. A. The conformation of the histone H3 tail inhibits association of the BPTF PHD finger with the nucleosome. *eLife* **2018**, *7*, No. e31481.
- (18) Zandian, M.; Gonzalez Salguero, N.; Shannon, M. D.; Purusottam, R. N.; Theint, T.; Poirier, M. G.; Jaroniec, C. P. Conformational Dynamics of Histone H3 Tails in Chromatin. *J. Phys. Chem. Lett.* **2021**, *12*, 6174–6181.
- (19) Morrison, E. A.; Baweja, L.; Poirier, M. G.; Wereszczynski, J.; Musselman, C. A. Nucleosome composition regulates the histone H3 tail conformational ensemble and accessibility. *Nucleic Acids Res.* **2021**, *49*, 4750–4767.
- (20) Shi, X.; Prasanna, C.; Soman, A.; Pervushin, K.; Nordenskiöld, L. Dynamic networks observed in the nucleosome core particles couple the histone globular domains with DNA. *Commun. Biol.* **2020**, *3*, 639.
- (21) Smrt, S. T.; Gonzalez Salguero, N.; Thomas, J. K.; Zandian, M.; Poirier, M. G.; Jaroniec, C. P. Histone H3 core domain in chromatin with different DNA linker lengths studied by ¹H-detected solid-state NMR spectroscopy. *Front Mol. Biosci.* **2023**, *9*, No. 1106588.
- (22) Xiang, S.; le Paige, U. B.; Horn, V.; Houben, K.; Baldus, M.; van Ingen, H. Site-Specific Studies of Nucleosome Interactions by Solid-State NMR Spectroscopy. *Angew. Chem., Int. Ed. Engl.* **2018**, *57*, 4571–4575.
- (23) Ohtomo, H.; Kurita, J. I.; Sakuraba, S.; Li, Z.; Arimura, Y.; Wakamori, M.; Tsunaka, Y.; Umehara, T.; Kurumizaka, H.; Kono, H.; Nishimura, Y. The N-terminal Tails of Histones H2A and H2B Adopt Two Distinct Conformations in the Nucleosome with Contact and Reduced Contact to DNA. *J. Mol. Biol.* **2021**, *433*, No. 167110.
- (24) Shoaib, M.; Chen, Q.; Shi, X.; Nair, N.; Prasanna, C.; Yang, R.; Walter, D.; Frederiksen, K. S.; Einarsson, H.; Svensson, J. P.; Liu, C. F.; Ekwall, K.; Lerdrup, M.; Nordenskiöld, L.; Sørensen, C. S. Histone H4 lysine 20 mono-methylation directly facilitates chromatin openness and promotes transcription of housekeeping genes. *Nat. Commun.* **2021**, *12*, 4800.
- (25) Jennings, C. E.; Zoss, C. J.; Morrison, E. A. Arginine anchor points govern H3 tail dynamics. *Front Mol. Biosci.* **2023**, *10*, No. 1150400.
- (26) Kim, T. H.; Nosella, M. L.; Bolik-Coulon, N.; Harkness, R. W.; Huang, S. K.; Kay, L. E. Correlating histone acetylation with nucleosome core particle dynamics and function. *Proc. Natl. Acad. Sci. U. S. A.* **2023**, *120*, No. e2301063120.
- (27) Musselman, C. A.; Kutateladze, T. G. Visualizing Conformational Ensembles of the Nucleosome by NMR. *ACS Chem. Biol.* **2022**, *17*, 495–502.
- (28) Gaponenko, V.; Howarth, J. W.; Columbus, L.; Gasmi-Seabrook, G.; Yuan, J.; Hubbell, W. L.; Rosevear, P. R. Protein global fold determination using site-directed spin and isotope labeling. *Protein Sci.* **2000**, *9*, 302–309.
- (29) Battiste, J. L.; Wagner, G. Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear Overhauser effect data. *Biochemistry* **2000**, *39*, 5355–5365.
