Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Molecular dynamics simulations of DNA–DNA and DNA–protein interactions

Jejoong Yoo[1,2], David Winogradoff[3] and Aleksei Aksimentiev[3,4]

The all-atom molecular dynamics method can characterize the molecular-level interactions in DNA and DNA–protein systems with unprecedented resolution. Recent advances in computational technologies have allowed the method to reveal the unbiased behavior of such systems at the microseconds time scale, whereas enhanced sampling approaches have matured enough to characterize the interaction free energy with quantitative precision. Here, we describe recent progress toward increasing the realism of such simulations by refining the accuracy of the molecular dynamics force field, and we highlight recent application of the method to systems of outstanding biological interest.

Addresses
[1] Department of Physics, Sungkyunkwan University, Suwon 16419, Republic of Korea
[2] Center for Self-assembly and Complexity, Institute for Basic Science, Pohang 37673, Republic of Korea
[3] Department of Physics and the Center for the Physics of Living Cells, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[4] Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Corresponding authors: Yoo, Jejoong (jejoong@skku.edu), Aksimentiev, Aleksei (aksiment@illinois.edu)

Current Opinion in Structural Biology 2020, **64**:88–96

This review comes from a themed issue on **Biophysical and computational methods**

Edited by **Nagasuma Chandra** and **Gautam I Menon**

For a complete overview see the Issue and the Editorial

Available online 15th July 2020

https://doi.org/10.1016/j.sbi.2020.06.007

Introduction

The very utility of DNA as the carrier of hereditary information is derived from its ability to undergo hybridization, whereby two DNA strands or a DNA and an RNA strand form a double-stranded (ds) DNA helix or a DNA/RNA duplex in accordance with the pattern of complementary interactions between the strands' bases, Figure 1. RNA is also known to form higher-order structures, such as riboswitches in mRNA and ribozymes, driven by both canonical base pairing and higher-order non-canonical interactions. Other types of DNA–DNA interactions, such as polycation-mediated interactions between dsDNA molecules and end-to-end base stacking interactions are central to the processes of DNA condensation [1] and the repair of double-stranded DNA breaks [2].

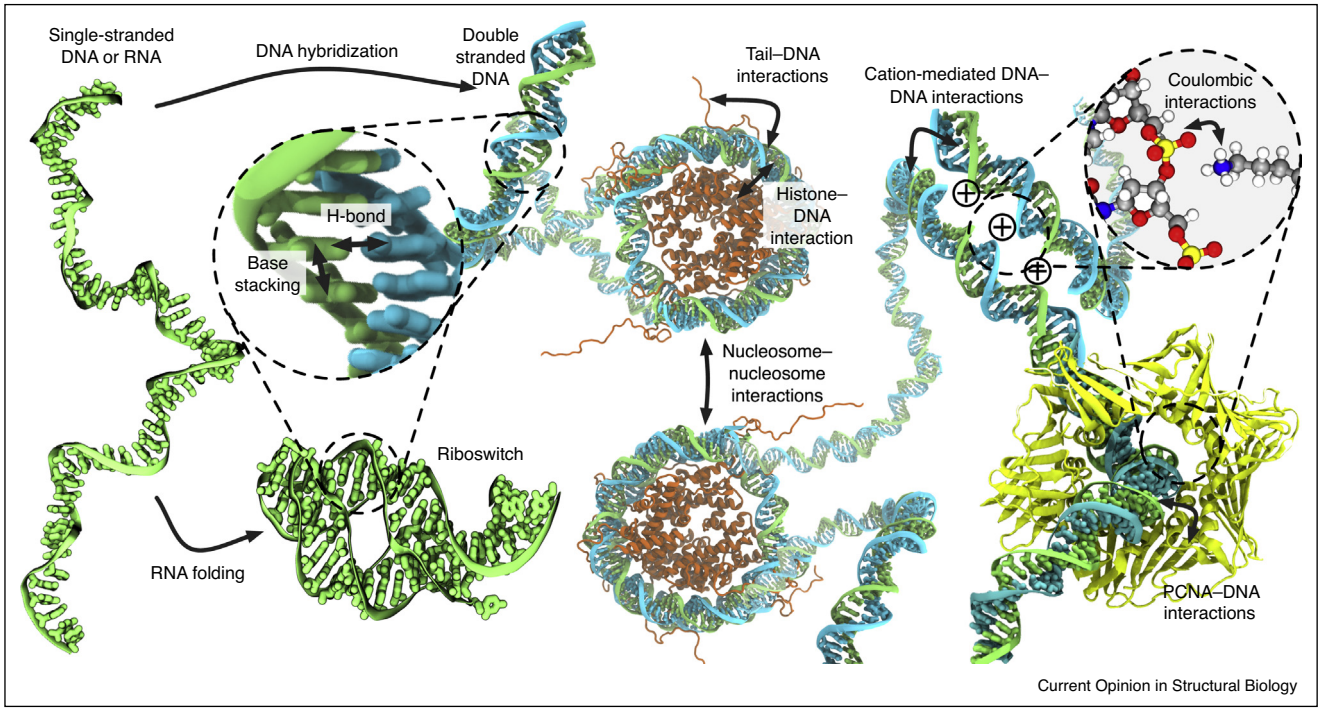
Protein–DNA interactions largely dominate the organization of DNA into higher-order structures. Thus, histone or histone-like proteins (in eukaryotes and prokaryotes, respectively) organize dsDNA at several scales, from individual protein–DNA complexes to higher-order clusters and fibers, and, ultimately, into entire chromosomes. The molecular mechanisms driving such multiscale organization are not yet fully understood, but they are known to play a central role in gene expression. Both transcription and translation require the concerted action of several protein and RNA species organized into massive supramolecular complexes [3]. Similarly, replication of genomic DNA requires the action of numerous DNA-binding proteins to synthesize DNA without losing the genome's integrity [4]. Protein–RNA interactions [5] play essential roles in emerging fields, such as gene editing by CRISPR–Cas [6] and in the formation of membraneless organelles in the nucleus and cytoplasm through phase separation [7]. The biological function of all the above systems critically depends on the strength of nucleic acid (NA)-protein interactions, as those should be strong enough to ensure stable binding of interaction partners that specifically recognize one another [8••] but also weak enough to enable partner seeking by diffusion [9].

