HHS Public Access

Author manuscript

Phys Chem Chem Phys. Author manuscript; available in PMC 2019 March 28.

Published in final edited form as:

Phys Chem Chem Phys. 2018 March 28; 20(13): 8432–8449. doi:10.1039/C7CP08185E.

New tricks for old dogs: Improving the accuracy of biomolecular force fields by pair-specific corrections to non-bonded interactions

Jejoong Yoo[a,b] and **Aleksei Aksimentiev**[a]

[a]Center for the Physics of Living Cells, Department of Physics and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, IL 61801, USA. Fax: +1 866 467 5398; Tel: +1 217 333 6495; aksiment@illinois.edu

[b]Center for Self-assembly and Complexity, Institute for Basic Science, Pohang, 37363, Republic of Korea

Abstract

In contrast to ordinary polymers, the vast majority of biological macromolecules adopt highly ordered three-dimensional structures that define their functions. The key to folding of a biopolymer into a unique 3D structure or to assembly of several biopolymers into a functional unit is a delicate balance between the attractive and repulsive forces that also makes such self-assembly reversible under physiological conditions. The all-atom molecular dynamics (MD) method has emerged as a powerful tool for studies of individual biomolecules and their functional assemblies, encompassing systems of ever increasing complexity. However, advances in parallel computing technology have outpaced development of the underlying theoretical models—the molecular force fields, pushing the MD method into untested territory. Recent tests of the MD method have found the most commonly used molecular force fields to be out of balance, overestimating attractive interactions between charged and hydrophobic groups, which can promote artificial aggregation in MD simulations of multi-component protein, nucleic acid, and lipid systems. One route to improving the force fields is through the NBFIX corrections method, in which the intermolecular forces are calibrated against experimentally measured quantities such as osmotic pressure by making atom pair-specific adjustments to the non-bonded interactions. In this article, we review development of the NBFIX corrections to the AMBER and CHARMM force fields and discuss their implications for MD simulations of electrolyte solutions, dense DNA systems, Holliday junctions, protein folding and lipid bilayer membranes.

1 Introduction

Since the 9.2 ps molecular dynamics (MD) simulation of the 58 residue bovine pancreatic trypsin inhibitor (BPTI)[1], the MD simulation method has been developing in sync with the computer power that doubles every two years according to the Moore's law[2,3]. Over the history of MD simulations, the ever-increasing computer power has been putting the

Correspondence to: Aleksei Aksimentiev.

underlying theoretical models—empirical molecular force fields—under evermore stringent tests, stimulating force field development that has made the next-generation simulations possible. The most popular biomolecular models, the CHARMM and AMBER force fields, have had several major updates to improve description of proteins[4–8], lipids[9,10], and nucleic acids[4,11–18]. Presently, microsecond simulations of monomeric proteins can be readily obtained using commodity computers[19–22]. The emergence of the special purpose hardware and massively parallel computer systems has pushed the time scale of all-atom MD simulations into milliseconds[23,24] and has made simulations of large biomolecular systems, such as an HIV virus, possible[25].

Recent spectacular improvements in single-node performance of the MD codes[26,27] have dramatically lowered the cost of ensemble simulations, making it computationally feasible to probe not only near-native conformations of biomolecules but also their conformations in transient states including those where the biomolecules are partially or completely denatured. The growing recognition of the functional roles played by denatured conformations, exemplified by intrinsically disordered proteins, have made the sampling of denatured ensembles even more important[28]. However, MD characterization of denatured conformations remains challenging in part because of the inaccuracies of the molecular force fields that have been calibrated to reproduce properties of folded biomolecules. For example, long time-scale simulations of small proteins and nucleic acids have shown that ensembles of denatured conformations obtained using either CHARMM or AMBER force fields are overly compact in comparison to the conformations observed in experiment[29–43].

The advent of massively parallel supercomputers have enabled simulations of large biomolecular assemblies[25,44–48] where multiple biomolecules—proteins, nucleic acids, lipids and polyglycans—form a functional unit that is held together by non-covalent interactions. All-atom simulations of such biomolecular assemblies have been an important milestone for the development of the MD method, setting the stage for truly molecular-level description of the most fundamental biological processes, such as transcription, translation, virus maturation, etc. However, quantitative studies of DNA–protein and DNA–DNA interactions revealed considerable problems with the standard parameterization of the molecular force field, which typically produce unrealistic aggregation of biomolecules[31,41,43,48–50].

Largely, the MD community is in agreement that the approximate description of molecular interactions employed by the fixed-charge models strengthens attractive solute–solute interactions, which causes artificial aggregation of denatured proteins or multimeric assemblies[29,32,35,39,41,42,53,54,60]. There is, however, no agreement on the best way for improving the force field models. One possibility is that the overly attractive interactions between solutes derive from the shortcomings of the standard water models, such as TIP3P and its variants[61]. Realizing that the older water models generally underestimate the solute–water binding affinity and instead promote solute–solute bindings, several new water models have been developed[39,62] while the older water models have been modified[32,35,63]. Using these improved water models in MD simulations produces more realistic conformations of denatured proteins[39] and intrinsically disordered peptides[35] and reduces artificial aggregation of amino acid monomers[32,54]. However, modification of the water model

requires significant validations as it affects *all* solute–solute interactions. Indeed, the improved water models that produce expanded conformations of denatured proteins also destabilize the folded conformations of the proteins[32,39] and may prevent protein folding[42]. In the case of nucleic acids, the effect might be amplified by the already underestimated strength of hydrogen bond interactions[64].

In this perspective, we discuss an alternative approach for improving the MD force fields that surgically corrects pair-specific Lennard-Jones (LJ) parameters without modifying solute–water interactions. This approach is usually referred as NBFIX (Non-Bonded FIX) following the convention of the CHARMM package[55], although NBFIX equivalents also exist for the AMBER and Gromacs packages (LJEDIT and nonbond_param, respectively). The need for NBFIX arises from the parameterization of LJ interactions employed in the biomolecular force fields[4,66]. An empirical force field defines more than a hundred different atom types to described biomolecules. For example, the CHARMM36 force field has 124 atom types in total (53, 42, and 29 atom types for proteins[7], nucleic acids[14], and lipids[10], respectively), and thus requires 7,626 sets of LJ parameters to describe all possible pair-wise LJ interactions. For convenience, only 124 atom-specific sets of LJ parameters are used to described solute–water interactions that directly affect the hydration free energy, whereas LJ parameters for solute–solute atom pairs are *estimated* according to the atom-specific LJ parameters and the Lorentz-Berthelot combination rule[67,68]. The LJ parameters for the AMBER force fields have been developed in a similar manner. The NBFIX correction overrides the default assignment of the LJ parameters and allows one to assign custom values to the LJ parameters of a specific atom pair. For example, Bernèche & Roux used NBFIX to correct LJ interactions between Na^+ and K^+ ions with carbonyl oxygen to simulate ion selectivity of the KcsA channel[55].

Development of NBFIX corrections requires reference experimental data. Osmotic pressure data for binary solutions[63,69,70] have been used extensively for NBFIX development as such data can be easily and accurately obtained both experimentally and computationally. Experimentally, osmotic pressure can be determined using either a conventional semipermeable membrane osmometer or modern osmometers based on the freezing point or vapor pressure measurements[71]. Computationally, osmotic pressure can be determined through an MD simulation using a two-compartment setup (see Section 2.1)[70]. Experimental osmotic pressure is known for most common solutions of electrolytes[69] and amino acids[72–76].

Thus, realizing that binding affinity of the Na–Cl and K–Cl ion pairs is overestimated, Luo and Roux calibrated the Na–Cl and K–Cl ion pair LJ interactions to reproduce the experimental osmotic pressure of the NaCl and KCl solutions[56], a procedure that we refer hereinafter as the NBFIX approach. Subsequently, osmotic pressure-calibrated NBFIX corrections have been developed for several other interaction types as summarized in Fig. 1 and Table 1 and 2[31,41,51,53,54,59]. The advantage of the NBFIX approach is that the calibration is surgically applied only to specific atom pairs, making the introduction of side effects, such as unintended destabilization of hydrogen bonds[32,35,39], unlikely. Thus, NFBIX calibration has been shown to have no affects on the hydration free energy, solution density, interactions with other solutes, and bonded parameters (e.g., backbone torsions and NMR J

coupling parameters in proteins)[42]. Below, we review the improvements brought to the MD method by the NBFIX approach and discuss its limitations and challenges.

2 Physical chemistry of the NBFIX approach

In the case of an ideal solution where osmolytes (e.g., solutes or ions) do not interact with each other, osmotic pressure, π, is linearly proportional to the solute concentration, c: $\pi(c) = cRT$ where R and T are the gas constant and temperature, respectively. In the case of interacting osmolytes, the osmotic pressure of a solution can be related to its concentration by introducing a concentration-dependent factor, the osmotic coefficient $\phi(c)$, $\pi(c) = \phi(c)cRT$. Broadly speaking, ϕ values smaller than 1 indicate attractive interactions between the osmolytes, which decrease the effective osmolyte concentration to ϕc as the osmolytes form complexes[71]. Conversely, ϕ values larger than 1 indicate repulsive interactions that effectively increase the solute concentration to ϕc. Thus, osmotic pressure (or equivalently osmotic coefficient) data present a convenient target for the calibration of solute–solute interactions.

2.1 Calibration of solute–solute interactions using osmotic pressure data

The first step in the calibration procedure is choosing a target interaction. For an illustrative example, here we choose the interaction between positively charged amine and negatively charged carboxylate groups,[41] which are common functional groups of amino acids (e.g., lysine, glutamate, and aspartate) and lipids (e.g., phosphatidylserine), Fig. 2A. First we note that using entire amino acids or lipid molecules for calibration is not ideal because the amine–carboxylate interaction are not guaranteed to dominate the overall inter-solute interactions. A much better alternative is using a methylammonium–acetate or glycine solution where amine–carboxylate interactions dominate over all other interactions. Here, we chose the glycine solution because experimental osmotic pressure data are available for a range of solution concentrations[72]. Similarly, methylguanidinium and dimethylphosphate can serve as good models for calibration of interactions between arginine side chain and phosphate groups of lipids or nucleic acids, respectively Fig. 2A, whereas carbohydrate monomers can be used to calibrate the interactions between carbohydrate groups of the nucleic acid backbone[59].

The osmotic pressure of a solution can be computed using a two-compartment simulation setup, Fig. 2B,C, where a rectangular simulation volume is divided, under periodic boundary conditions, into two compartments separated by two virtual semi-permeable membranes (dashed lines in Fig. 2B,C). The effect of the two semi-permeable membranes is typically realized by using two planar half-harmonic potentials that exert forces on non-hydrogen atoms of the solutes. The presence of solutes only in one of the two compartments—the solute compartment—generates the osmotic pressure. At equilibrium, the pressure exerted by the semi-permeable membrane on the atoms of solutes equals to the osmotic pressure in the solute compartment[56,70] and can be conveniently evaluated by recording the forces applied to the solutes by the half-harmonic potentials.

When simulated using the standard CHARMM36 force field, a ~3 m glycine solution is observed to form clusters, Fig. 2B, which is inconsistent with the high experimental

solubility of glycine that exceeds 3 M at room temperature and the near-ideal dependence of the experimental osmotic pressure on glycine concentration, Fig. 2D. Accordingly, the simulated osmotic pressure is half or even less of the experimental value over a range of glycine concentration, Fig. 2D, indicating an overestimation of the attractive interaction between amine and carboxylate groups. Although the attractive interaction is of the electrostatic nature, modifying the partial charges of the amine or carboxylate groups would lead to uncontrolled changes in the solvation energy and charge–charge interactions with other groups of the force field. An alternative approach is to change the interaction that opposes the electrostatic attraction—the steric repulsion between the amine nitrogen (N) and carboxylate oxygen (O) described by the LJ potential:

$$U_{\mathrm{LJ}}=\varepsilon\left[(R_{\min}/r)^{12}-(R_{\min}/r)^{6}\right]=4\varepsilon\left[(\sigma/r)^{12}-(\sigma/r)^{6}\right] \quad (1)$$

where r indicates the N–O distance and $R_{\min}$ (= $2^{1/6}\sigma$) and ε are pair-specific LJ parameters. Indeed, increasing the LJ R_{min} (or equivalently σ) parameter for amine nitrogen-carboxylate oxygen pairs gradually increases the osmotic pressure of a glycine solution at a fixed concentration and alleviates the clustering problem, Fig. 2C[41]. Increasing the R_{min} value by only 0.08 Å brings the simulated osmotic pressure values in agreement with experiment Fig. 2D[41]. Fig. 2E displays the amine nitrogen–carboxylate oxygen LJ potentials before and after the NBFIX correction. Fig. 2F characterizes the effect of NBFIX correction on the radial distribution function (RDF) of the inter-molecular nitrogen-oxygen pair: increasing $R_{\min}$ monotonically decreases the height of the first RDF peak, g_1. By plotting $-k_{\mathrm{B}}T\log(g_1)$ as a function of $\Delta R_{\min}$, we obtain an estimate of the free energy change introduced by the NBFIX correction, Fig. 2G. Thus, $\Delta R_{\min}$ of 0.08 Å increases the free energy of a contact ion pair by about 0.7 kcal/mol with respect to the standard value.

