How well do force fields capture the strength of salt bridges in proteins?

Mustapha Carab Ahmed[1], Elena Papaleo[1,2] and Kresten Lindorff-Larsen[1]

[1] Structural Biology and NMR Laboratory, Linderstrøm-Lang Centre for Protein Science, Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, DK-2200 Copenhagen, Denmark

Corresponding Author

Kresten Lindorff-Larsen
Email address: lindorff@bio.ku.dk

Present Address

[2] Computational Biology Laboratory, Danish Cancer Society Society Research Center, Copenhagen, Denmark

Abstract

Salt bridges form between pairs of ionisable residues in close proximity and are important interactions in proteins. While salt bridges are known to be important both for protein stability, recognition and regulation, we still do not have fully accurate predictive models to assess the energetic contributions of salt bridges. Molecular dynamics simulations is one technique that may be used study the complex relationship between structure, solvation and energetics of salt bridges, but the accuracy of such simulations depend on the force field used. We have used NMR data on the B1 domain of protein G (GB1) to benchmark molecular dynamics simulations. Using enhanced sampling simulations, we calculated the free energy of forming a salt bridge for three possible ionic interactions in GB1. The NMR experiments showed that these interactions are either not formed, or only very weakly formed, in solution. In contrast, we show that the stability of the salt bridges is slightly overestimated in simulations of GB1 using six commonly used combinations of force fields and water models. We therefore conclude that further work is needed to refine our ability to model quantitatively the stability of salt bridges through simulations, and that comparisons between experiments and simulations will play a crucial role in furthering our understanding of this important interaction.

Introduction

Proteins are stabilized via the concerted action of numerous weak forces including those that arise from hydrogen bonds, the hydrophobic effect and salt bridges (Dill, 1990; Zhou & Pang, 2017). While we now know much about the relative contributions and physical origins of these effects, we still do not have quantitative models that allow us, for example, to predict accurately the overall stability of a protein given its three-dimensional structure. A quantitative understanding of protein stability would aid both to design proteins with improved stabilities (Foit et al., 2009), as well as understand how loss of protein stability may give rise to disease (Casadio et al., 2011; Nielsen et al., 2017).

We here focus our attention on salt bridges in proteins, where they occur between oppositely charged ionisable groups, most commonly between the negatively charged side chains of aspartate and glutamate or the carboxy-terminus, and positive charges in arginine, lysine or histidine side chains, or with the N-terminal ammonium group. A salt bridge is formed when the two residues involved are spatially close enough to form energetically favourable electrostatic interaction between the (partially) charged atoms (Barlow & Thornton, 1983). Although the first quantitative models of protein electrostatics are almost one hundred years old (Linderstrøm-Lang, 1924), our ability to predict electrostatic properties remain incomplete. Indeed, while ionic interactions were also early suggested to contribute substantially to the stability of certain proteins (Speakman, J. B., & Hirst, 1931), their importance and energetic contribution was controversial already before they had been observed in experimentally-derived protein structures (Jacobsen & Linderstrøm-Lang, 1949). The stability of a salt bridge depends on the environment around the residues, the pH, and the distance and geometric orientation between the involved residues. Salt bridges may be located on the protein surface, where they are exposed to the solvent, or they can be found buried within the hydrophobic interior of the folded protein. Studies on different protein families have found that buried salt-bridges are more likely to be conserved and functionally relevant than surface exposed ones (Schueler & Margalit, 1995; Takano et al., 2000). Surface exposed salt bridges are generally weaker, more variable in their contribution to stability, and more difficult to predict (Sarakatsannis & Duan, 2005). Double mutant cycles in barnase illustrate the distinction between forming a salt bridge in the low dielectric environment of

the protein core versus the higher dielectric on the protein surface. These experiments show that an internal salt bridge stabilizes the protein by more than 3 kcal mol^{-1} (Vaughan et al., 2002), whereas a surface exposed salt bridge is about 10 times weaker, ~ 0.3 kcal mol^{-1} at low ionic strength, a value that decreases to zero at higher ionic strength (Serrano et al., 1990).