Matching their biological importance, NA–NA and protein–NA interactions have been studied extensively using the all-atom molecular dynamics (MD) approach. Paramount to the success of such studies is the precision of the molecular force field that prescribes the interaction strength between the chemical groups. However, recent long time-scale, quantitative characterization of NA–NA and NA–protein interactions revealed considerable imperfections of the existing molecular mechanics models [10,11••,12,13•]. Here, we describe recent advances in characterizing NA–NA and protein–NA interactions using all-atom MD simulation and increasing the realism of such simulations. Readers interested in a more comprehensive description of computational studies of DNA–protein systems are directed to a recent review [14••].

Molecular force fields

All-atom MD simulations of nucleic acids systems are typically performed using the AMBER or CHARMM force fields, which have been validated and improved

Figure 1



Examples of nucleic acid interactions in biological systems. From left to right: nucleic acids form secondary and tertiary structures; histone proteins organize DNA into nucleosomes and higher order structures; DNA interacts with a replication protein.

through multiple cycles of revisions. Presently, AMBER bsc0 [15], bsc1 [16], OL15 [17] and CHARMM36 parameters for RNA [18] and DNA [19] are the most up-to-date choices. A potential alternative is the OPLS force field [20] with the recently developed NA parameters [21], though it requires further validation by the community. For MD simulations of protein-NA systems, AMBER ff99SB [22] or ff14SB [23] or CHARMM36m [24] parameter sets for proteins contain the latest updates.

Historically, the improvement of NA force fields has been focused on the refinements of backbone and glycosidic torsion parameters [15,22,25]. Presently, both AMBER and CHARMM-based simulations can maintain the experimental double helical structure in tens of microseconds simulations [26], although some artifacts with the CHARMM36 simulations of longer dsDNA fragments have been reported recently [27]. It has long been recognized that MD simulations of unfolded proteins exhibit overly collapsed protein conformations [28–30,31•,32–34], in part, because the TIP3P water model (used by both AMBER and CHARMM) energetically favors water–water interactions over water–protein interactions. Conversely, overly strong NA–NA and protein–NA interactions can be potential artifacts of standard AMBER and CHARMM simulations, which, in the case of peptide-mediated DNA–DNA interactions, leads to qualitatively incorrect mulation outcomes [12]. Ion-specific effects are another area of concern, in particular, to simulations of dense NA systems [35].

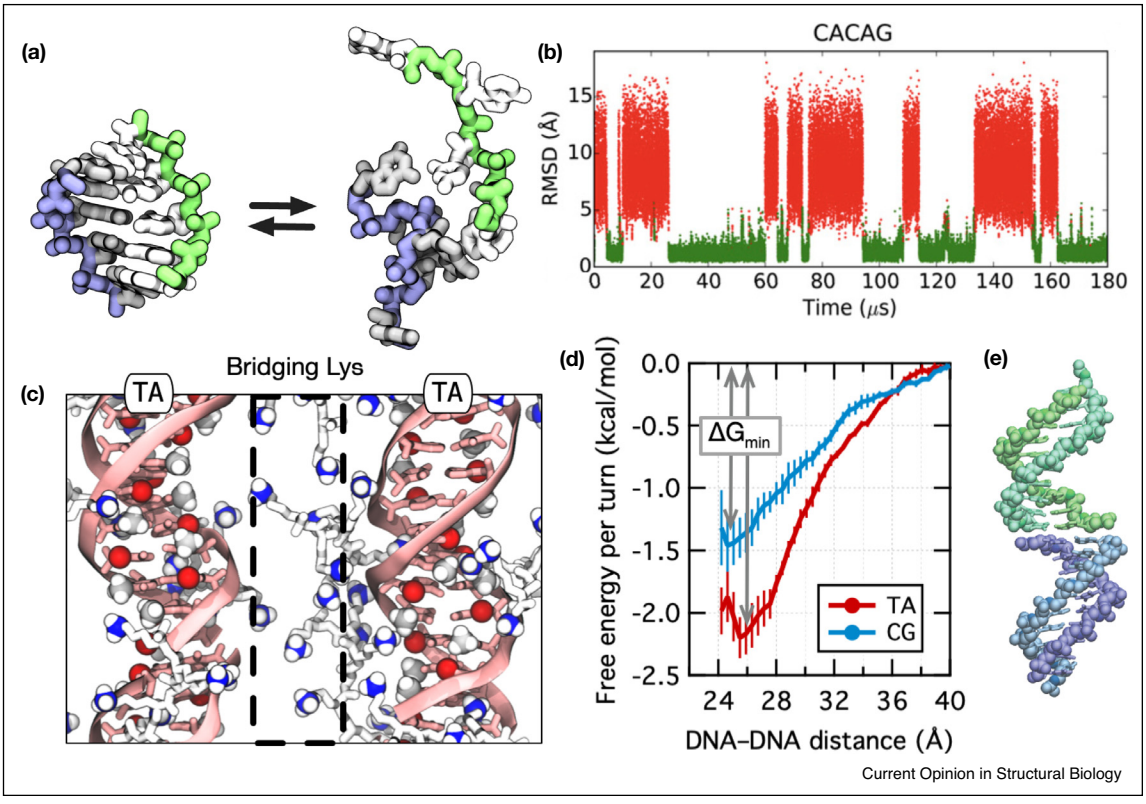
Recognizing the problem, several approaches have been developed to increase the realism of long time-scale simulations of NA systems, which we describe in the subsequent sections.

NA–NA interactions

Hybridization is the most fundamental type of DNA and RNA self-assembly, a process in which hydrogen bonds and base stacking interactions stabilize the double helical structure in a nucleotide sequence-specific manner, Figure 2a. Because of its complexity and biological significance, the hybridization process has become a test ground for advanced simulation methods and force field refinement.

A typical problem encountered in the simulations of short oligomers hybridization or small RNA hairpin folding is the emergence of an intercalated base conformation [10]. Recognizing that the intercalation occurs because of the overestimated base–base stacking and base–backbone interactions, in comparison to base pairing, Chen and Garcia refined the Lennard-Jones (LJ) parameters of nucleobase atoms to weaken the stacking and strengthen

Figure 2



Hybridization of RNA and condensation of DNA. **(a)** Illustration of the hybridization process. **(b)** Root mean squared deviation of a CACAG RNA duplex from its folded configuration observed in a simulated tempering MD simulation. Panel (b) adapted from Ref. [11••]. **(c)** Condensation of DNA duplexes mediated by poly-lysine peptides. **(d)** Condensation free energy of TA and CG repeat DNA duplexes. Panels (c) and (d) adapted from Ref. [40••]. **(e)** End-to-end stacking of two DNA duplexes [47].

base pairing [10]. Following that, several revisions of the LJ parameters have been suggested by several groups [36–38].