2.2 Why NBFIX works?

Although the NBFIX correction does not explicitly alter the solute-water interactions, it affects the behavior of water molecules in proximity of the solutes. Specifically, the NBFIX correction affects the ability of water molecules to mediate the solute-solute interaction, controlling the relative abundance of contact ion pairs (CIP), Fig. 3A, and solvent-shared ion pairs (SIP), Fig. 3B. We illustrate this effect here by analyzing the simulations of an ammonium–acetate pair, for which the NBFIX correction was derived in Ref. 41. Fig. 3C shows the potential of mean force (PMF) between an ammonium and acetate in pure water computed using the umbrella sampling method [78,79] and the AMBER ff99 force field with and without the NBFIX correction. In the absence of NBFIX, the free energy of a CIP configuration is about 1 kcal/mol lower than that of a SIP configuration, Fig. 3C, whereas the two configurations have similar energy in the NBFIX-optimized simulation. Thus, the NBFIX correction destabilizes the CIP state, increasing the population of SIP states. NBFIX corrections to the CHARMM force field have similar effects on the energetics of CIP to SIP[41].

To examine the destabilization of CIP by the NBFIX correction more quantitatively, we simulated the ammonium–acetate system having the distance between the ammonium nitrogen (N) and carboxylate carbon (CC) restrained to 4.2 Å using a weak harmonic potential (spring constant $k = 0.5$ kcal/mol·Å^2). The equilibrium restrain distance (4.2 Å) chosen for these simulations corresponds to a local maximum in the PMF that separates the CIP and SIP states, Fig. 3C. During the simulations, the system was observed to undergo frequent transitions between the CIP and SIP states, Fig. 3D,E. In the simulation performed using the standard force field model, the CIS state was dominant (72% of all conformations). The application of the NBFIX correction changed the relative occupancy of the CIS and SIP states, making the SIP state more probable (60% of all conformations), Fig. 3E.

To elucidate the role of water, we counted the number of water molecules that simultaneously formed hydrogen bonds with both ammonium and acetate groups, Fig. 3B. The average number of such shared water molecules exhibits a strong correlation with the N–CC distance, Fig. 3F. Most data points obtained using the NBFIX correction have the number of shared water molecules larger than 0.3 (140 out of 154 points) and N–CC distance larger than 4.2 Å (146 out of 154 points), Fig. 3F. Thus, our analysis suggests that the primary effect of the NBFIX correction on a solute pair is in changing the probability of having a water molecule between the two solutes. Consequently, the presence of a shared water molecule between the solutes increases the solute–solute distance roughly by the size of a water molecule.

2.3 Potential concerns with the NBFIX approach

Although the NBFIX corrections developed using the experimental osmotic pressure data of model solutes have so far been found to be transferable to larger biomacromolecules, e.g., nucleic acids[31,41,43], proteins[42,58], and lipid bilayer membranes[41,57], there are several potential concerns about the NBFIX approach.

First, the NBFIX corrections might be overfitted to the osmotic pressure data and hence require careful cross-validation against diverse experimental data sets. Potential target data sets for NBFIX validation include Ramachandran plots, NMR J-coupling parameters, and X-ray scattering data[8,35,42]. Second, all current NBFIX corrections (see Table 1 and 2) are applied to pair-specific LJ interactions, not partial charges, largely for the convenience. Given the fact that both CHARMM and AMBER force fields have a well-defined procedure for the derivation of solute partial charges and solute–water LJ parameters whereas solute–solute LJ parameters are largely estimated without validation, modifying pair-specific solute–solute LJ parameters appears to be a logical approach. It is, nevertheless, true that a better optimization can be achieved by modifying both LJ parameters and partial charges. Third, the transferability of the NBFIX corrections from one chemical context to the other is not guaranteed and requires additional validation. For example, although the standard CHARMM and AMBER force fields prescribe the same LJ parameters for the car-boxylate and phosphate oxygen atoms, reproducing osmotic pressure data for both required different sets of NBFIX corrections[41]. Thus, NBFIX corrections for the same chemical groups can be different depending on their chemical context.

3 Improvements and challenges

3.1 Biological electrolyte solutions

The physical properties of electrolyte solutions is a topic of growing interest for both physical chemists and biophysicists[71]. Many inorganic and organic salts dissociate in water into positively and negatively charged ions, cations and anions, making the electrolyte solution electrically conducting. The interaction of cations and anions with water molecules and with each other affects the structure and dynamics of the electrolyte solutions and determine their physical properties such as ionic conductivity, osmotic and activity coefficients. Furthermore, the valence and concentration of ions are known to considerably affect the properties of biomacromolecules. For example, the strength of salt bridges in proteins is determined by the complex interactions of ionizable side chains, cations, and anions[80]. Below, we discuss the application of the NBFIX approach toward improving the classical MD description of ion effects and discuss remaining challenges.

3.1.1 NBFIX corrections improve description of ion pairing—The key process determining the structure and dynamics of an electrolyte solution is ion pair formation[63,77,81]. Although various spectroscopic and scattering methods have provided important insights into the behavior of ions and water leading to ion pair formation[82–84], these methods cannot presently provide direct information about the structural dynamics of ion pairs as it occurs at the picosecond and angstrom time and length scales[85]. The all-atom MD method can thus provide unique information about the process of ion pair formation, including the properties of the ions' hydration shells[63,81,86,87]. Here, we examine the effect of NBFIX corrections on ion pairs containing functional analogs of biomolecules such as ionizable side chains and nucleic acid backbone. Following that, we briefly discuss the present status of NBFIX corrections for monatomic ions.

As previously discussed (see Section 2.2 and Fig. 3), the effective interaction between two oppositely charged ion species sensitively depends on the microscopic configuration of water molecules at short (2–5 Å) ion–ion separations. At such separations, the microscopic configurations adopted by an ion pair can be classified as either forming a direct contact (CIP) or sharing a water molecule (SIP). Furthermore, the energetics of an ion pair formation is conditioned by the relative abundance of the CIP and SIP states. Fig. 4 shows the free-energy of an ion pair as a function of ion–ion distance computed using the standard CHARMM36 force field for six biologically relevant ion pairs: sodium–acetate (Fig. 4A), methylammonium–acetate (Fig. 4B), methylguanidinium–acetate (Fig. 4C), sodium–dimethylphosphate (Fig. 4D), methylammonium–dimethylphosphate (Fig. 4E), and methylguanidinium–dimethylphosphate (Fig. 4F). Note that acetate, methylammonium, and methylguanidinium, and dimethylphosphate ions are analogous to functional groups of ionizable side chains, nucleic acid backbone, and lipid head groups, Fig. 2A. Just as with ammonium–acetate (Fig. 3), the standard parameterization favors the CIP state over the SIP state for all ion pairs except sodium–dimethylphosphate, with the maximum free energy difference of −2 kcal/mol, Fig. 4C. The excessive stability of CIPs leads to artificial clustering of ions in MD simulations performed using standard force fields[31,41].

For comparison, we plot in Fig. 4 the free energy of the ion pairs computed using the NBFIX corrections to the CHARMM36 force field. The application of corrections reduces stability of CIPs by 1 – 3 kcal/mol but does not affect stability of SIPs. As a result of the corrections, the SIP state is energetically more stable than the CIP state for all six ion pairs. Because of the reduced energy barrier between CIP to SIP, the application of the NBFIX corrections also increases the exchange rate between the two states.

An alternative approach to correcting the artificially strong attraction between molecular ion pairs is the modification of their partial atomic charges. Thus, the Shaw group modified the partial charges of the side chains of guanidinium, acetate, and glutamate residues in the CHARMM22 force field to reduce the attractive interactions of arginine side chain with aspartate and glutamate side chains, producing a force field variant known as CHARMM22*[30]. To determine how the two approaches compare, we computed the PMF for methylguanidinium–acetate and methylammonium–acetate pairs using CHARMM22*; the results are shown as black curves in Fig. 4B,C. Similar to the effect of the NBFIX corrections to CHARMM36, the charge reduction in CHARMM22* reduces stability of CIPs without affecting stability of SIPs, however, the CIP states remains the global energy minimum for both ion pairs. At the quantitative level, the free energy of the CIP state is 0.5 and 2 kcal/mol more favorable in CHARMM22* than in NBIFX-CHARMM36 for the methylammonium–acetate, Fig. 4B, and methylguanidinium–acetate, Fig. 4C, pairs, respectively. Thus, the two calibration approaches produce qualitatively similar results that nevertheless differ substantially at a quantitative level.

For a pair of monatomic ions, the refinement of ion-ion interactions using the NBFIX approach is straightforward because each ion carries a well-defined electrical charge. The NBFIX corrections for ion pairs involving a Cl^- ion and one of the monovalent cations (Li^+, Na^+, K^+) were previously reported for both CHARMM36 and AMBER99 force fields[31,56]. The NBFIX corrections for ion pairs containing larger monatomic ions such as Rb^+, Cs^+, Br^-, and I^- are reported in this article, see Fig. S1 to S7 for the results of the calibration simulations. For divalent monoatomic cations such as Mg^{2+} and Ca^{2+}, addressing the polarization effect of first solvation water molecules was as important as the refinement of Lennard-Jones parameters[88]. Treating Mg^{2+} and Ca^{2+} as hexahydrate and heptahydrate complexes, respectively, and adjusting the dipole moment of the first solvation shell water molecules turned out to be an effective approach for implementation of the polarization effects in the framework of non-polarizable force fields[31,51]. Further optimization of the hydrated Mg^{2+} and Ca^{2+} ion models was done by adjusting the LJ parameters describing the interactions of the first solvation shell oxygen and anions[31,51]. Table 3 and 4 provides a summary of all NBFIX corrections developed by Yoo and Aksimentiev.

3.1.2 Parameterization of guanidinium group leaves room for improvement—

To illustrate the limitation of the NBFIX approach, here we review NBFIX calibration of the guanidinium–acetate interaction—a proxy for evaluation of non-bonded interactions between positively charged arginine side chains and anionic moieties, such as the side chains of aspartic or glutamic acids. The calibration was performed using a two-compartment system similar to that shown in Fig. 2B,C. The osmotic pressure computed using the standard CHARMM36 force field (red marks in Fig. 5A) was significantly lower than

prescribed by experiment[89]. For example, 139 bar was measured in experiment at 2 m concentration whereas the simulated osmotic pressure was only 18 bar. The NBFIX correction was introduced by gradually increasing the LJ R_{min} (or equivalently σ) parameter for the guanidinium nitrogen–carboxylate oxygen atom pair.

Increasing R_{min} by 0.08 Å, which was sufficient to bring the non-bonded interaction between amine and carboxylate groups in accord with experiment (see Table 4), markedly increased the osmotic pressure of the guanidinium–acetate solution (orange marks in Fig. 5A), however, the simulated pressure remained less than a half of the experimental value (black marks in Fig. 5A). Increasing the R_{min} value further increased the simulated osmotic pressure, however, no change in osmotic pressure was observed beyond ΔR_{min} = 0.20 Å, Fig. 5A. The saturation occurs because of the finite range of the effect of NBFIX corrections on the formation of contact ion pairs: once all contact ions pairs have been removed, further increase in R_{min} value has little effect on the behavior of the system (see Section 3.1.1). At concentrations lower than 1 m, ΔR_{min} of 0.20 Å is seen to produce a reasonable agreement with experiment, see inset of Fig. 5A. Thus, the NBFIX correction alone was not sufficient to bring the simulated osmotic pressure of the system in accordance with experiment at higher electrolyte concentrations.

The calibration simulations performed using the AMBER99 force field, Fig. 5B, revealed similar limitations of the NBFIX approach. The significantly underestimated osmotic pressure obtained using the standard parameterization (red marks in Fig. 5B) could be gradually increased by increasing the σ parameter of the LJ potentail. However, no further improvement could be made beyond $\Delta\sigma$ = 0.20 Å.

Importantly, the experimental osmotic pressure of the guanidinium–acetate solution is much higher than that of an ideal solution, Fig. 5A, suggesting that guanidinium and carboxylate groups at least do not form tight CIP complexes. Accordingly, the CIP state of the methylguanidinium–acetate pair is not even a local minimum of the ion pair's PMF computed using the NBFIX correction, blue line in Fig. 4C. At the same time, the osmotic pressure of the glycine solution is slightly lower than that of an ideal solution, Fig. 2D, suggesting that the interaction between ammonium and carboxylate groups is slightly more attractive than in the ideal non-interacting case. Accordingly, we find the CIP state to be a local minimum of the methylammonium–acetate PMF, blue line in Fig. 4B. Thus, our NBFIX corrections capture the overall difference between interactions of acetate with guanidinium and amine groups. To describe such a difference more accurately, it might be necessary to explicitly consider the sphere–plane and plane–plane geometry of ammonium–acetate and guanidinium–acetate interactions.

The discrepancy between the computed and experimental osmotic pressure of the guanidinium–acetate solution at a higher concentration highlights the limitations of the NBFIX appoach. The similarity of limitations observed for both AMBER99 and CHARMM36 force fields suggests that the problem could be intrinsic to the standard water models, TIP3P and modified TIP3P for the AMBER99 and CHARMM36 force fields, respectively. Calibration simulation using different water models could be essential for simulations of high-concentration guanidinium–acetate solutions. Another possible source of

discrepancy is the fundamental limitation of the non-polarizable models in describing the π–π stacking interactions[90–92] between guanidiniums in planar configurations[93]. Computing the osmotic pressure of the guanidinium–acetate solution using a polarizable force field model[94,95] is an interesting direction for future research efforts in this area. Corrections that affect formation of solvent-shared ion pairs might also be necessary to address the remaining discrepancy.