- (30) Iwahara, J.; Schwieters, C. D.; Clore, G. M. Ensemble approach for NMR structure refinement against ¹H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. *J. Am. Chem. Soc.* **2004**, *126*, 5879–5896.
- (31) Kato, H.; van Ingen, H.; Zhou, B.-R.; Feng, H.; Bustin, M.; Kay, L. E.; Bai, Y. Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 12283–12288.
- (32) Zhou, B.-R.; Feng, H.; Kato, H.; Dai, L.; Yang, Y.; Zhou, Y.; Bai, Y. Structural insights into the histone H1-nucleosome complex. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 19390–19395.
- (33) Zhou, B.-R.; Feng, H.; Ghirlando, R.; Li, S.; Schwieters, C. D.; Bai, Y. A Small Number of Residues Can Determine if Linker Histones Are Bound On or Off Dyad in the Chromatosome. *J. Mol. Biol.* **2016**, *428*, 3948–3959.
- (34) Solomon, I.; Bloembergen, N. Nuclear magnetic interactions in the HF molecule. *J. Chem. Phys.* **1956**, *25*, 261–266.
- (35) Lee, L. K.; Rance, M.; Chazin, W. J.; Palmer, A. G. Rotational diffusion anisotropy of proteins from simultaneous analysis of N-15 and C-13(alpha) nuclear spin relaxation. *J. Biomol. NMR* **1997**, *9*, 287–298.
- (36) Iwahara, J.; Clore, G. M. Structure-independent analysis of the breadth of the positional distribution of disordered groups in macromolecules from order parameters for long, variable-length vectors using NMR paramagnetic relaxation enhancement. *J. Am. Chem. Soc.* **2010**, *132*, 13346–13356.
- (37) Lietzow, M. A.; Jamin, M.; Dyson, H. J.; Wright, P. E. Mapping long-range contacts in a highly unfolded protein. *J. Mol. Biol.* **2002**, *322*, 655–662.
- (38) Teilum, K.; Kragelund, B. B.; Poulsen, F. M. Transient structure formation in unfolded acyl-coenzyme A-binding protein observed by site-directed spin labelling. *J. Mol. Biol.* **2002**, *324*, 349–357.
- (39) Wu, K.-P.; Kim, S.; Fela, D. A.; Baum, J. Characterization of conformational and dynamic properties of natively unfolded human and mouse alpha-synuclein ensembles by NMR: Implication for aggregation. *J. Mol. Biol.* **2008**, *378*, 1104–1115.
- (40) Gillespie, J. R.; Shortle, D. Characterization of long-range structure in the denatured state of staphylococcal nuclease. 1. Paramagnetic relaxation enhancement by nitroxide spin labels. *J. Mol. Biol.* **1997**, *268*, 158–169.
- (41) Huang, J. R.; Grzesiek, S. Ensemble calculations of unstructured proteins constrained by RDC and PRE data: a case study of urea-denatured ubiquitin. *J. Am. Chem. Soc.* **2010**, *132*, 694–705.
- (42) Salmon, L.; Nodet, G.; Ozenne, V.; Yin, G. W.; Jensen, M. R.; Zweckstetter, M.; Blackledge, M. NMR characterization of long-range order in intrinsically disordered proteins. *J. Am. Chem. Soc.* **2010**, *132*, 8407–8418.

- (43) Dedmon, M. M.; Lindorff-Larsen, K.; Christodoulou, J.; Vendruscolo, M.; Dobson, C. M. Mapping long-range interactions in α -synuclein using spin-label NMR and ensemble molecular dynamics simulations. *J. Am. Chem. Soc.* **2005**, *127*, 476–477.
- (44) Allison, J. R.; Varnai, P.; Dobson, C. M.; Vendruscolo, M. Determination of the free energy landscape of α -synuclein using spin label nuclear magnetic resonance measurements. *J. Am. Chem. Soc.* **2009**, *131*, 18314–18326.