The above-mentioned intercalation artifact is another manifestation of the imbalance between water–water and water–solute interactions. To remedy the problem, the Shaw group introduced the TIP4P-D water model and revised the AMBER ff14 RNA force field by optimizing the partial charges, the LJ and torsional parameters [11••]. Note that a similar strategy was used previously by the Shaw group to refine the protein force field [31•]. Using the revised force field, the Shaw group demonstrated dramatic improvements in the simulations of reversible hybridization of RNA duplexes, Figure 2b, as well as in the simulations of unfolded long ssRNAs, tetraloops, and riboswitches [11••]. Although it has not been explicitly shown in the original study, one can expect to observe similar improvements for MD simulations of DNA hybridization.

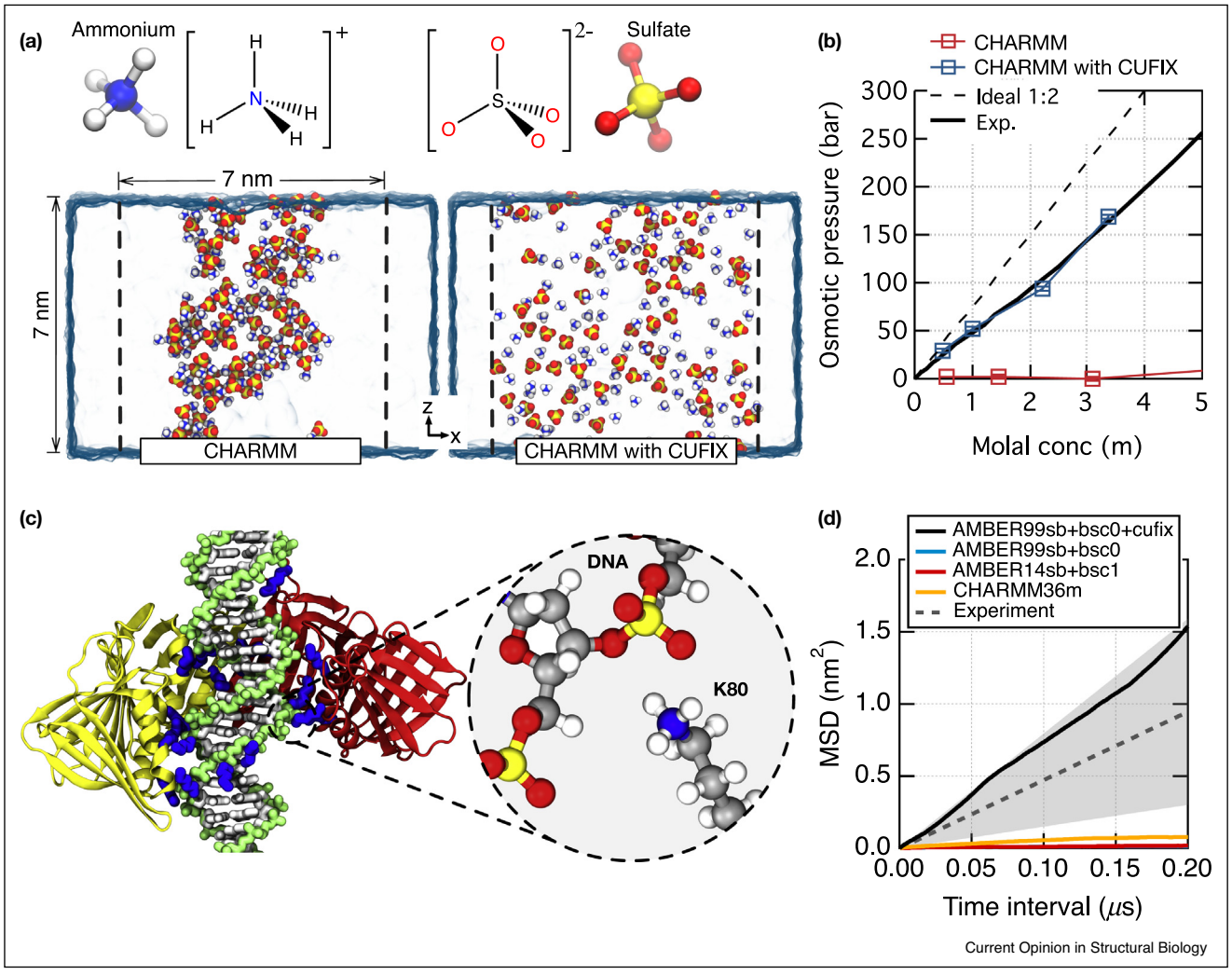
The behavior of densely packed DNA systems, such as in a fully packed viral capsid, a folded chromosome or a synthetic DNA nanostructure, sensitively depends on its ionic environment. When submerged in a monovalent cation solution, dsDNA molecules do not aggregate, regardless of the cation concentration. However, in the presence of tetravalent (or longer) basic peptides, for example, Lys_4, dsDNA can spontaneously form a condensate (equivalently coacervate), in which dsDNA helices form a solvated, ordered phase [1,39]. This self-assembly process is driven by the Coulombic attractions between the basic groups of the peptide and the phosphate groups of DNA.

All-atom MD simulations of dsDNA molecules in aqueous solution of monovalent basic amino acids (e.g. Lys) showed pronounced aggregation of dsDNA for both standard AMBER and CHARMM parameter sets [12], a simulation artifact. The underlying cause of the aggregation was considerable overestimation of the Coulombic attractions between lysine sidechains and DNA [12]. Yoo and Aksimentiev refined the interaction strength using the experimental osmotic pressure of ammonium sulfate as a reference [12]. Briefly, ammonium sulfate solutions of various concentrations were simulated in a semi-permeable membrane setup,

Figure 3a, measuring the effect of the force field corrections on the osmotic pressure. The experimental osmotic pressure was recovered by increasing the LJ R_{min} parameter for the amine nitrogen–sulfate oxygen pairs by about 0.16 Å, Figure 3b [12]. The simulations carried out using this non-bonded correction matched semi-quantitatively the experimental data on the magnitude of the DNA–DNA forces [12]. Readers interested in the development of the so-called CUFIX corrections are referred to a recent review [13•].