In summary, recent MD studies demonstrate that all-atom MD simulations equipped with a refined force field can quantitatively reproduce the thermodynamic properties of electrolyte solutions (e.g., osmotic pressure[31,63,86]) and some thermodynamics effects of ions on biomolecules (e.g., condensation of DNA[31,41]). However, it is still an open question whether MD simulations can provide an atomistic description of a thermodynamic observation. An outstanding challenge is the simulation of the Hofmeister series[96,97]. Although the “salting in” and “salting out” phenomena have been known and used in practice for more than a hundred of years, we still do not have a complete theoretical model explaining their physical origin[97]. A careful parameterization of all ions in the Hofmeister series using the NBFIX approach could provide an opportunity for resolving this long-standing puzzle.

3.2 Nucleic acid systems

For efficient storage and gene regulation, DNA in living cells and viruses is present in a highly condensed state[98,99]. Experimental studies of DNA condensation *in vitro* have provided invaluable insights into the physics of DNA–DNA interactions, revealing surprising effects that originate from the DNA's polymer-like behaviors and complex electrostatics[100–105]. As each nucleotide of a DNA strand contains a negatively charged phosphate group at the strand's backbone, screening of the DNA charge by counter ions and ion mediation of DNA–DNA interactions are the key effects determining the structure and dynamics of DNA systems[106,107]. Below, we discuss development and application of NBFIX corrections for MD simulations of condensed DNA systems and illustrate remaining challenges using the Holliday junction (HJ) system.

3.2.1 NBFIX corrections increase realism of DNA simulations—

DNA condensation *in vitro* can be produced through the application of external pressure or by condensing agents such as positively charged polyamine ions. Regardless of the condensation method, the DNA helices adopt an ideal hexagonal packing in the condensed state, with the average inter-DNA distance being a well-defined function of the external pressure and ionic condition[100,101]. Unlike the DNA structure, the concentration and arrangement of couterions within such *in vitro* condensates is not experimentally known, which makes them an attractive target for MD studies[108].

Prior to the development of NBFIX corrections to ion–DNA phosphate interactions[31,41,51], MD characterization of DNA condensates had been challenging. At the level of a single DNA helix, both standard CHARMM and AMBER force fields overestimate the binding affinities of cations to DNA phosphate groups, which is a consequence of overstabilization of the CIP states[31]. For a pair of DNA helices, the artificially strong interactions between cations and DNA phosphates facilitate formation of phosphate–cation–phosphate bridges

and produce attractive inter-DNA forces even for monovalent and divalent electrolytes[109] that do not condense DNA in experiment[100]. In the simulations of DNA array systems[43], such as the one shown in Fig. 6A, the overly favorable interaction between DNA helices produces artificial clustering of DNA helices and significantly underestimate the internal pressure, Fig. 6B–D, inconsistent with experimental observations[100,101].

The application of NBFIX corrections brings MD description of ion atmosphere of DNA in agreement with experiment. In comparison to standard parameterization, the NBFIX corrections reduce the number of cations directly bound to the phosphate groups of the DNA backbone, increasing the population of loosely bound ions in the atmosphere. Direct quantitative validation of ion atmosphere description is obtained through comparison of the simulation results to the ion counting data obtained using the buffer equilibration and atomic emission spectroscopy (BE-AES)[106]. Specifically, NBFIX-enabled MD simulations were found to reproduce results of the BE-AES measurements even in the case when two cation species, such as Li^+, Na^+, K^+, and Mg^{2+}, compete for DNA binding[107]. MD simulations of DNA translocation through solid-state nanopores[51,110] provided additional validation of ion type-dependent cation–DNA interactions[51,110].

NBFIX-enabled simulations of DNA array systems reproduce the experimental dependence of the internal pressure on the inter-DNA distance for a range of electrolyte compositions and concentrations Fig. 6B–D[31,43]. Inside an array, DNA helices are found to arrange on a regularly spaced hexagonal lattice[43], as prescribed by experiment[100,101]. Most importantly, the simulations identified indirectly bound bridging cations as the source of cohesive interactions that give rise to DNA condensation and characterized translational and rotational diffusion of DNA in such arrays, which is critical for modeling virus packaging and ejection[43]. The balanced description of ion binding to DNA enabled characterization of more subtle effects, such as sequence and methylation dependence of inter-DNA forces. Thus, simulation predicted AT-rich DNA helices to attract stronger than CG-rich ones and methylation of cytosines to strengthen such attraction. These predictions were validated by single molecule experiment[111].

Another example of dense DNA systems are self-assembled DNA nanostructures (e.g., DNA origami[112]) that consist of many DNA helices joined through multiple HJs. In such systems, overly strong inter-DNA interactions can cause mechanical distortions including excessive twisting or bending. It has been shown that the MD method equipped with NBFIX corrections can be used to characterize structural and mechanical properties of DNA origami[113] and DNA bricks[114] objects. Such simulations are particularly suitable for characterization of the transport properties of DNA channels[115–119], where NBFIX corrections are essential to capture the ionic conductivity, electro-osmosis, and diffusion of small solutes.

3.2.2 MD simulations of Holliday junctions exemplify challenges in capturing conformation dynamics—MD simulation of a HJ system, Fig. 7A[121], presents a stringent yet straightforward test of an MD force field. In this system, four DNA double helix arms interact with one another through two types of interactions: ion-mediated electrostatics of side-by-side DNA duplexes and base-stacking at the crossing of the four

duplexes. Depending on ionic conditions, a HJ is known to adopt several conformations that are well characterized through biochemical[121] and single-molecule[122,123] studies, Fig. 7B–D. At a low concentration of monovalent cations and in the absence of multivalent cation, electrostatic side-by-side repulsion wins over base-stacking attraction. As a result, a HJ adopts an open conformation, in which four arms are fully extended in a square-planar configuration, Fig. 7C[122,124,125]. At a high concentration of monovalent cations (> 1 M) or in the presence of multivalent cation, side-by-side electrostatic repulsion is screened while the base-stacking interactions remain unaffected, Fig. 7B,D. As a result, a HJ adopts a right-handed antiparallel stacked-X configuration, Fig. 7B,D[126]. Depending on the the stacking partner, two isomeric stacked-X configurations (iso I and II) are possible [121,123]

Using the AMBER ff99bsc0 force field[13] and the NBFIX corrections for ion-phosphate interactions[31], we investigated the effect of ion conditions on the conformation of a HJ. Starting from the crystal structure of an open configuration of a HJ (PDB: 2QNC)[120], Fig. 7A, we built an all-atom model of the junction having the nucleotide sequence of the experimentally studied system[123]. The all-atom model was submerged in a dodecahedron volume of TIP3P water[61] ($a = b = c = 10$ nm $\alpha = \beta = 60°$ $\gamma = 90°$) containing a neutralizing amount (84) of sodium ions, which corresponds to a low ion concentration condition that, in experiment, strongly favors an open configuration of the junction. During the first 120 ns, the junction maintained its open configuration, Fig. 7E, but soon after underwent an unexpected conformational transition to a stacked-X conformation, Fig. 7E. A much more rapid transition to a stacked conformation was observed in fourteen additional simulations carried out in 50 mM Mg electrolyte, Fig. 7F. The initial conformations for these simulations were randomly selected from the first 120 ns of the HJ trajectory at low Na concentrations. In 13 out 14 such simulations, the con-formational change to stacked-X conformations occurred within 10–50 ns. Among the 13 stacked-X conformations, we found 3 right-handed iso I, 4 right-handed iso II, 3 left-handed iso I, and 3 left-handed iso II conformers.

The above simulations of the HJ system revealed two problems. First, a stable stacked conformation was observed in MD simulations at low sodium conditions whereas experiment indicates the dominance of open conformation even at 400 mM Na[121–123]. Second, six out of thirteen stacked conformations at 50 mM Mg (and the stacked conformation at low sodium) were left-handed, whereas experiment found left-handed stacked conformations to be unlikely[121,126]. Previously, we found NBFIX-enabled ion-mediated electrostatic interactions between DNA helices to provide quantitatively accurate description of experiments, Fig. 6[43]. Thus the disagreement between simulation and experiment is likely to originate from inaccurate description of stacking forces between DNA bases[34,127–129]. Indeed, an overestimation of base-stacking forces can increase the population of left-handed stacked-X conformations that have more steric contacts between the DNA arms than the right-handed or open conformations.

A potential problem with the present parameterization of the nucleic acids is the description of the nonbonded interactions of their sugar groups. Recently, the Elcock group showed that carbohydrate monomers attract one another too strongly when simulated using standard parameters and developed a set of NBFIX corrections for both the CHARMM and AMBER force fields that brought such interactions in agreement with experimental observations[52,59].

It is thus possible to improve the description of the nucleic acid carbohydrate groups in the CHARMM and AMBER force fields by adopting the NBFIX corrections developed by the Elcock group. A combination of NBFIX corrections describing non-bonded interactions of backbone phosphates[31,41], sugars[52,59] and base-stacking[34] together with refinement of bonded interactions[13,14,130] may produce the most accurate computational model of nucleic acids.

3.3 Protein folding

The all-atom MD method has been instrumental in elucidating the mechanisms of protein folding-unfolding transitions[23,24] and even prediction of the folded structure of short globular proteins[131]. It has been, however, established that conformations of unfolded or disordered proteins are too compact in MD simulations in comparison to experimental conformations[29–42].

Several types of molecular interactions drive folding of a polypeptide chain into a unique 3D structure[132]. Among those interactions, salt bridges between amine and carboxylate groups[80,133] and hydrophobic interactions between nonpolar residues[134,135] have been found to be too strong in the standard CHARMM and AMBER force fields[32,41] whereas interactions between polar groups appear to be accurate[41].

To elucidate the effect of NBFIX corrections on protein simulations, we used the replica-exchange MD (REMD) method[42] to simulate folding of a villin head piece[136] and WW domain[137]. Our simulations demonstrated that the NBFIX corrections to AMBER ff99sb-ildn-phi force field reduce artificial aggregation of both proteins in their disordered states[42]. Despite such reduction, both proteins fold into correct 3D structures, whereas the folding time is reduced by a factor of two, on average, in comparison to the results obtained using the standard force field[42].

Although corrections to non-bonded parameters have improved sampling of disordered conformations[37,42], it has been argued that the radius of gyration, R_g, of a disordered protein is still smaller in an MD simulation than in experiment[37,38]. In particular, the Shaw group found the simulated R_g values of unfolded proteins to be roughly half the value of chemically denatured ones for proteins containing from 10 to 90 residues[139]. For example, the experimental R_g value of a 76-residue ubiquitin protein is 27 Å[139] at 6 M GuHCl whereas the R_g value of unfolded ubiquitin at 390 K computed using the CHARMM22* force field is only ~12 Å[37,140].

It is, however, unclear if the above comparison is fair as the conditions realized in experiment and simulation differ substantially. Unlike an MD simulation of an unfolded protein, which typically is performed at an elevated temperature[37,140], an experimental ensemble of denatured proteins is realized through addition of detergent at high concentration (typically in excess of 5 M). For a better comparison, we performed an REMD simulation of unfolded ubiquitin in pure water (0 M urea) and in 8 M urea solution using the AMBER ff99sb-ildn-phi force field and NBFIX corrections[41,42]. The R_g value in the absence of urea converged to a value slightly below 20 Å (blue line in Fig. 8), which is about 10 Å smaller than the experimental R_g value at 6 M GuHCl[139] (black line in Fig. 8).

However, simulated R_g of ubiquitin at 8 M urea is close to the experimental value, Fig. 8. Although the above result is encouraging, it is still possible that NBFIX-enabled force fields underestimate R_g values of unfolded or intrinsically disordered proteins[40].

3.4 Lipid membranes

Phospholipid is a major class of lipids that are the building blocks of cell membranes. Each phospholipid (hereinafter lipid for brevity) contains two hydrophobic hydrocarbon tails and a polar or charged head group. The head groups in particular are known to interact with membrane and membrane-associated proteins through electrostatic interactions involving negatively charged phosphate and carboxylate groups and positively charged amine and choline groups. Thus, it is natural to expect NBFIX corrections to have a major effect on the simulations of lipid bilayers and protein–lipid systems.

Indeed, we found the application of the NBFIX corrections to amine–phosphate and amine–carboxylate interactions to loosen the packing of lipid molecules in a lipid bilayer membrane[41]. The CHARMM27r lipid force field is known to underestimate the simulated area per lipid molecule, suggesting that the head groups of the lipid molecules attract one another too strongly[9]. The application of the NBFIX corrections reduces attraction between the head groups, bringing the simulated area per lipid molecule in close agreement with experiment for both phosphatidylethanolamine (PE) and phosphatidylserine (PS) lipid bilayers[41]. As all lipid types contain electronegative groups at high concentrations, their structure and dynamics is affected by the presence of cations. As the introduction of NBFIX corrections improves description of cation–carboxylate and cation–phosphate interactions, it can contribute to more accurate simulations of the cation effects on lipid membrane structure and dynamics[41,57].