- (45) Ganguly, D.; Chen, J. H. Structural interpretation of paramagnetic relaxation enhancement-derived distances for disordered protein states. *J. Mol. Biol.* **2009**, *390*, 467–477.
- (46) Marsh, J. A.; Neale, C.; Jack, F. E.; Choy, W.-Y.; Lee, A. Y.; Crowhurst, K. A.; Forman-Kay, J. D. Improved structural characterizations of the drkN SH3 domain unfolded state suggest a compact ensemble with native-like and non-native structure. *J. Mol. Biol.* **2007**, *367*, 1494–1510.
- (47) Beier, A.; Schwarz, T. C.; Kurzbach, D.; Platzer, G.; Tribuzio, F.; Konrat, R. Modulation of Correlated Segment Fluctuations in IDPs upon Complex Formation as an Allosteric Regulatory Mechanism. *J. Mol. Biol.* **2018**, *430*, 2439–2452.
- (48) Xue, Y.; Podkorytov, I. S.; Rao, D. K.; Benjamin, N.; Sun, H. L.; Skrynnikov, N. R. Paramagnetic relaxation enhancements in unfolded proteins: Theory and application to drkN SH3 domain. *Protein Sci.* **2009**, *18*, 1401–1424.
- (49) Xue, Y.; Skrynnikov, N. R. Motion of a disordered polypeptide chain as studied by paramagnetic relaxation enhancements, ^{15}N relaxation, and Molecular Dynamics simulations: how fast is segmental diffusion in denatured ubiquitin? *J. Am. Chem. Soc.* **2011**, *133*, 14614–14628.
- (50) Stone, J. E.; Phillips, J. C.; Freddolino, P. L.; Hardy, D. J.; Trabuco, L. G.; Schulten, K. Accelerating molecular modeling applications with graphics processors. *J. Comput. Chem.* **2007**, *28*, 2618–2640.
- (51) Huang, J.; MacKerell, A. D., Jr. Force field development and simulations of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2018**, *48*, 40–48.
- (52) Mu, J.; Liu, H.; Zhang, J.; Luo, R.; Chen, H.-F. Recent Force Field Strategies for Intrinsically Disordered Proteins. *J. Chem. Inf. Model.* **2021**, *61*, 1037–1047.
- (53) Lowary, P. T.; Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **1998**, *276*, 19–42.
- (54) Altenbach, C.; Oh, K. J.; Trabanino, R. J.; Hideg, K.; Hubbell, W. L. Estimation of inter-residue distances in spin labeled proteins at physiological temperatures: experimental strategies and practical limitations. *Biochemistry* **2001**, *40*, 15471–15482.
- (55) Hubbell, W. L.; Gross, A.; Langen, R.; Lietzow, M. A. Recent advances in site-directed spin labeling of proteins. *Curr. Opin. Struct. Biol.* **1998**, *8*, 649–656.
- (56) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
- (57) Piana, S.; Donchev, A. G.; Robustelli, P.; Shaw, D. E. Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *J. Phys. Chem. B* **2015**, *119*, 5113–5123.
- (58) Peng, Y.; Li, S.; Onufriev, A.; Landsman, D.; Panchenko, A. R. Binding of regulatory proteins to nucleosomes is modulated by dynamic histone tails. *Nat. Commun.* **2021**, *12*, 5280.
- (59) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building water models: a different approach. *J. Phys. Chem. Lett.* **2014**, *5*, 3863–3871.
- (60) Shabane, P. S.; Izadi, S.; Onufriev, A. V. General purpose water model can improve atomistic simulations of intrinsically disordered proteins. *J. Chem. Theory Comput.* **2019**, *15*, 2620–2634.
- (61) Vasudevan, D.; Chua, E. Y. D.; Davey, C. A. Crystal structures of nucleosome core particles containing the '601' strong positioning sequence. *J. Mol. Biol.* **2010**, *403*, 1–10.