The improved realism of MD characterization of DNA–DNA interactions enabled the prediction of previously unknown phenomena. MD simulations with the updated force field predicted that AT-rich dsDNA molecules would attract each other more strongly than GC-rich ones when poly-lysine peptides mediate the inter-DNA interactions, Figure 2c,d [40••]. This prediction immediately suggests that dsDNA molecules can undergo phase separation in a manner that depends on the DNA sequence. Both predictions were validated by single-molecule experiment [40••,41] and bulk liquid–liquid phase separation assays [42]. Given that basic residues are critical functional groups of intrinsically disordered peptides that interact with DNA or RNA [43], the CUFIX corrections may improve MD description of such systems. A potential alternative approach to realizing accurate simulations of peptide-mediated DNA condensation could be the combination of the protein and nucleic acid force fields with improved water models [31•,44•], though this approach has not yet been validated.

Figure 3



Refinement of non-bonded interactions for MD simulation of a protein/DNA complex. **(a,b)** Calibration of charge-charge interactions against osmotic pressure data for aqueous solution of ammonium sulfate. Here, ammonium and sulfate are used as analogs of the protein's amine and DNA's phosphate groups, respectively. **(c)** All-atom model of a PCNA/DNA complex based on a crystal structure [51]. Lysine and arginine residues that form contacts with the DNA phosphates are highlighted in blue. **(d)** Mean squared displacement of PCNA on DNA simulated using four force field models and measured from experiment. Figures in panels (a)–(d) are adapted from Refs. [12] and [52], respectively.

In addition to its role in non-homologous double-strand DNA break repair, end-to-end stacking of dsDNA molecules, Figure 2e, has come to light as a possible mechanism driving the origin of life [45] and as a method of assembling complex synthetic DNA nanostructures [46]. MD simulations of the end-to-end stacking were found to overestimate the absolute magnitude of such interactions but correctly account for the effect of the termination chemistry [47]. Recent characterization of nucleotide type-specific base-stacking interactions [48] has set the stage for future refinement of the MD base-stacking models.

Protein–NA interactions

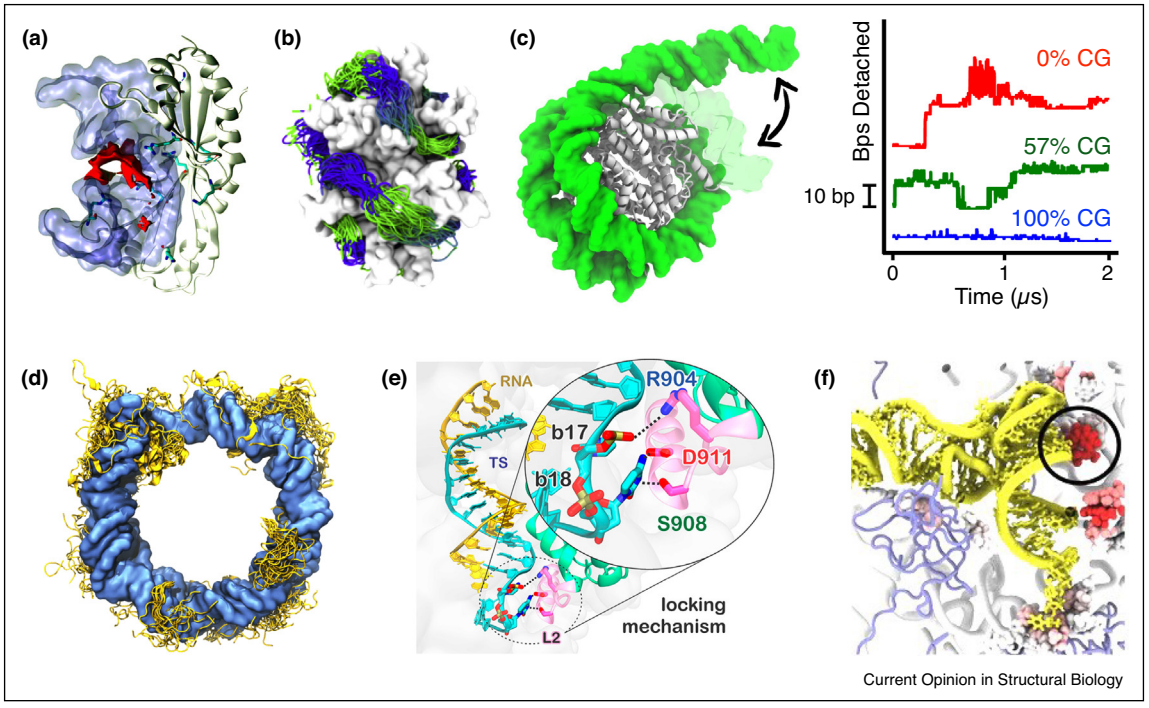
Because of the enormous size of the genome, diffusion along DNA is essential for most DNA-binding proteins to find their targets [9,49]. An outstanding example is the DNA clamp, a ring-shaped protein complex that encircles dsDNA to ensure processive DNA replication in both eukaryotes and prokaryotes, Figure 3c. Because the diffusion coefficient of the eukaryotic clamp, proliferating cell nuclear antigen (PCNA), is greater than 1 $nm^2/\mu s$ [50], a noticeable displacement of PCNA should be observed in a microsecond-long MD simulation. Contrary to that expectation, when a PCNA-dsDNA system is simulated using standard AMBER and CHARMM force fields, the mean-squared displacement (MSD) of the protein relative to dsDNA is orders of magnitude smaller than in experiment, Figure 3d [51,52]. Such a dramatic underestimation of MSD occurs because the lysine and arginine residues at the PCNA–DNA interface form long-lasting contact pairs with DNA phosphates, Figure 3c. In contrast, MSDs obtained from simulations carried out using the CUFIX corrections quantitatively match the experimental value, Figure 3d [52].

The natural outcome of a diffusive search is stable binding of the protein to a DNA fragment carrying a specific nucleotide sequence. Transcription factors (TFs) are one class of such sequence recognition proteins [9,49] that have been studied extensively through all-atom MD approaches [53,54•,55,56•,57••,58•]. An outstanding question in this area is the microscopic mechanism(s) enabling the recognition of the target DNA sequence. Extensive analysis of a large collection of protein-DNA complexes [8••] has recently identified the sequence-specific deformation of a DNA fragment as a critical factor enabling the target recognition by a DNA binding protein, Figure 4a. The free energy of TF binding has been determined using advanced MD sampling methods such as alchemy perturbation [53] and umbrella sampling [54•,55]. One study [52], however, implied that either standard AMBER or standard CHARMM force field systematically overestimates the binding free energy because the DNA–protein interactions typically involve direct charge-charge contacts at the protein-DNA interface [56•], which standard MD force fields are known to overestimate. In spite of such strong binding, TFs were observed to slide along dsDNA, albeit by a small amount, in a recent MD study [59]. Using the CUFIX corrections to charge–charge interactions is expected to reduce the strength of TF–DNA binding and to accelerate TF sliding along DNA.