Just like with development of corrections to amino acid side chains (recall Fig. 4B,C and related paragraphs), there could be multiple different approaches for improving parameterization of lipid head groups. One such approach is development of NBFIX corrections[41,57]. An alternative approach is modification of partial charges, similar to that used by the Shaw group for development of CHARMM22*. With regard to simulations of lipid bilayers, Berger and coworkers have shown that modifications of head group partial charges in the Gromos lipid force field brings the simulated area per lipid in agreement with experiment[141]. Yet another approach is modification of torsion parameters, which was used in the development of the CHARMM36 lipid force field[10]. As all three approaches can reproduce the correct area per lipid values, it is premature to conclude which of the specific approaches is better than the others. A next generation lipid force field might as well be produced through combination of all these approaches after systematic validation against experimental data.

3.5 Molecular recognition and signaling

Regulation of biochemical processes is paramount to survival, growth, and development of living cells. At the molecular level, regulation of cellular processes relies on physical recognition of signaling ligands by their biomolecular targets. Molecular mechanisms of such recognition have been an area of active research for the MD community[144]. Because

molecular recognition is largely electrostatic in nature, the application of NBFIX corrections to charge-charge interactions could have a major effect on the outcome of a molecular recognition simulation.

The application of NBFIX corrections can be expected to provide more accurate simulations of ligand docking or protein–protein assembly, especially when mediated by salt bridges between amine and carboxylate groups. For example, the stability of β-amyloid fibrils implicated in the Alzheimer disease[145] and the ligand binding behavior of the extracellular domain of G-protein coupled receptors (GPCR)[146,147] are highly regulated by salt bridges between lysine and aspartate residues. The salt bridge formation is also essential for the function of ligand-gated ion channels and transporters[142]. For example, Fig. 9A shows the structure of a dopamine transporter (PDB: 4XP9), in which the bound dopamine is in direct contact with the carboxylate group of an aspartate side chain[142]. Several important neurotransmitters such as AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid), NMDA (N-methyl-D-aspartate), GABA (γ-aminobutyric acid), glutamate, and serotonin contain multiple amine and carboxylate groups[148,149], Fig. 9B. Based on our simulation data[41,42], we expect that the standard CHARMM and AMBER force fields are likely to overestimate the binding affinity of such ligands to their respective target sites and that applying NBFIX corrections may improve the simulation outcomes.

In the case of protein–DNA complexes, electrostatic interactions play an important role in sequence-specific binding of transcription factors via leucine zipper, zinc finger and helix-turn-helix mechanisms[150,151], DNA-binding domains of ATP-driven motor proteins[152], and DNA–histone assembly of nucle-osomes[153,154]. One can thus expect that NBFIX corrections to interactions between positively charged residues (Lys, Arg, His) and the phosphate groups of a DNA backbone will improve the accuracy of protein–DNA complex simulations. Indeed, we observed such improvement in simulations of lysine-mediated interactions between DNA molecules[41]. In qualitative disagreement with experiment[155], the simulations carried out using the standard CHARMM and AMBER models predict attractive interaction between DNA helices in the presence lysine monomers or trimers[41], which derive primarily from the overestimation of amine–phosphate attraction. Applying NBFIX corrections to the same simulations considerably improved the realism of the simulation outcome, reproducing, for example, a crossover from repulsive to attractive interactions with increasing length of polylysine peptides[41].

Considering that charged groups are functional moieties of the molecules involved in cell signaling, the NBFIX corrections to amine–carboxylate and amine–phosphate interactions might be of relevance to simulations of cell signaling. The interactions between charged amino acids and lipid head groups are essential for maintaining the structure[156–158] and biological function[159,160] of channel proteins. One celebrated example is the voltage sensor of potassium or sodium channels, in which guanidinium–phosphate interactions play a central role[159,161].

Lipid molecules containing phosphatidylinositide (PI) groups have been implicated in cell signaling[160,162–165] and have recently become a subject of MD studies[166,167]. Accurate simulations of such systems will require validation of interactions between the phosphate

groups of PI lipids and the charged side chains of the signaling proteins. For example, the crystal structure of inositol 1,3-diphosphate bound to EEA1 homodimer of C-terminal FYVE domain shows a coordination of a PI group by four different arginine side chains, Fig. 9C. The NBFIX corrections developed to describe the amine–phosphate and guanidinium–phosphate interactions should provide a reasonable initial parameterization for the MD study of PI lipid molecules and cell signaling.

4 Conclusion

Over the past decades, the field of MD simulations has been experiencing a tremendous growth, fueled by continuous efforts of the MD community to improve accuracy and efficiency of the MD method. Until recently, significant efforts were put into refinement of the chemical bond interactions through systematic quantum mechanical calculations and comparison to experimental structural data. Presently, the MD community recognizes the urgent need for refinement of the non-bonded interactions. One particular challenge is to balance the strength of non-bonded interactions without compromising stability of the native structures. In this perspective, we reviewed one of the possible solutions to that challenge—the NBFIX approach, which has the advantage of offering a significant improvement to the force field accuracy by introducing minimal adjustments to the force field parameterization. However, it remains to be seen if the NBFIX approach can indeed offer solutions to the majority of the remaining force field problems or if the MD community will need to build a drastically different force field model. Answering this open question will require critical assessment of force fields accuracy, development of experimental approaches tailored to the force field parameterization and a framework for systematic force field refinement and validation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the grants from National Science Foundation (PHY-1430124) and the National Institutes of Health (P41-GM104601). Supercomputer time was provided through The Extreme Science and Engineering Discovery Environment allocation MCA05S028 and the Blue Waters petascale supercomputer system at the University of Illinois at Urbana-Champaign. J.Y. was also supported by Institute for Basic Science (IBS-R007-Y1).

References

1. McCammon JA, Gelin BR, Karplus M. Nature. 1977; 267:585–590. [PubMed: 301613]
2. Moore GE. Electronics. 1965; 38:114–117.
3. Vendruscolo M, Dobson CM. Curr Biol. 2011; 21:R68–R70. [PubMed: 21256436]
4. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. J Am Chem Soc. 1995; 117:5179–5197.
5. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Proteins: Struct, Funct, Bioinf. 2006; 65:712–25.
6. Buck M, Bouguet-Bonnet S, Pastor RW, MacKerell AD. Biophys J. 2006; 90:L36–38. [PubMed: 16361340]

7. Best RB, Zhu X, Shim J, Lopes PEM, Mittal J, Feig M, MacKerell AD Jr. J Chem Theory Comput. 2012; 8:3257–3273. [PubMed: 23341755]
8. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE. Proteins: Struct, Func, Bioinf. 2010; 78:1950–8.
9. Klauda JB, Brooks BR, MacKerell AD Jr, Venable RM, Pastor RW. J Phys Chem B. 2005; 109:5300–11. [PubMed: 16863197]
10. Klauda JB, Venable RM, Freites JA, O'Connor JW, Tobias DJ, Mondragon-Ramirez C, Vorobyov I, Mac Kerell AD Jr, Pastor RW. J Phys Chem B. 2010; 114:7830–7843. [PubMed: 20496934]
11. MacKerell D Jr, Banavali NK. J Comput Chem. 2000; 21:105–120.
12. Foloppe N, MacKerell AD Jr. J Comput Chem. 2000; 21:86–104.
13. Perez A, Marchan I, Svozil D, Sponer J, Cheatham TE, Laughton CA, Orozco M. Biophys J. 2007; 92:3817–3829. [PubMed: 17351000]
14. Hart K, Foloppe N, Baker CM, Denning EJ, Nilsson L, MacKerell AD Jr. J Chem Theory Comput. 2012; 8:348–362. [PubMed: 22368531]
15. Denning EJ, Priyakumar UD, Nilsson L, MacKerell AD Jr. J Comput Chem. 2011; 32:1929–1943. [PubMed: 21469161]
16. Zgarbová M, Otyepka M, Šponer J, Mládek A, Banáš P, Cheatham TE, Jurečka P. J Chem Theory Comput. 2011; 7:2886–2902. [PubMed: 21921995]
17. Fadrna E, Spackova N, Sarzynska J, Koca J, Orozco M, Cheatham TE, Kulinski T, Šponer J. J Chem Theory Comput. 2009; 5:2514–2530. [PubMed: 26616629]
18. Zgarbova M, Sponer J, Otyepka M, Cheatham TE, Galindo-Murillo R, Jurecka P. J Chem Theory Comput. 2015; 11:5723–5736. [PubMed: 26588601]
19. Hess B, Kutzner C, van der Spoel D, Lindahl E. J Chem Theory Comput. 2008; 4:435–447. [PubMed: 26620784]
20. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. J Comput Chem. 2005; 26:1781–1802. [PubMed: 16222654]
21. Salomon-Ferrer R, GÃŭtz AW, Poole D, Le Grand S, Walker RC. J Chem Theory Comput. 2013; 9:3878–3888. [PubMed: 26592383]
22. Comer J, Phillips JC, Schulten K, Chipot C. J Chem Theory Comput. 2014; 10:5276–5285. [PubMed: 26583211]
23. Shaw DE, Maragakis P, Lindorff-Larsen K, Piana S, Dror RO, Eastwood MP, Bank JA, Jumper JM, Salmon JK, Shan Y, Wriggers W. Science. 2010; 330:341–346. [PubMed: 20947758]
24. Lindorff-Larsen K, Piana S, Dror RO, Shaw DE. Science. 2011; 334:517–520. [PubMed: 22034434]
25. Zhao G, Perilla JR, Yufenyuy EL, Meng X, Chen B, Ning J, Ahn J, Gronenborn AM, Schulten K, Aiken C, Zhang P. Nature. 2013; 497:643–646. [PubMed: 23719463]
26. Pierce LC, Salomon-Ferrer R, Augusto C, de Oliveira F, McCammon JA, Walker RC. J Chem Theory Comput. 2012; 8:2997–3002. [PubMed: 22984356]
27. Kutzner C, Pall S, Fechner M, Esztermann A, de Groot BL, Grubmuller H. J Comput Chem. 2015; 36:1990–2008. [PubMed: 26238484]
28. Theillet FXX, Binolfi A, Frembgen-Kesner T, Hingorani K, Sarkar M, Kyne C, Li C, Crowley PB, Gierasch L, Pielak GJ, Elcock AH, Gershenson A, Selenko P. Chem Rev. 2014; 114:6661–6714. [PubMed: 24901537]
29. Johnson ME, Malardier-Jugroot C, Murarka RK, Head-Gordon T. J Phys Chem B. 2009; 113:4082–4092. [PubMed: 19425247]
30. Piana S, Lindorff-Larsen K, Shaw DE. Biophys J. 2011; 100:L47–9. [PubMed: 21539772]
31. Yoo J, Aksimentiev A. J Phys Chem Lett. 2012; 3:45–50.
32. Nerenberg PS, Jo B, So C, Tripathy A, Head-Gordon T. J Phys Chem B. 2012; 116:4524–4534. [PubMed: 22443635]
33. Lindorff-Larsen K, Maragakis P, Piana S, Eastwood MP, Dror RO, Shaw DE. PLoS One. 2012; 7:e32131. [PubMed: 22384157]
34. Chen A, García AE. Proc Natl Acad Sci U S A. 2013; 110:16820–16825. [PubMed: 24043821]

35. Best RB, Zheng W, Mittal J. J Chem Theory Comput. 2014; 10:5113–5124. [PubMed: 25400522]
36. Petrov D, Zagrovic B. PLoS Comput Biol. 2014; 10:e1003638. [PubMed: 24854339]
37. Piana S, Klepeis JL, Shaw DE. Curr Opin Struct Biol. 2014; 24:98–105. [PubMed: 24463371]
38. Skinner JJ, Yu W, Gichana EK, Baxa MC, Hinshaw JR, Freed KF, Sosnick TR. Proc Natl Acad Sci U S A. 2014; 111:15975–80. [PubMed: 25349413]
39. Piana S, Donchev AG, Robustelli P, Shaw DE. J Phys Chem B. 2015; 119:5113–5123. [PubMed: 25764013]
40. Rauscher S, Gapsys V, Gajda MJ, Zweckstetter M, de Groot BL, Grubmüller H. J Am Chem Soc. 2015; 11:5513–5524.
41. Yoo J, Aksimentiev A. J Chem Theory Comput. 2016; 12:430–443. [PubMed: 26632962]
42. Yoo J, Aksimentiev A. J Phys Chem Lett. 2016; 7:3812–3818. [PubMed: 27617340]
43. Yoo J, Aksimentiev A. Nucleic Acids Res. 2016; 44:2036–2046. [PubMed: 26883635]
44. Freddolino PL, Arkhipov AS, Larson SB, McPherson A, Schulten K. Structure. 2006; 14:437–449. [PubMed: 16531228]
45. Biswas M, Voltz K, Smith JC, Langowski J. PLoS Comput Biol. 2011; 7:e1002279. [PubMed: 22207822]
46. Perilla JR, Goh BC, Cassidy CK, Liu B, Bernardi RC, Rudack T, Yu H, Wu Z, Schulten K. Curr Opin Struct Biol. 2015; 31:64–74. [PubMed: 25845770]
47. Sanbonmatsu KY. Curr Opin Struct Biol. 2012; 22:168–174. [PubMed: 22336622]
48. Shaytan K, Armeev GA, Goncearenco A, Zhurkin VB, Landsman D, Panchenko AR. J Mol Biol. 2016; 428:221–237. [PubMed: 26699921]
49. Dai L, Mu Y, Nordenskiöld L, van der Maarel JRC. Phys Rev Lett. 2008; 100:118301. [PubMed: 18517834]
50. Andrews CT, Campbell BA, Elcock AH. J Chem Theory Comput. 2017; 13:1794–1811. [PubMed: 28288277]
51. Yoo J, Wilson J, Aksimentiev A. Biopolymers. 2016; 105:752–763. [PubMed: 27144470]
52. Lay WK, Miller MS, Elcock AH. J Chem Theory Comput. 2017; 13:1874–1882. [PubMed: 28437100]
53. Miller M, Lay W, Elcock AH. J Phys Chem B. 2016; 120:8217–8229. [PubMed: 27052117]
54. Miller MS, Lay WK, Li S, Hacker WC, An J, Ren J, Elcock AH. J Chem Theory Comput. 2017; 13:1812–1826. [PubMed: 28296391]
55. Bernèche S, Roux B. Nature. 2001; 414:73–7. [PubMed: 11689945]
56. Luo Y, Roux B. J Phys Chem Lett. 2010; 1:183–9.
57. Venable RM, Luo Y, Gawrisch K, Roux B, Pastor RW. J Phys Chem B. 2013; 117:10183–10192. [PubMed: 23924441]
58. Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, Grubmuller H, MacKerell AD Jr. Nat Methods. 2017; 14:71–73. [PubMed: 27819658]
59. Lay WK, Miller MS, Elcock AH. J Chem Theory Comput. 2016; 12:1401–1407. [PubMed: 26967542]
60. Chapman DE, Steck JK, Nerenberg PS. J Chem Theory Comput. 2014; 10:273–281. [PubMed: 26579910]
61. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. J Chem Phys. 1983; 79:926–935.
62. Izadi S, Anandakrishnan R, Onufriev AV. J Phys Chem Lett. 2014; 5:3863–3871. [PubMed: 25400877]
63. Hess B, Holm C, van der Vegt N. J Chem Phys. 2006; 124:164509. [PubMed: 16674148]
64. Kuhrova P, Best RB, Bottaro S, Bussi G, Sponer J, Otyepka M, Banas P. J Chem Theory Comput. 2016; 12:4534–4548. [PubMed: 27438572]
65. Joung S, Cheatham TE. J Phys Chem B. 2008; 112:9020–9041. [PubMed: 18593145]
66. MacKerell D Jr, Bashford D, Bellott M, Dunbrack RL Jr, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE III, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J,

Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M. J Phys Chem B. 1998; 102:3586–3616. [PubMed: 24889800]

67. Lorentz HA. Annalen der Physik. 1881; 248:127–136.
68. Berthelot D. Comptes rendus hebdomadaires des séances de l'Académie des Sciences. 1898; 126:1703–1706.
69. Robinson, RA., Stokes, RH. Electrolyte Solutions. Butterworths scientific publications; 1959.
70. Murad S, Powles JG. J Chem Phys. 1993; 99:7271–7272.
71. Dill, A., Bromberg, S. Molecular Driving Forces: Statistical Thermodynamics in Biology, Chemistry, Physics, and Nanoscience. Garland Science; New York: 2003.
72. Tsurko E, Neueder R, Kunz W. J Solut Chem. 2007; 36:651–672.
73. Romero CM, GonzÃąlez ME. Fluid Phase Equilib. 2006; 250:99–104.
74. Zafarani-Moattar MT, Sadeghi R. Fluid Phase Equilib. 2002; 203:177–191.
75. Apelblat A, Korin E. J Chem Thermodynamics. 2008; 40:906–908.
76. Held C, Cameretti LF, Sadowski G. Ind Eng Chem Res. 2011; 50:131–141.
77. Fennell CJ, Bizjak A, Vlachy V, Dill KA. J Phys Chem B. 2009; 113:6782–6791. [PubMed: 19206510]
78. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. J Comput Chem. 1992; 13:1011–1021.
79. Gumbart JC, Roux B, Chipot C. J Chem Theory Comput. 2013; 9:794–802. [PubMed: 23794960]
80. Perutz M. Science. 1978; 201:1187–1191. [PubMed: 694508]
81. Masunov A, Lazaridis T. J Am Chem Soc. 2003; 125:1722–30. [PubMed: 12580597]
82. Fulton JL, Heald SM, Badyal YS, Simonson JM. J Phys Chem A. 2003; 107:4688–4696.
83. Badyal YS, Barnes AC, Cuello GJ, Simonson JM. J Phys Chem A. 2004; 108:11819–11827.
84. Jalilehvand F, Spångberg D, Lindqvist-Reis P, Hermansson K, Persson I, Sandström M. J Am Chem Soc. 2001; 123:431–441. [PubMed: 11456545]
85. van der Vegt NFA, Haldrup K, Roke S, Zheng J, Lund M, Bakker HJ. Chem Rev. 2016; 116:7626–41. [PubMed: 27153482]
86. Hess B, van der Vegt N. Proc Natl Acad Sci U S A. 2009; 106:13296–300. [PubMed: 19666545]
87. Ganguly P, Schravendijk P, Hess B, van der Vegt N. J Phys Chem B. 2011; 115:3734–9. [PubMed: 21410261]
88. Jiao D, King C, Grossfield A, Darden TA, Ren P. J Phys Chem B. 2006; 110:18553–9. [PubMed: 16970483]
89. Kumar A. J Phys Chem B. 2003; 107:2808–2814.
90. Chipot C, Jaffe R, Maigret B, Pearlman DA, Kollman PA. J Am Chem Soc. 1996; 118:11217–11224.
91. Sponer J, Leszczynski J, Hobza P. Biopolymers. 2001; 61:3–31. [PubMed: 11891626]
92. Lankas F, Sponer J, Langowski J, Cheatham TE. Biophys J. 2003; 85:2872–2883. [PubMed: 14581192]
93. Mason PE, Neilson GW, Enderby JE, Saboungi ML, Dempsey CE, MacKerell AD, Brady JW. J Am Chem Soc. 2004; 126:11462–11470. [PubMed: 15366892]
94. Gkionis K, Kruse H, Platts JA, Mladek A, Koca J, Sponer J. J Chem Theory Comput. 2014; 10:1326–1340. [PubMed: 26580197]
95. Li H, Ngo V, Da Silva MC, Salahub DR, Callahan K, Roux B, Noskov SY. J Phys Chem B. 2015; 119:9401–9416. [PubMed: 25578354]
96. Hofmeister F. Arch Exp Pathol Pharmacol. 1888; 24:247–260.
97. Lo Nostro P, Ninham BW. Chem Rev. 2012; 112:2286–2322. [PubMed: 22251403]
98. Knobler CM, Gelbart WM. Annu Rev Phys Chem. 2009; 60:367–83. [PubMed: 19046126]
99. Qiu X, Rau DC, Parsegian VA, Fang LT, Knobler CM, Gelbart WM. Phys Rev Lett. 2011; 106:028102. [PubMed: 21405253]
100. Rau DC, Lee B, Parsegian VA. Proc Natl Acad Sci U S A. 1984; 81:2621–2625. [PubMed: 6585818]

101. Todd BA, Parsegian VA, Shirahata A, Thomas TJ, Rau DC. Biophys J. 2008; 94:4775–82. [PubMed: 18326632]
102. Bloomfield VA. Biopolymers. 1997; 44:269–282. [PubMed: 9591479]
103. Smith DE, Tans SJ, Smith SB, Grimes S, Anderson DL, Bustamante C. Nature. 2001; 413:748–752. [PubMed: 11607035]
104. Raspaud E, de la Cruz MO, Sikorav JL, Livolant F. Biophys J. 1998; 74:381–93. [PubMed: 9449338]
105. Kornyshev A, Lee D, Leikin S, Wynveen A. Rev Mod Phys. 2007; 79:943–996.
106. Bai Y, Greenfeld M, Travers KJ, Chu VB, Lipfert J, Doniach S, Herschlag D. J Am Chem Soc. 2007; 129:14981–14988. [PubMed: 17990882]
107. Yoo J, Aksimentiev A. J Phys Chem B. 2012; 116:12946–12954. [PubMed: 23016894]
108. Qiu X, Giannini J, Howell SC, Xia Q, Ke F, Andresen K. Biophys J. 2013; 105:984–992. [PubMed: 23972850]
109. Luan B, Aksimentiev A. J Am Chem Soc. 2008; 130:15754–15755. [PubMed: 18975864]
110. Kowalczyk SW, Wells DB, Aksimentiev A, Dekker C. Nano Lett. 2012; 12:1038–1044. [PubMed: 22229707]
111. Yoo J, Kim H, Aksimentiev A, Ha T. Nat Commun. 2016; 7:11045. [PubMed: 27001929]
112. Rothemund PWK. Nature. 2006; 440:297–302. [PubMed: 16541064]
113. Yoo J, Aksimentiev A. Proc Natl Acad Sci U S A. 2013; 110:20099–20104. [PubMed: 24277840]
114. Slone S, Li CY, Yoo J, Aksimentiev A. New J Phys. 2016; 18:055012.
115. Li CY, Hemmig EA, Kong J, Yoo J, Hernández-Ainsa S, Keyser UF, Aksimentiev A. ACS Nano. 2015; 9:1420–1433. [PubMed: 25623807]
116. Yoo J, Aksimentiev A. J Phys Chem Lett. 2015; 6:4680–4687. [PubMed: 26551518]
117. Maingi V, Lelimousin M, Howorka S, Sansom MSP. ACS Nano. 2015; 9:11209–11217. [PubMed: 26506011]
118. Göpfrich K, Li CY, Mames I, Bhamidimarri SP, Ricci M, Yoo J, Mames A, Ohmann A, Winterhalter M, Stulz E, Aksimentiev A, Keyser UF. Nano Lett. 2016; 16:4665–4669. [PubMed: 27324157]
119. Göpfrich K, Li CY, Ricci M, Bhamidimarri SP, Yoo J, Gyenes B, Ohmann A, Winterhalter M, Aksimentiev A, Keyser UF. ACS Nano. 2016; 10:8207–8214. [PubMed: 27504755]
120. Biertümpfel C, Yang W, Suck D. Nature. 2007; 449:616–20. [PubMed: 17873859]
121. Lilley DM, Clegg RM. Annu Rev Biophys Biomol Struct. 1993; 22:299–328. [PubMed: 8347993]
122. Joo C, McKinney SA, Lilley DMJ, Ha T. J Mol Biol. 2004; 341:739–51. [PubMed: 15288783]
123. Hohng S, Zhou R, Nahas MK, Yu J, Schulten K, Lilley DMJ, Ha T. Science. 2007; 318:279–83. [PubMed: 17932299]
124. Duckett DR, Murchie AI, Lilley DM. EMBO J. 1990; 9:583–590. [PubMed: 2303044]
125. Clegg RM, Murchie AI, Lilley DM. Biophys J. 1994; 66:99–109. [PubMed: 8130350]
126. Murchie I, Clegg RM, von Kitzing E, Duckett DR, Diekmann S, Lilley DM. Nature. 1989; 341:763–6. [PubMed: 2797209]
127. Maffeo C, Luan B, Aksimentiev A. Nucleic Acids Res. 2012; 40:3812–3821. [PubMed: 22241779]
128. Brown RF, Andrews CT, Elcock AH. J Chem Theory Comput. 2015; 11:2315–2328. [PubMed: 26574427]
129. Hase F, Zacharias M. Nucleic Acids Res. 2016; 44:7100. [PubMed: 27407106]
130. Ivani I, Dans PD, Noy A, Pérez A, Faustino I, Hospital A, Walther J, Andrio P, Goñi R, Balaceanu A, Portella G, Battistini F, Gelpí JL, González C, Vendruscolo M, Laughton CA, Harris SA, Case DA, Orozco M. Nat Methods. 2015; 13:55. EP –. [PubMed: 26569599]
131. Piana S, Sarkar K, Lindorff-Larsen K, Guo M, Gruebele M, Shaw DE. J Mol Biol. 2011; 405:43–8. [PubMed: 20974152]
132. Israelachvili, JN. Intermolecular and Surface Forces: Revised Third Edition. Elsevier Science; 2011.