In between stable binding and free diffusion lies a situation where a protein-DNA complex forms with a high affinity but remains amenable to rearrangement in response to external factors. *In vivo*, ssDNA is almost always sequestered by single-stranded DNA binding (SSB) proteins that bind ssDNA with high affinity. Despite the strong binding, SSBs can diffuse along ssDNA. The elementary steps of such a diffusion process — formation and diffusion of a small DNA bulge, Figure 4b — was recently observed in MD simulations [60]. Another example is the binding of dsDNA to histone proteins. The resulting assembly, the nucleosome, sequesters 147 bases of dsDNA, and many such nucleosomes can form higher-order structures guided by the interactions between DNA and intrinsically disordered histone tails or chromatin remodeling factors. Recent MD studies investigated DNA unwrapping from a protein core [61,62,63•], and interactions between intrinsically disordered histone tails and the surrounding DNA [62,64,65,66,67•]. Spontaneous and reversible nucleosome unravelling, Figure 4c, was observed in MD simulations carried out at elevated magnesium concentration [63•]. Combined with NMR measurements, MD simulations uncovered how histone tails, and their chemical modifications, impede a zinc finger domain from binding to a nucleosome [64,67•], Figure 4d. Other simulations characterized the electrostatic environment at the histone-DNA interface [68•], and the effect of a centromere-specific histone variant on nucleosome elasticity [69•]. Several recent all-atom simulations of multi-nucleosome systems investigated histone tail bridging interactions between two separate nucleosomes [65], and, combined with experiment and coarse-grained simulation, determined how a chaperone protein binds to a di-nucleosome [70]. All-atom MD simulations have also examined how relatively small histone-like proteins can strongly bend [71•] or bridge [72] regions of bacterial DNA.

RNA–protein interactions feature prominently in both CRISPR-Cas9 and ribosomes. The CRISPR-Cas9 system has received significant attention because of its ability to edit genetic information in live cells [73,74]. Several recent all-atom MD studies have investigated the source of CRISPR-Cas9's unintended interaction with off-target DNA sequences [75••,76,77], one of which [75••] identified a conformational 'locking mechanism', Figure 4e, which could be enhanced further through mutagenesis. Lastly, we highlight a simulation study of ssRNA's interaction with the protein and RNA components of a ribosome, Figure 4f [78•], which found diffusion of aminoacyl-tRNA to be significantly impeded by relatively few direct contacts with the ribosome.

Figure 4



All-atom simulations of protein–nucleic acid systems. **(a)** Binding of a protein (white) bends a segment of dsDNA (blue). Red shows a -7.4 kcal mol^{-1} interaction potential isosurface. Adapted from Ref. [8••]. **(b)** Ensemble of conformations explored by ssDNA bound to a single-stranded binding protein (white) within 10 μs. The DNA is depicted as a tube that is colored from green to blue every 10 nt. Adapted from Ref. [60]. **(c)** Left, partial unwrapping of 601L DNA (green) from a histone core (white). Right, spontaneous unwrapping for three DNA sequences. Adapted from Ref. [63•]. **(d)** Non-specific binding of histone tails (gold) to nucleosomal DNA (blue). Adapted from Ref. [67•]. **(e)** MD simulation of a CRISPR-Cas9 system. The inset illustrates the interactions between the target strand of an RNA:DNA hybrid (TS, cyan) with loop 2 (L2, pink) of the protein catalytic domain. Adapted from Ref. [75••]. **(f)** An 'elbow' of aminoacyl-tRNA (yellow) interacts with the RNA (white tubes) and protein (blue tubes) components of a ribosome. The black circle highlights a prominent contact between rRNA and tRNA. Adapted from Ref. [78•].

Conclusions

As the scope of all-atom MD simulations evolves from individual proteins or NAs to systems containing hundreds of such biomolecules, so too will the need to refine the underlying computational models to accurately describe non-bonded interactions between those molecules. An outstanding challenge for the field lies in describing intrinsically disordered regions (IDRs) of the DNA-binding proteins, which are omnipresent and play important roles in the recruitment and formation of transient biomolecular complexes. One particularly exciting and challenging class of systems are NA-IDR condensates [40••,79], the biological roles of which we are just beginning to grasp. Here, in addition to further refinement of the all-atom molecular force field, methods that permit adequate sampling of the conformational space in an entangled polymer melt environment will be needed [80,81]. In general, the field would greatly benefit from a closer integration of coarse-grained models, such as Martini [82], oxDNA [83], 3SPN.2 [84] and ABSINTH [85], with all-atom approaches, allowing for mixed-resolution, accuracy-when-needed types of description of very large protein-NA systems.

Conflict of interest statement

Nothing declared.

Acknowledgements

This work was supported by the Institute for Basic Science [IBS-R007-Y1] and the National Science Foundation (USA) through the Center for the Physics of Living Cells grant PHY-1430124.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Rau DC, Lee B, Parsegian VA: **Measurement of the repulsive force between polyelectrolyte molecules in ionic solution: hydration forces between parallel DNA double helices**. *Proc Natl Acad Sci U S A* 1984, **81**:2621-2625.

2. Duckett DR, Murchie AI, Diekmann S, Kitzing von E, Kemper B, Lilley DM: **The structure of the Holliday junction, and its resolution**. *Cell* 1988, **55**:79-89.

3. He Y, Yan C, Fang J, Inouye C, Tjian R, Ivanov I, Nogales E: **Near-atomic resolution visualization of human transcription promoter opening**. *Nature* 2016, **533**:359-365.

4. Johnson A, O'Donnell M: **Cellular DNA replicases: components and dynamics at the replication fork**. *Annu Rev Biochem* 2005, **74**:283-315.

5. Ghaemi Z, Guzman I, Gnutt D, Luthey-Schulten Z, Gruebele M: **Role of electrostatics in protein-RNA binding: the global vs the local energy landscape**. *J Phys Chem B* 2017, **121**:8437-8446.

6. Knott GJ, Doudna JA: **CRISPR-Cas guides the future of genetic engineering**. *Science* 2018, **361**:866-869.

7. Shin Y, Brangwynne CP: **Liquid phase condensation in cell physiology and disease**. *Science* 2017, **357**:eaaf4382-13.

8. •• Battistini F, Hospital A, Buitrago D, Gallego D, Dans PD, Gelpí JL, Orozco M: **How B-DNA dynamics decipher sequence-selective protein recognition**. *J Mol Biol* 2019, **431**:3845-3859
Comprehensive computational study of DNA sequence recognition by DNA binding proteins.

9. Blainey PC, Luo G, Kou SC, Mangel WF, Verdine GL, Bagchi B, Xie XS: **Nonspecifically bound proteins spin while diffusing along DNA**. *Nat Struct Mol Biol* 2009, **16**:1224-1229.

10. Chen AA, Garcia AE: **High-resolution reversible folding of hyperstable RNA tetraloops using molecular dynamics simulations**. *Proc Natl Acad Sci U S A* 2013, **110**:16820-16825.

11. •• Tan D, Piana S, Dirks RM, Shaw DE: **RNA force field with accuracy comparable to state-of-the-art protein force fields**. *Proc Natl Acad Sci U S A* 2018, **115**:E1346-E1355
A revised AMBER-based RNA force field tested with the TIP4P-D water model. Could improve description of protein-RNA complexes.

12. Yoo J, Aksimentiev A: **Improved parameterization of amine-carboxylate and amine-phosphate interactions for molecular dynamics simulations using the CHARMM and AMBER force fields**. *J Chem Theory Comput* 2016, **12**:430-443.

13. • Yoo J, Aksimentiev A: **New tricks for old dogs: improving the accuracy of biomolecular force fields by pair-specific corrections to non-bonded interactions**. *Phys Chem Chem Phys* 2018, **20**:8432-8449
Describes Champaign-Urbana non-bonded fix (CUFIX) corrections to atomistic AMBER and CHARMM force fields.

14. •• Maffeo C, Chou H-Y, Aksimentiev A: **Molecular mechanisms of DNA replication and repair machinery: insights from microscopic simulations**. *Adv Theory Simul* 2019, **2**:1800191
Comprehensive review of atomistic and coarse-grained molecular dynamics studies of DNA replication and repair machineries.

15. Perez A, Marchan IAN, Svozil D, Sponer J, Cheatham TE, Laughton CA, Orozco M: **Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers**. *Biophys J* 2007, **92**:3817-3829.

16. Ivani I, Dans PD, Noy A, Perez A, Faustino I, Hospital A, Walther J, Andrio P, Goñi R, Balaceanu A *et al.*: **Parmbsc1: a refined force field for DNA simulations**. *Nat Methods* 2016, **13**:55-58.

17. Zgarbová M, Sponer J, Otyepka M, Cheatham TE, Galindo-Murillo R, Jurečka P: **Refinement of the sugar-phosphate backbone torsion beta for AMBER force fields improves the description of Z- and B-DNA**. *J Chem Theory Comput* 2015, **11**:5723-5736.

18. Denning EJ, Priyakumar UD, Nilsson L, Mackerell AD: **Impact of 2'-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA**. *J Comput Chem* 2011, **32**:1929-1943.

19. Hart K, Foloppe N, Baker CM, Denning EJ, Nilsson L, Mackerell AD: **Optimization of the CHARMM additive force field for DNA: improved treatment of the BI/BII conformational equilibrium**. *J Chem Theory Comput* 2012, **8**:348-362.

20. Jorgensen WL, Tirado-Rives J: **The OPLS potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin**. *J Am Chem Soc* 1988, **110**:1657-1666.

21. Robertson MJ, Tirado-Rives J, Jorgensen WL: **Improved treatment of nucleosides and nucleotides in the OPLS-AA force field**. *Chem Phys Lett* 2017, **683**:276-280.

22. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA: **A second generation force field for the simulation of proteins, nucleic acids, and organic molecules**. *J Am Chem Soc* 1995, **117**:5179-5197.

23. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C: **ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB**. *J Chem Theory Comput* 2015, **11**:3696-3713.

24. Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, Grubmuller H, MacKerell AD Jr: **CHARMM36m: an improved force field for folded and intrinsically disordered proteins**. *Nat Methods* 2017, **14**:71-73.

25. Zgarbova M, Otyepka M, Sponer J, Mladek A, Banas P, Cheatham TE III, Jurecka P: **Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles**. *J Chem Theory Comput* 2011, **7**:2886-2902.

26. Galindo-Murillo R, Robertson JC, Zgarbová M, Sponer J, Otyepka M, Jurečka P, Cheatham TE: **Assessing the current state of amber force field modifications for DNA**. *J Chem Theory Comput* 2016, **12**:4114-4127.

27. Minhas V, Sun T, Mirzoev A, Korolev N, Lyubartsev AP, Nordenskiöld L: **Modeling DNA flexibility: comparison of force fields from atomistic to multiscale levels**. *J Phys Chem B* 2020, **124**:38-49.

28. Piana S, Klepeis JL, Shaw DE: **Assessing the accuracy of physical models used in protein-folding simulations: quantitative evidence from long molecular dynamics simulations**. *Curr Opin Struct Biol* 2014, **24**:98-105.

29. Best RB, Zheng W, Mittal J: **Balanced protein-water interactions improve properties of disordered proteins and non-specific protein association**. *J Chem Theory Comput* 2014, **10**:5113-5124.

30. Yoo J, Aksimentiev A: **Refined parameterization of nonbonded interactions improves conformational sampling and kinetics of protein folding simulations**. *J Phys Chem Lett* 2016, **7**:3812-3818.

31. • Robustelli P, Piana S, Shaw DE: **Developing a molecular dynamics force field for both folded and disordered protein states**. *Proc Natl Acad Sci U S A* 2018, **119** 201800690–9
Revised AMBER protein force field that, together with a revised RNA force field, might improve description of RNA-protein systems. Needs further testing.

32. Best RB: **Computational and theoretical advances in studies of intrinsically disordered proteins**. *Curr Opin Struct Biol* 2017, **42**:147-154.

33. Huang J, Mackerell AD: **Force field development and simulations of intrinsically disordered proteins**. *Curr Opin Struct Biol* 2018, **48**:40-48.

34. Nerenberg PS, Head-Gordon T: **New developments in force fields for biomolecular simulations**. *Curr Opin Struct Biol* 2018, **49**:129-138.

35. Yoo J, Aksimentiev A: **Improved parametrization of Li+, Na+, K +, and Mg2+ ions for all-atom molecular dynamics simulations of nucleic acid systems**. *J Phys Chem Lett* 2012, **3**:45-50.

36. Bergonzo C, Cheatham TE III: **Improved force field parameters lead to a better description of RNA structure**. *J Chem Theory Comput* 2015, **11**:3969-3972.