133. Sheinerman FB, Norel R, Honig B. Curr Opin Struct Biol. 2000; 10:153–159. [PubMed: 10753808]
134. Nicholls A, Sharp KA, Honig B. Proteins: Struct, Funct, Bioinf. 1991; 11:281–296.
135. Dill KA, Bromberg S, Yue K, Chan HS, Ftebig KM, Yee DP, Thomas PD. Protein Sci. 1995; 4:561–602. [PubMed: 7613459]
136. Kubelka J, Chiu TK, Davies DR, Eaton WA, Hofrichter J. J Mol Biol. 2006; 359:546–53. [PubMed: 16643946]
137. Jäger M, Zhang Y, Bieschke J, Nguyen H, Dendle M, Bowman ME, Noel JP, Gruebele M, Kelly JW. Proc Natl Acad Sci U S A. 2006; 103:10648–10653. [PubMed: 16807295]
138. Vijay-Kumar S, Bugg CE, Cook WJ. J Mol Biol. 1987; 194:531–544. [PubMed: 3041007]
139. Kohn JE, Millett IS, Jacob J, Zagrovic B, Dillon TM, Cingel N, Dothager RS, Seifert S, Thiyagarajan P, Sosnick TR, Hasan MZ, Pande VS, Ruczinski I, Doniach S, Plaxco KW. Proc Natl Acad Sci U S A. 2004; 101:12491–12496. [PubMed: 15314214]
140. Piana S, Lindorff-Larsen K, Shaw DE. Proc Natl Acad Sci U S A. 2012; 109:17845–17850. [PubMed: 22822217]
141. Berger O, Edholm O, Jähnig F. Biophys J. 1997; 72:2002–13. [PubMed: 9129804]
142. Wang KH, Penmatsa A, Gouaux E. Nature. 2015; 521:322–327. [PubMed: 25970245]
143. Dumas JJ, Merithew E, Sudharshan E, Rajamani D, Hayes S, Lawe D, Corvera S, Lambright DG. Mol Cell. 2001; 8:947–958. [PubMed: 11741531]
144. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE. Acc Chem Res. 2000; 33:889–897. [PubMed: 11123888]
145. Petkova T, Ishii Y, Balbach JJ, Antzutkin ON, Leapman RD, Delaglio F, Tycko R. Proc Natl Acad Sci U S A. 2002; 99:16742–16747. [PubMed: 12481027]
146. Ji TH, Grossmann M, Ji I. J Biol Chem. 1998; 273:17299–17302. [PubMed: 9651309]
147. Bokoch MP, Zou Y, Rasmussen SGF, Liu CW, Nygaard R, Rosenbaum DM, Fung JJ, Choi HJJ, Thian FS, Kobilka TS, Puglisi JD, Weis WI, Pardo L, Prosser RS, Mueller L, Kobilka BK. Nature. 2010; 463:108–12. [PubMed: 20054398]
148. Betz H. Neuron. 1990; 5:383–392. [PubMed: 1698394]
149. Armstrong N, Sun Y, Chen GQ, Gouaux E. Nature. 1998; 395:913–917. [PubMed: 9804426]
150. Pabo CO, Sauer RT. Annu Rev Biochem. 1984; 53:293–321. [PubMed: 6236744]
151. Pabo CO, Sauer RT. Annu Rev Biochem. 1992; 61:1053–1095. [PubMed: 1497306]
152. Thomsen ND, Berger JM. Cell. 2009; 139:523–34. [PubMed: 19879839]
153. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. Nature. 1997; 389:251–260. [PubMed: 9305837]
154. Bintu L, Ishibashi T, Dangkulwanich M, Wu YYY, Lubkowska L, Kashlev M, Bustamante C. Cell. 2012; 151:738–49. [PubMed: 23141536]
155. DeRouchey J, Hoover B, Rau DC. Biochemistry. 2013; 52:3000–3009. [PubMed: 23540557]
156. Kyte J, Doolittle RF. J Mol Biol. 1982; 157:105–132. [PubMed: 7108955]
157. Hessa T, Kim H, Bihlmaier K, Lundin C, Boekel J, Andersson H, Nilsson I, White SH, von Heijne G. Nature. 2005; 433:377–381. [PubMed: 15674282]
158. von Heijne G. Nat Rev Mol Cell Biol. 2006; 7:909–18. [PubMed: 17139331]
159. Schmidt D, Jiang QX, MacKinnon R. Nature. 2006; 444:775–779. [PubMed: 17136096]
160. McLaughlin S, Murray D. Nature. 2005; 438:605–11. [PubMed: 16319880]
161. Yoo J, Cui Q. Biophys J. 2010; 99:1529–38. [PubMed: 20816065]
162. McLaughlin S, Wang J, Gambhir A, Murray D. Annu Rev Biophys Biomol Struct. 2002; 31:151–75. [PubMed: 11988466]
163. Suh BCC, Hille B. Annu Rev Biophys. 2008; 37:175–95. [PubMed: 18573078]
164. Lemmon MA. Nat Rev Mol Cell Biol. 2008; 9:99–111. [PubMed: 18216767]
165. van den Bogaart G, Meyenberg K, Risselada HJ, Amin H, Willig KI, Hubrich BE, Dier M, Hell SW, Grubmüller H, Diederichsen U, Jahn R. Nature. 2011; 479:552–5. [PubMed: 22020284]

166. Slochower DR, Huwe PJ, Radhakrishnan R, Janmey PA. J Phys Chem B. 2013; 117:8322–8329. [PubMed: 23786273]

167. Wu EL, Qi Y, Song KC, Klauda JB, Im W. J Phys Chem B. 2014; 118:4315–4325. [PubMed: 24689790]

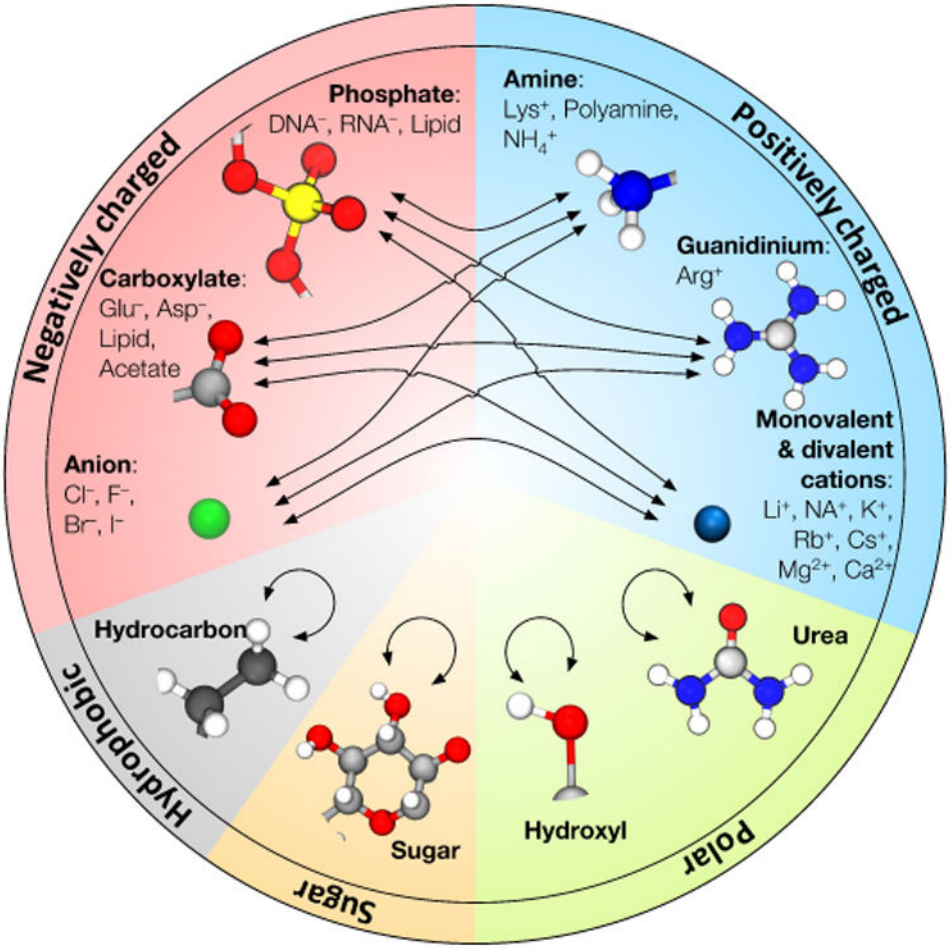

Fig. 1.

Visual summary of NBFIX corrections to the AMBER99 and CHARMM force fields. Tables 1 and 2 summarize efforts of various groups in this area. Tables 3 and 4 provide specific values of the Champaign-Urbana NBFIX (CUFIX) corrections to the AMBER99 and CHARMM force fields.

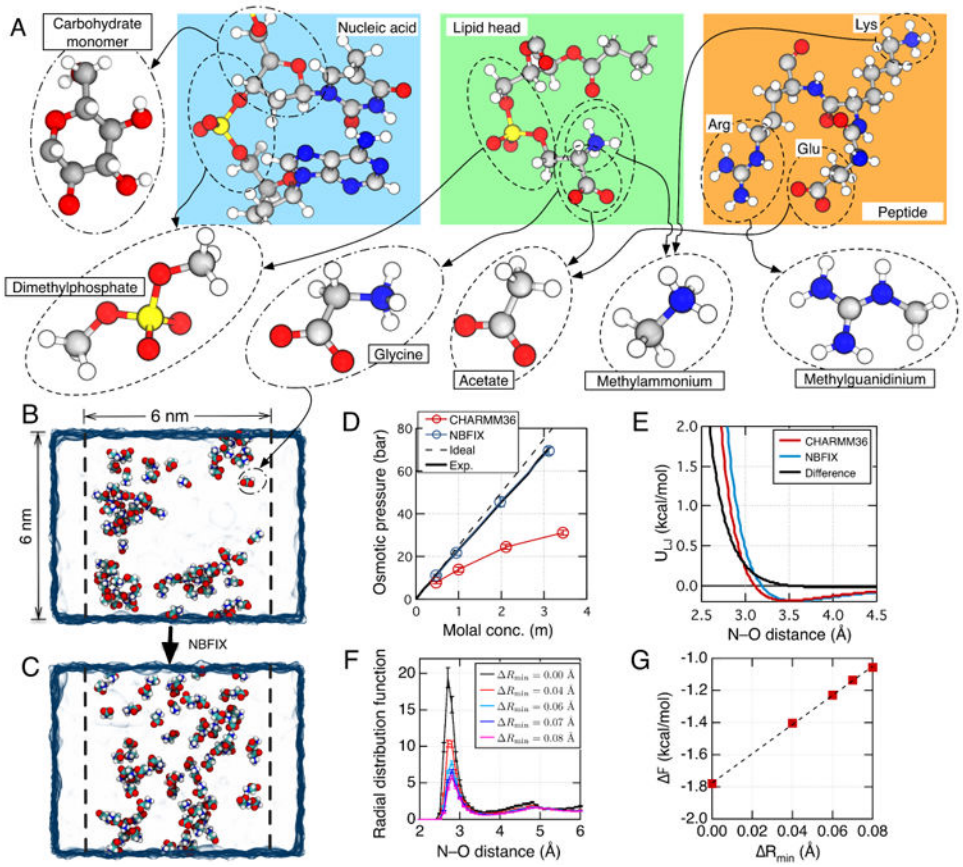

Fig. 2.

NBFIX calibration using model compounds. (A) To perform NBFIX calibration, complex biological molecules are represented by simpler model compounds. The compounds are chosen to isolate a specific type of inter-solute interactions, which can be modulated by the NBFIX correction. The figure shows a possible decomposition of three representative biomolecules—a nucleic acid fragment, a lipid molecule and a peptide. Molecular structures are shown using the ball-and-stick representation where each atom is colored according to its type: hydrogen, white; carbon, gray; nitrogen, blue; oxygen, red; phosphorus, yellow. (B,C) Calibration of amine–carboxylate interactions using the osmotic pressure of a glycine solution. The simulated system includes a volume of water (blue semi-transparent surface) divided into two compartments by two planar half-harmonic potentials (depicted by dashed lines). Glycine monomers (shown as spheres colored by atom: carbon, cyan; hydrogen, white; oxygen, red; nitrogen, blue) are confined to remain within one compartment by the half-harmonic potentials while water can exchange between the compartments freely. The difference in the glycine concentration between the two compartments generates osmotic pressure. The equilibrium osmotic pressure value is determined from the average force exerted by the confining potentials on the glycine molecules and the cross-section area of the system. Panels B and C illustrate instantaneous configurations of a 3 m glycine solution observed at the end of two 25 ns simulations performed using the standard CHARMM36 force field without (B) and with (C) the NBFIX correction. (D) Osmotic pressure of a glycine solution as a function of its concentration obtained from MD simulations performed

using the CHARMM force field with (blue) and without (red) the NBFIX corrections. Experimentally determined[72] and ideal solution (osmotic coefficient = 1) dependences are shown as solid and dashed black lines, respectively. (E) The effect of an NBFIX correction (ΔR_{min} = 0.08 Å) on the LJ interaction potential between amine nitrogen and carboxylate oxygen. The LJ potentials are plotted using the standard CHARMM force field with (blue) and without (red) the NBFIX correction. The difference between the NBFIX and standard potentials is shown as a black line. (F) Inter-molecular radial distribution function (RDF) of glycine nitrogen with respect to glycine oxygen in MD simulations of a ~3 m glycine solution at several values of ΔR_{min}. Error bars indicate standard error. The value of ΔR_{min} affects the height of the first peak, g_1. (G) Correlation between ΔR_{min} and $\Delta F = -k_B T \log(g_1)$. Dashed line indicates a linear fit. Figures in panels B–D were adapted with permission from Ref. 41.

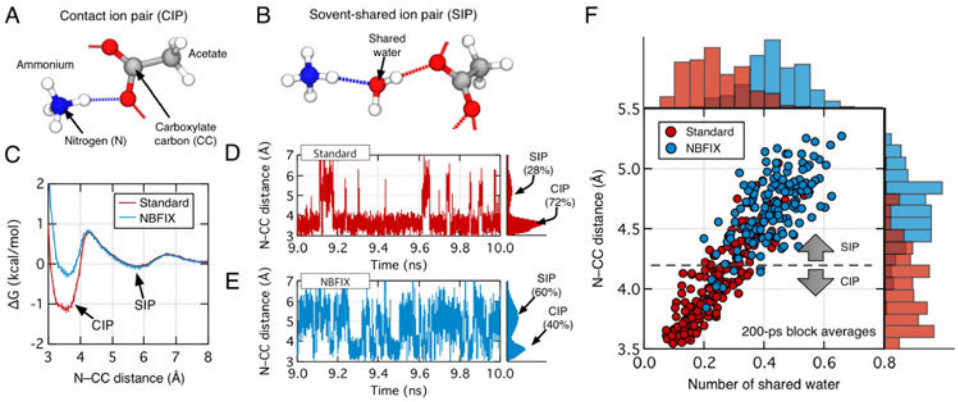

Fig. 3.