37. Yang C, Lim M, Kim E, Pak Y: **Predicting RNA structures via a simple van der Waals correction to an all-atom force field**. *J Chem Theory Comput* 2017, **13**:395-399.

38. Kührová P, Mlýnský V, Zgarbová M, Krepl M, Bussi G, Best RB, Otyepka M, Sponer J, Banáš P: **Improving the performance of the amber RNA Force field by tuning the hydrogen-bonding interactions**. *J Chem Theory Comput* 2019, **15**:3288-3305.

39. DeRouchey J, Hoover B, Rau DC: **A comparison of DNA compaction by arginine and lysine peptides: a physical basis for arginine rich protamines**. *Biochemistry* 2013, **52**:3000-3009.

40. •• Kang H, Yoo J, Sohn B-K, Lee S-W, Lee HS, Ma W, Kee J-M, Aksimentiev A, Kim H: **Sequence-dependent DNA condensation as a driving force of DNA phase separation**. *Nucleic Acids Res* 2018, **128**:787
Combined computational and experimental study that demonstrated phase-separation regulated by DNA sequence and methylation.

41. Yoo J, Kim H, Aksimentiev A, Ha T: **Direct evidence for sequence-dependent attraction between double-stranded DNA controlled by methylation**. *Nat Commun* 2016, **7**:11045.

42. Shakya A, King JT: **DNA local-flexibility-dependent assembly of phase-separated liquid droplets**. *Biophys J* 2018, **115**:1840-1847.

43. van der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, Fuxreiter M, Gough J, Gsponer J, Jones DT *et al.*: **Classification of intrinsically disordered regions and proteins**. *Chem Rev* 2014, **114**:6589-6631.

44. • Shabane PS, Izadi S, Onufriev AV: **General purpose water model can improve atomistic simulations of intrinsically disordered proteins**. *J Chem Theory Comput* 2019, **15**:2620-2634
Describes a 4-point explicit water model (OPC) that improves description of disordered proteins in comparison to the TIP3P water model.

45. Smith GP, Fraccia TP, Todisco M, Zanchetta G, Zhu C, Hayden E, Bellini T, Clark NA: **Backbone-free duplex-stacked monomer nucleic acids exhibiting Watson-Crick selectivity**. *Proc Natl Acad Sci U S A* 2018, **115**:E7658-E7664.

46. Dietz H, Douglas SM, Shih WM: **Folding DNA into twisted and curved nanoscale shapes**. *Science* 2009, **325**:725-730.

47. Maffeo C, Luan B, Aksimentiev A: **End-to-end attraction of duplex DNA**. *Nucleic Acids Res* 2012, **40**:3812-3821.

48. Kilchherr F, Wachauf C, Pelz B, Rief M, Zacharias M, Dietz H: **Single-molecule dissection of stacking forces in DNA**. *Science* 2016, **353** aaf5508–aaf5508.

49. Elf J, Li G-W, Xie XS: **Probing transcription factor dynamics at the single-molecule level in a living cell**. *Science* 2007, **316**:1191-1194.

50. Kochaniak AB, Habuchi S, Loparo JJ, Chang DJ, Cimprich KA, Walter JC, van Oijen AM: **Proliferating cell nuclear antigen uses two distinct modes to move along DNA**. *J Biol Chem* 2009, **284**:17700-17710.

51. De March M, Merino N, Barrera-Vilarmau S, Crehuet R, Onesti S, Blanco FJ, De Biasio A: **Structural basis of human PCNA sliding on DNA**. *Nat Commun* 2017, **8**:13935.

52. You S, Lee H, Kim K, Yoo J: **Improved parameterization of protein–DNA interactions for molecular dynamics simulations of PCNA diffusion on DNA**. *J Chem Theory Comput* 2020 http://dx.doi.org/10.1021/acs.jctc.0c00241.

53. Gapsys V, de Groot BL: **Alchemical free energy calculations for nucleotide mutations in protein-DNA complexes**. *J Chem Theory Comput* 2017, **13**:6275-6289.

54. • Jakubec D, Vondrášek J: **Can all-atom molecular dynamics simulations quantitatively describe homeodomain-DNA binding equilibria?** *J Chem Theory Comput* 2019, **15**:2635-2648
Comprehensive MD study that determined the potentials of mean force for binding transcription factors to DNA.

55. Wieczór M, Czub J: **How proteins bind to DNA: target discrimination and dynamic sequence search by the telomeric protein TRF1**. *Nucleic Acids Res* 2017, **45**:7643-7654.

56. • Liao Q, Lüking M, Krüger DM, Deindl S, Elf J, Kasson PM, Kamerlin SCL: **Long time-scale atomistic simulations of the structure and dynamics of transcription factor-DNA recognition**. *J Phys Chem B* 2019, **123**:3576-3590
A simulation study uncovering the specific and non-specific binding of protein LacI to DNA.

57. •• Morgunova E, Yin Y, Das PK, Jolma A, Zhu F, Popov A, Xu Y, Nilsson L, Taipale J: **Two distinct DNA sequences recognized by transcription factors represent enthalpy and entropy optima**. *eLife* 2018, **7**:213
The study characterized the role of water molecules at the protein-DNA interfaces.

58. • Baird-Titus JM, Thapa M, Doerdelmann T, Combs KA, Rance M: **Lysine side-chain dynamics in the binding site of homeodomain/DNA complexes as observed by NMR relaxation experiments and molecular dynamics simulations**. *Biochemistry* 2018, **57**:2796-2813
Combined NMR and simulation study uncovering lysine's role in protein-DNA interactions.

59. Zacharias M: **Atomic resolution insight into Sac7d protein binding to DNA and associated global changes by molecular dynamics simulations**. *Angew Chem Int Ed Engl* 2019, **58**:5967-5972.

60. Maffeo C, Aksimentiev A: **Molecular mechanism of DNA association with single-stranded DNA binding protein**. *Nucleic Acids Res* 2017, **45**:12125-12139.

61. Kono H, Sakuraba S, Ishida H: **Free energy profiles for unwrapping the outer superhelical turn of nucleosomal DNA**. *PLoS Comput Biol* 2018, **14**:e1006024.

62. Chakraborty K, Kang M, Loverde SM: **Molecular mechanism for the role of the H2A and H2B histone tails in nucleosome repositioning**. *J Phys Chem B* 2018, **122**:11827-11840.

63. • Winogradoff D, Aksimentiev A: **Molecular mechanism of spontaneous nucleosome unraveling**. *J Mol Biol* 2019, **431**:323-335
Describes spontaneous unravelling of nucleosomal DNA from histone core proteins and its dependence on DNA sequence.