The effects of NBFIX on ion pair formation. (A, B) Ammonium and acetate ions forming a contact ion pair (CIP, panel A) and a solvent-shared ion pair (SIP, panel B)[77]. Nitrogen, oxygen, carbon, and hydrogen atoms are shown in blue, red, gray, and white colors. Hydrogen bonds are depicted as dotted lines. (C) Potential of mean force (PMF) between ammonium and acetate in pure water computed using the standard (red) and NBFIX-optimized (blue) force field models (data taken from Ref. 41). The PMF was computed using the umbrella sampling method; the force constants of the umbrella restraints were 10 kcal/mol·Å^2. The reaction coordinate was the distance between the ammonium nitrogen (N) and carboxylate carbon (CC) atoms. (D,E) Variations of the N–CC distance in the simulations performed using the standard (D) and NBFIX-optimized (E) force field models. Each plot illustrates the last 1 ns segment of a 30 ns simulation performed having the N–CC distance harmonically constrained to 4.2 Å using a weak harmonic potential (0.5 kcal/mol·Å^2 force constant). The instantaneous N–CC distance was recorded every 20 fs. The unnormalized probability density functions computed from the entire 30 ns trajectories are shown at the right vertical axis of each panel. (F) Correlations between the N–CC distance and the number of shared water molecules. Each data point represents a 200 ps average. Here, a water molecule is considered to be shared if it simultaneously forms hydrogen bonds with both ammonium and acetate (see panel B). The occurrence of a hydrogen bond was determined using the g_hbond program of the Gromacs package with the distance and angle cutoffs of 3.5 Å and 30°, respectively. The bar graphs at the right and the top axes show the unnormalized probability distributions of the N–CC distance and the number of shared water molecules. All data reported in this figure were obtained using the AMBER ff99 force field.

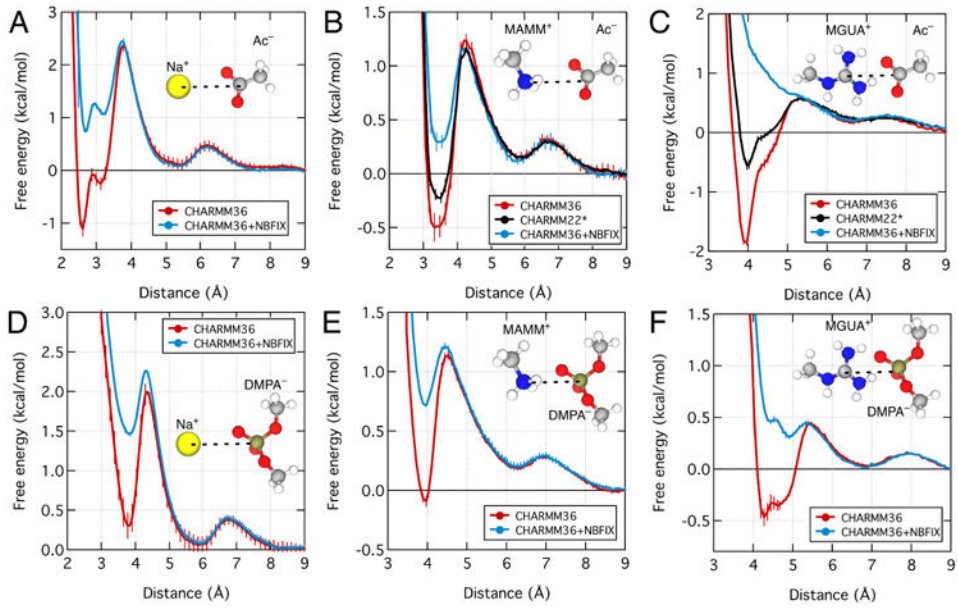

Fig. 4.

Effect of NBFIX corrections on the PMF between sodium and acetate (A), methylammonium and acetate (B), methylguanidinium and acetate (C), sodium and dimethylphosphate (D), methylammonium and dimethylphosphate (E), and methylguanidinium and dimethylphosphate (F). The reaction coordinate was defined as the distance between the following atoms of the ion pairs: sodium, nitrogen of methylammonium, nitrogen of methylguanidinium, carboxylate carbon of acetate, and phosphorus of dimethylphosphate. Each panel shows the PMFs computed using the standard CHARMM36[66] (red) and the CHARMM36 force field with the NBFIX correction[41] (blue). In panels B and C, the black lines show the PMFs computed using CHARMM22*[30]. The inset of each plot shows the chemical structure of the ion pair. Except for panel C, all NBFIX-corrected data were obtained using parameters developed in Ref. 41. Fig. 5 and Section 3.1.2 describe the development of the NBFIX correction for the guanidinium–carboxylate pair.

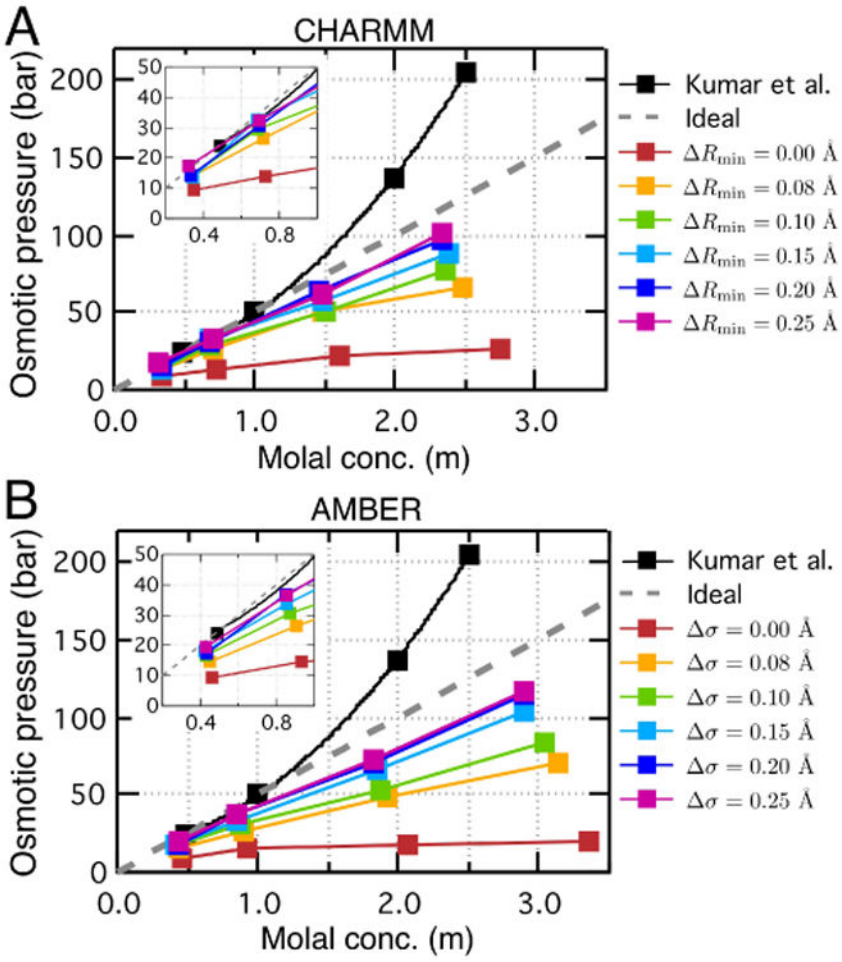

Fig. 5.

Calibration of the guanidinium–acetate interaction using the osmotic pressure method. (A) Osmotic pressure of the guanidinium–acetate solution as a function of the solution molality obtained using the CHARMM36 force field with the specified value of the NBFIX correction (colored symbols) and experiment[89] (black symbols). Gray dashed line shows the dependence of an ideal 1:1 solution. For the purpose of calibration, the LJ R_{min} parameter describing the interactions between the guanidinium nitrogen and acetate oxygen was increased by ΔR_{min}. The inset shows a zoomed-in view of the dependences at concentrations lower than 1 m. (B) Same as in panel A but for the AMBER99 force field. To be consistent with the AMBER parameterization convention, the calibration of the LJ interaction is done by adjusting the σ parameter ($R_{min} = 2^{(1/6)}\sigma$).

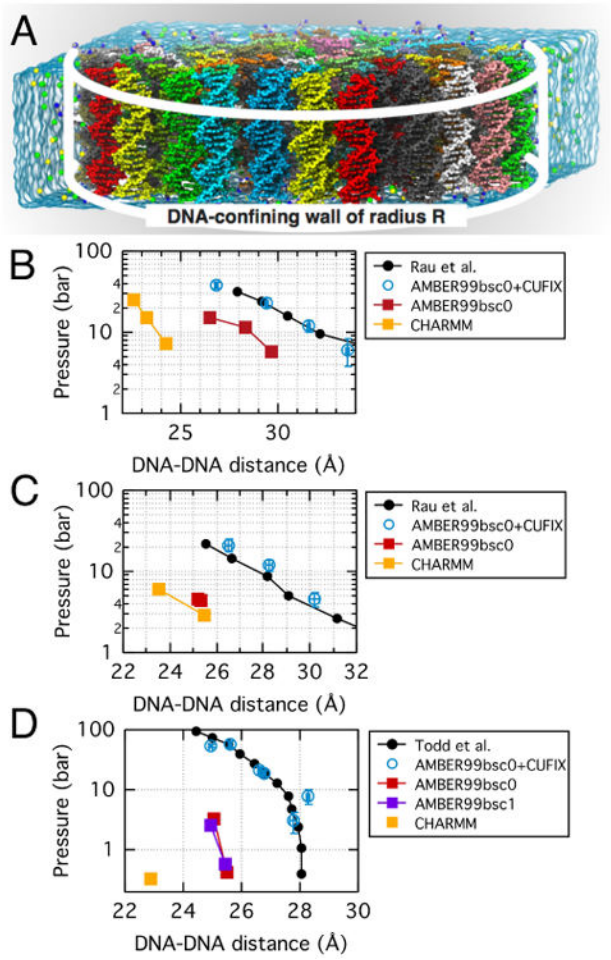

Fig. 6.

Effect of NBFIX corrections on MD simulations of DNA array systems [41,43]. (A) DNA array system containing 64 20-bp double-stranded DNA molecules (each shown in distinct color) confined to a cylinder of radius R (white lines) by a half-harmonic potential[31,43]. Under periodic boundary conditions, each DNA molecule is effectively infinite. The blue semi-transparent surface indicates the volume occupied by electrolyte in a unit simulation cell; water and ions are not subject to external potential and can exchange between buffer and DNA array volumes. The internal pressure of a DNA array is measured from the total force applied to DNA by the confining potential divided by the array's area. A set of simulations carried out at different values of R yields a dependence of the internal pressure on the inter-DNA distance. (B) Experimentally determined (black) and simulated (colors) pressure in a DNA array as a function of the mean distance between the nearest DNA helices. Both experimental and simulated dependences were obtained at 250 mM Na^+ [43,100]. (C) Same as in panel B but for $MgCl_2$ solution. The experimental dependence was measured at 20 mM concentration of Mg^{2+} ions[100]. The simulated pressure was obtained at 20 mM Mg^{2+} and 200 mM Na^+ [43]. (D) Same as in panel B but in the presence of spermine. The experimental dependence was obtained at 2 mM spermine concentration[101]. The charge of the spermine molecules in the MD systems was equal by magnitude to the charge of the DNA array and no additional ions were introduced to the system[43].

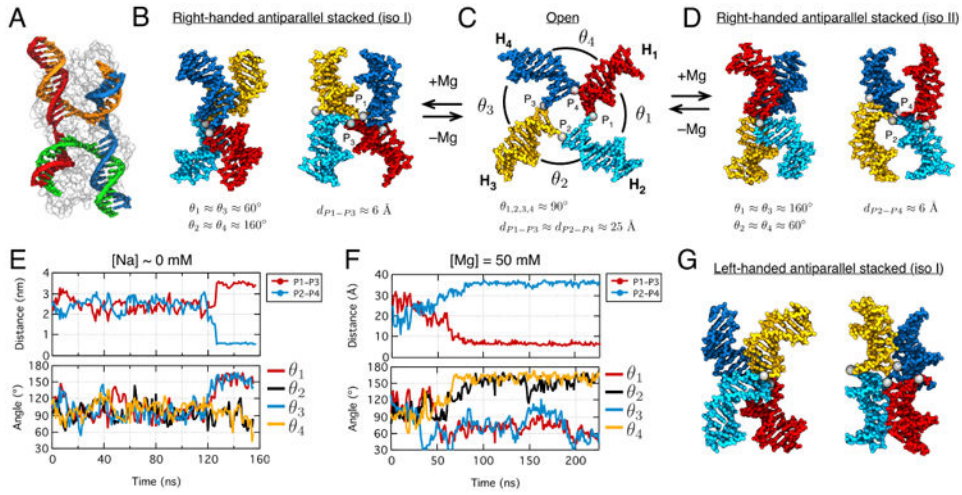

Fig. 7.

MD simulation of conformational transitions in a Holliday junction (HJ). (A) Crystal structure (PDB: 2QNC) of an open HJ conformation (each DNA strand uniquely colored) co-crystallized with a T4 endonuclease (semitransparent surface)[120]. (B–D) Cation-dependent conformational transitions in HJ. At low ionic strength, HJ forms an open conformation (C). At high ionic strength, HJ forms one of the two, panels B and D, right-handed antiparallel stacked-X conformations. The HJ conformations can be described using the distances d_{P1-P3} and d_{P2-P4} between the phosphate groups P_1, …,P_4 (gray spheres) at the junction, and angles θ_1, …,θ_4 between the neighboring arms. Typical values of d_{P1-P3}, d_{P2-P4} and θ_1, …,θ_4 are shown at the bottom of each panel. (E) The inter-phosphate distances d_{P1-P3} and d_{P2-P4} (top) and the θ_1, …,θ_4 angles (bottom) in an MD simulation initiated from an open HJ conformation and carried out at 0 mM Na^+. (F) Same as in panel E but for a simulation that began from an open HJ conformation but was carried out at 50 mM Mg^{2+}. (G) An example of a left-hand antiparallel stacked-X conformation found in an MD simulation.

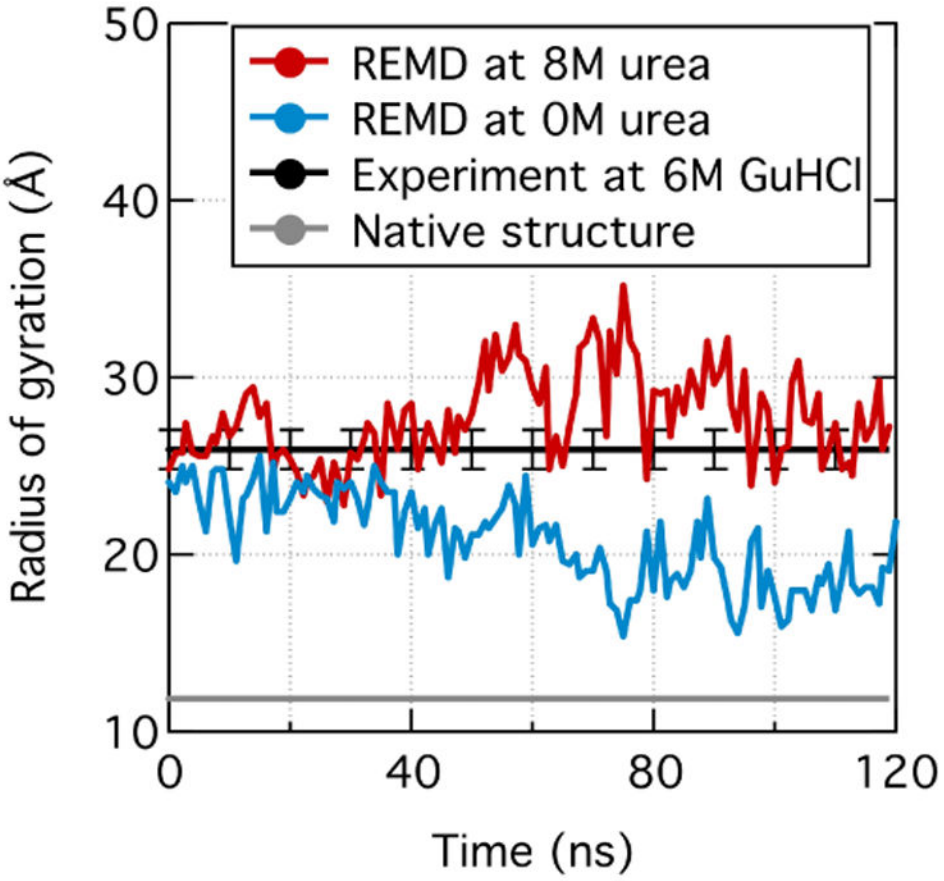

Fig. 8.

Experimental and simulated radius of gyration, R_g, of an unfolded ubiquitin. R_g values observed in the simulations carried out at 0 and 8 M urea concentrations are shown in blue and red, respectively. The R_g values shown were measured using the 300 K replica of a 53-temperature replica-exchange MD (REMD) simulation. Experimental R_g values for the folded (gray) and unfolded (black, 6 M GuHCl) ubiquitin were taken from PDB 1UBQ[138] and Ref. 139, respectively.

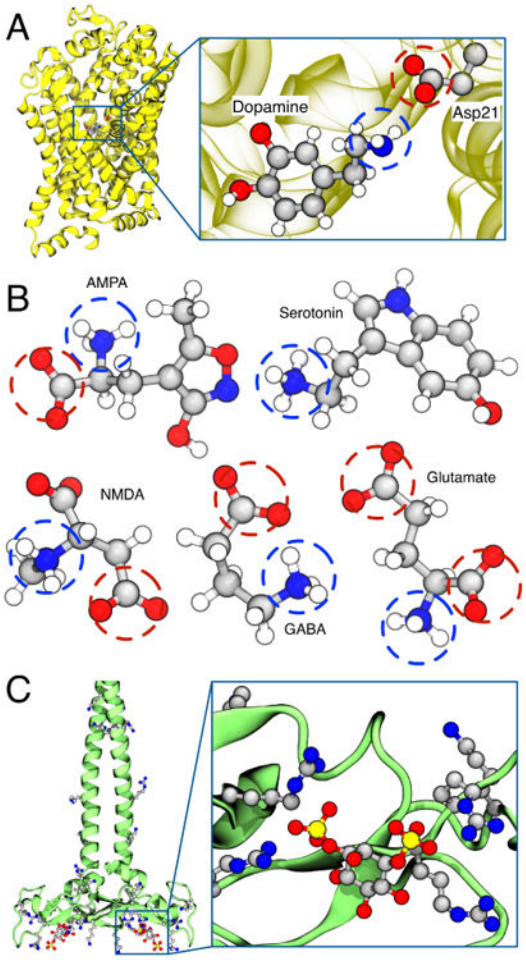

Fig. 9.

Interactions between charged groups in cell signaling. (A) Crystal structure of a dopamine transporter having a dopamine molecule bound to the binding site (PDB: 4XP9)[142]. The inset shows a close-up view of the dopamine molecule and the interacting aspartate side chain. (B) Structures of five representative neurotransmitters: AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid), NMDA (N-methyl-D-aspartate), GABA (γ-aminobutyric acid), serotonin, and glutamate. The chemical structures are shown using the ball-and-stick representation colored by the atom type: carbon, gray; oxygen, red; nitrogen, blue; hydrogen, white. Charged amine and carboxylate groups are indicated by blue and red circles, respectively. (C) Crystal structure of EEA1 homodimer of C-terminal FYVE domain bound to inositol 1,3-diphosphate (PDB: 1JOC)[143]. The inset shows the structure of inositol 1,3-diphosphate coordinated by four neighboring arginine side chains.

Table 1

Summary of NBFIX corrections for the AMBER ff99-derivative force fields. COO, PO_4, and SO_4 indicate carboxylate, phosphate, and sulfate groups, respectively.

Reference	Type of NBFIX corrections
Yoo & Aksimentiev, 2012[31]	Li/Na/K/Mg–Cl/COO/PO_4
Yoo & Aksimentiev, 2016[41]	Amine–COO/PO_4/SO_4
Yoo & Aksimentiev, 2016[41]	Urea–urea
Yoo & Aksimentiev, 2016[42]	Hydrocarbon–hydrocarbon
Yoo, Wilson, Aksimentiev, 2016[51]	Ca–Cl/COO/PO_4
Lay, Miller, Elcock, 2017[52]	Carbohydrate–carbohydrate
Miller, …, Elcock, 2017[53,54]	Hydroxyl–hydroxyl
This article (Fig. 5A)	Guanidinium–COO
This article (Figs. S1 to S5)	Ion pairs including Cs/Rb/Br/I

Table 2

Summary of NBFIX corrections for the CHARMM force fields. COO, PO_4, and SO_4 indicate carboxylate, phosphate, and sulfate groups, respectively.

Reference	Type of NBFIX corrections
Bernèche & Roux, 2001[55]	Na/K–carbonyl
Luo & Roux, 2010[56]	Na/K–Cl
Yoo & Aksimentiev, 2012[31]	Li/Na/K/Mg–Cl/COO/PO_4
Venable, …, Pastor, 2013[57]	Na–COO
Yoo & Aksimentiev, 2016[41]	Amine–COO/PO_4/SO_4
Yoo & Aksimentiev, 2016[41]	Urea–urea
Yoo, Wilson, Aksimentiev, 2016[51]	Ca–Cl/COO/PO_4
Huang, ï, MacKerell, 2017[58]	Guanidinium–COO
Lay, Miller, Elcock, 2016[59]	Carbohydrate–carbohydrate
This article (Fig. 5B)	Guanidinium–COO
This article (Figs. S6, S7)	Ion pairs including Rb/Cs

Table 3

Pair-specific NBFIX corrections to AMBER ff99-derivative force fields. The table lists corrections to the LJ *a* parameters of the specified atom pairs. The corrections are given relative to the standard force field values and have units of Angstroms. The standard σ values for monovalent ions are those of Joung and Cheatham[65]. The standard σ values for the divalent cations are from the CHARMM36 force field [14]. Mg^{2+} and Ca^{2+} ions are described as permanent hexahydrate and heptahydrate complexes, respectively. The dipole moment of each water molecule within each Mg^{2+} hexahydrate or Ca^{2+} heptahydrate complex is set to be 1 or 0.5 Debye, respectively larger than the standard value of TIP3P water[31,51]. The NBFIX corrections are for the atom pairs formed by monovalent ions, water oxygens of Mg^{2+} hexahydrates and Ca^{2+} heptahydrates, nitrogen atom of a positively charged amine group (N–H), nitrogen atom of a guanidinium group (N=C), acetate oxygen (O=C), phosphate oxygen (O=P), sulfate oxygen (O=S), and methyl carbon (C–H). Superscripts §, ¶, *, and ‡ correspond to Refs. 31, 41, 51, and 42, respectively, which detail validation of the specific corrections. The value of 0.00 indicates that no NBFIX correction is needed to reproduce the osmotic pressure experiment using the standard parameter set. The CUFIX-corrected parameter sets in the Gromacs, AMBER, and Anton formats are available at http://bionano.physics.illinois.edu/CUFIX.

	Li^+	**Na^+**	**K^+**	**Rb^+**	**Cs^+**	**Mg^{2+}**	**Ca^{2+}**	**N–H**	**N=C**	**C–H**
Cl^-	0.12§	0.015§	0.02§	0.02 Fig. S1A	0.00 Fig. S1A	0.02§	0.01*	0.00¶	0.00¶	N/A
Br^-	0.15 Fig. S2A	0.00 Fig. S2B	0.02 Fig. S2C	0.00 Fig. S2D	0.00 Fig. S2E	0.00 Fig. S5A	0.00 Fig. S5C	N/A	N/A	N/A
I^-	0.00 Fig. S3A	0.00 Fig. S3B	0.00 Fig. S3C	0.00 Fig. S3D	−0.04 Fig. S3E	−0.08 Fig. S5B	0.00 Fig. S5D	N/A	N/A	N/A
O=C	0.17§	0.10§	0.30§	0.15Fig. S4A	0 10Fig. S4B	0.06§	0.07*	0.08¶	0.20Fig. 5A	N/A
O=P	0.17§	0.10§	0.30§	0.15 Fig. S4A	0.10 Fig. S4B	0.06§	0.07*	0.14¶	0.14¶	N/A
O=S	0.17§	0.10§	0.30§	0.15 Fig. S4A	0.10 Fig. S4B	0.06§	0.07*	0.14¶	0.14¶	N/A
C–H	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	−0.40‡

Table 4

Pair-specific NBFIX corrections to the CHARMM36 force field[14]. The table lists corrections to the LJ R_{min} parameters of the specified atom pairs. The corrections are given relative to the standard force field values and have units of Angstroms. Mg^{2+} and Ca^{2+} are modeled as permanent hexahydrate and heptahydrate complexes, respectively, with the water dipole moments increased by 1 or 0.5 Debye, respectively, above the standard TIP3P value[31,51]. The NBFIX corrections are for the atom pairs formed by monovalent ions, water oxygens of Mg^{2+} hexahydrates and Ca^{2+} heptahydrates, nitrogen atom of a positively charged amine group (N–H), nitrogen atom of a guanidinium group (N=C), acetate oxygen (O=C), phosphate oxygen (O=P), and sulfate oxygen (O=S). Superscripts §, ¶ and * correspond to Refs. 31, 41, and 51, respectively, which detail validation of the specific corrections. The value of 0.00 indicates that no NBFIX correction is needed to reproduce the osmotic pressure experiment using the standard parameter set. Br^- and I^- anions are not listed because they are not included in the standard CHARMM36 force field. The CUFIX-corrected parameter set in the CHARMM format is available at http://bionano.physics.illinois.edu/CUFIX.

	Li^+	**Na^+**	**K^+**	**Rb^+**	**Cs^+**	**Mg^{2+}**	**Ca^{2+}**	**N–H**	**N=C**
Cl^-	0.23§	0.06§	0.05§	0.022 Fig. S6A	0.00 Fig. S6B	0.015§	0.011*	0.03¶	0.03¶
O=C	0.19§	0.09§	0.08§	0.056 Fig. S7A	0.00 Fig. S7B	0.055§	0.079*	0.08¶	0.20 Fig. 5B
O=P	0.19§	0.09§	0.08§	0.056 Fig. S7A	0.00 Fig. S7B	0.055§	0.079*	0.16¶	0.16¶
O=S	0.19§	0.09§	0.08§	0.056 Fig. S7A	0.00 Fig. S7B	0.055§	0.079*	0.16¶	0.16¶