64. Gatchalian J, Wang X, Ikebe J, Cox KL, Tencer AH, Zhang Y, Burge NL, Di L, Gibson MD, Musselman CA *et al.*: **Accessibility of the histone H3 tail in the nucleosome for binding of paired readers**. *Nat Commun* 2017, **8**:1-10.

65. Ishida H, Kono H: **H4 tails potentially produce the diversity in the orientation of two nucleosomes**. *Biophys J* 2017, **113**:978-990.

66. Fu I, Cai Y, Geacintov NE, Zhang Y, Broyde S: **Nucleosome histone tail conformation and dynamics: impacts of lysine acetylation and a nearby minor groove Benzo[a]pyrene-derived lesion**. *Biochemistry* 2017, **56**:1963-1973.

67. • Morrison EA, Bowerman S, Sylvers KL, Wereszczynski J, Musselman CA: **The conformation of the histone H3 tail inhibits association of the BPTF PHD finger with the nucleosome**. *eLife* 2018, **7**:e78587
Combined NMR and simulation study investigating how histone tails impede a zinc finger from binding to a nucleosome.

68. • Elbahnsi A, Retureau R, Baaden M, Hartmann B, Oguey C: **Holding the nucleosome together: a quantitative description of the DNA-histone interface in solution**. *J Chem Theory Comput* 2018, **14**:1045-1058
Maps the electrostatic environment and residue-nucleotide contacts at the histone-DNA interface of a nucleosome.

69. • Melters DP, Pitman M, Rakshit T, Dimitriadis EK, Bui M, Papoian GA, Dalal Y: **Intrinsic elasticity of nucleosomes is encoded by histone variants and calibrated by their binding partners**. *Proc Natl Acad Sci U S A* 2019, **116**:24066-24074
Simulation and AFM study reporting the Young's modulus of wild type and centromere-specific nucleosomes.

70. Watanabe S, Mishima Y, Shimizu M, Suetake I, Takada S: **Interactions of HP1 bound to H3K9me3 dinucleosome by molecular simulations and biochemical assays**. *Biophys J* 2018, **114**:2336-2351.

71. • Hognon C, Garaude S, Timmins J, Chipot C, Dehez F, Monari A: **Molecular bases of DNA packaging in bacteria revealed by all-atom molecular dynamics simulations: the case of histone-like proteins in Borrelia burgdorferi**. *J Phys Chem Lett* 2019, **10**:7200-7207

MD study on the stability and energetics of highly bent DNA in complex with a histone-like protein.

72. Riccardi E, van Mastbergen EC, Navarre WW, Vreede J: **Predicting the mechanism and rate of H-NS binding to AT-rich DNA**. *PLoS Comput Biol* 2019, **15**:e1006845.

73. Palermo G, Miao Y, Walker RC, Jinek M, McCammon JA: **CRISPR-Cas9 conformational activation as elucidated from enhanced molecular simulations**. *Proc Natl Acad Sci U S A* 2017, **114**:7260-7265.

74. Huai C, Li G, Yao R, Zhang Y, Cao M, Kong L, Jia C, Yuan H, Chen H, Lu D *et al.*: **Structural insights into DNA cleavage activation of CRISPR-Cas9 system**. *Nat Commun* 2017, **8**:1375-1379.

75. •• Ricci CG, Chen JS, Miao Y, Jinek M, Doudna JA, McCammon JA, Palermo G: **Deciphering off-target effects in CRISPR-Cas9 through accelerated molecular dynamics**. *ACS Cent Sci* 2019, **5**:651-662

MD study detailing how mismatched DNA that is not being 'locked' by a catalytic CRISPR-Cas9 domain leads to off-target DNA binding.

76. Tangprasertchai NS, Di Felice R, Zhang X, Slaymaker IM, Vazquez Reyes C, Jiang W, Rohs R, Qin PZ: **CRISPR-Cas9 mediated DNA unwinding detected using site-directed spin labeling**. *ACS Chem Biol* 2017, **12**:1489-1493.

77. Zeng Y, Cui Y, Zhang Y, Zhang Y, Liang M, Chen H, Lan J, Song G, Lou J: **The initiation, propagation and dynamics of CRISPR-SpyCas9 R-loop complex**. *Nucleic Acids Res* 2018, **46**:350-361.

78. • Yang H, Bandarkar P, Horne R, Leite VBP, Chahine J, Whitford PC: **Diffusion of tRNA inside the ribosome is position-dependent**. *J Chem Phys* 2019, **151**:085102

Simulation study investigating how an 'elbow' of aminoacyl-tRNA interacts with the protein and RNA components of a ribosome.

79. Boeynaems S, Alberti S, Fawzi NL, Mittag T, Polymenidou M, Rousseau F *et al.*: **Protein phase separation: a new phase in cell biology**. *Trends Cell Biol* 2018, **28**:420-435 http://dx.doi.org/10.1016/j.tcb.2018.02.004.

80. Chou H-Y, Aksimentiev A: **Single-protein collapse determines phase equilibria of a biological condensate**. *J Phys Chem Lett* 2020, **11**:4923-4929 http://dx.doi.org/10.1021/acs.jpclett.0c01222.

81. Zeng X, Holehouse AS, Chilkoti A, Mittag T, Pappu RV: **Connecting coil-to-globule transitions to full phase diagrams for intrinsically disordered proteins**. *Biophys J* 2020 http://dx.doi.org/10.1016/j.bpj.2020.06.014.

82. Uusitalo JJ, Ing o Ifsson HI, Akhshi P, Tieleman DP, Marrink SJ: **Martini coarse-grained force field: extension to DNA**. *J Chem Theory Comput* 2015, **11**:3932-3945.

83. Snodin BEK, Randisi F, Mosayebi M, Šulc P, Schreck JS, Romano F, Ouldridge TE, Tsukanov R, Nir E, Louis AA *et al.*: **Introducing improved structural properties and salt dependence into a coarse-grained model of DNA**. *J Chem Phys* 2015, **142**:234901.

84. Hinckley DM, Freeman GS, Whitmer JK, de Pablo JJ: **An experimentally-informed coarse-grained 3-Site-Per-nucleotide model of DNA: structure, thermodynamics, and dynamics of hybridization**. *J Chem Phys* 2013, **139**:144903.

85. Vitalis A, Pappu RV: **ABSINTH: a new continuum solvation model for simulations of polypeptides in aqueous solutions**. *J Comput Chem* 2009, **30**:673-699.