Although solvent-exposed salt bridges are often observed in crystal structures of proteins, it is thus not always clear whether these interactions are stable in solution, in particular because their dynamic and transient nature make them difficult to study. With its site-specific resolution and ability to detect even transient interactions, NMR spectroscopy can, however, be used to study salt bridges in solution, and to determine the extent to which different residues interact. In one intriguing study, a range of different NMR experiments were used to examine three different potential salt bridges in the B1 domain of Protein G (hereafter called GB1) (Tomlinson et al., 2009). Six different crystal structures of GB1 show that three of six solvent-exposed lysines, K12, K39 and K58, form salt-bridges to the nearby acidic residues E12, E23 and D55 (Tomlinson et al., 2009) (Fig. 1A). To examine whether these ionic interactions are present in solution, the authors performed a series of experiments that include monitoring both the lysine nitrogen and proton chemical shifts, and the hydrogen-deuterium isotope effects on the ammonium group, while titrating the carboxylates to protonate them. Surprisingly, for two of the putative ion pairs (K12–E23 and K58–D55) the results showed essentially no observable changes at the lysines as the carboxylates changed protonation state, suggesting no substantial formation of a salt bridge. For one pair, the intra-helical K39–E35, there was a small change in the ^{15}N chemical shift of the ammonium group as pH was varied to protonate E35 (and other carboxylates), though no change was observed for the ^{1}H chemical shift nor for the isotope effects. These observations were also supported by the determination of the pKa values of the carboxylate and ammonium groups in these residues. The authors therefore conclude that two of the salt bridges (K12–E23 and K58–D55) are not formed in solution, while the third (K39–E35) may be weakly formed (though likely at a low population).

Computational methods provide an alternative approach to study the structure and dynamics of ionic interactions in solution (Kumar & Nussinov, 2002). All-atom, explicit solvent molecular dynamics (MD) simulations, in particular, can be used to provide insight into the structure and energetics of biomolecules in an aqueous environment. For such simulations to provide an accurate description of salt bridges it is, however, important that the energy function (force field) used provides an accurate description of the balance between the many forces that determine salt bridge strengths. As part of the re-parameterization and validation of one of the more accurate force fields, CHARMM22*, we discovered that ionic interaction between guanidinium and acetate ions were ~10 times too strong in the original CHARMM22 force field (Piana, Lindorff-Larsen & Shaw, 2011). This observation led to a re-parameterization (the “DER correction”) of the partial charges in Asp, Glu and Arg side chains, to bring the interactions closer to experiment (but still slightly overestimated) (Piana, Lindorff-Larsen & Shaw, 2011). Building upon these ideas, Debiec and colleagues examined the interactions between analogues of both arginine, lysine and histidine side chains with carboxylates in a range of force field and water models (Debiec, Gronenborn & Chong, 2014). Although they found considerable variability between different force fields, the general observation from the CHARMM force fields hold true, namely that the ionic interactions between side chain analogues are overstabilized relative to experiment. Such overstabilization is not without consequence when studying protein dynamics using simulations. For example, in a study of voltage gating in potassium channels the simulations were performed with the DER correction, or the even more substantial DER2 correction, in order to be able to observe the switch in ionic interactions that are central in channel gating

(Jensen et al., 2012). Similarly, when using free energy perturbation MD simulations to predict the change in protein stability from a mutation, the results are less accurate when the mutations involve a change in charge (Steinbrecher et al., 2017).
Here we build on the work described above by examining the extent to which five different force fields capture the behaviour of the three putative salt bridges in GB1. In particular, we conducted explicit-solvent molecular simulations of GB1 using the Amber ff99SB*-ILDN, Amber ff03, CHARMM27, CHARMM22* and OPLS_2005 force fields. In contrast to studies of side chain analogues, simulations of the salt bridges in the context of a folded protein takes into account the natural geometrical and energetic constraints imposed by the protein scaffold. To accelerate sampling and ensure convergence we used metadynamics simulations to map the free energy surface of salt bridge formation. The results reveal that also in the context of a folded protein do these force fields overestimate the population of ionic interactions, and also provide insight into the geometry of salt bridges in proteins.

Materials and Methods

In order to investigate potential salt-bridging interactions between residues K12-E23, K39-E35 and K58-D55 in GB1 we performed both standard MD simulations as well as enhanced sampling simulations using well-tempered metadynamics (WT-MetaD) (Barducci, Bussi & Parrinello, 2008). To test the effect of the force field we used five different force fields: Amber ff99SB*-ILDN (Hornak et al., 2006; Best & Hummer, 2009; Lindorff-Larsen et al., 2010), Amber ff03w (Best & Mittal, 2010), CHARMM27 (CHARMM22 with the CMAP correction) (MacKerell et al., 1998; Mackerell, Feig & Brooks, 2004), CHARMM22* (Piana, Lindorff-Larsen & Shaw, 2011) and OPLS_2005 (Banks et al., 2005) (Table 1). Each force field was combined with its "native" water model, namely the CHARMM-specific TIP3P water model for CHARMM force fields (MacKerell et al., 1998) (herein called TIPS3P), the original TIP3P water model (Jorgensen et al., 1983) for the Amber ff99SB*-ILDN force fields, TIP4P/2005 (Abascal & Vega, 2005) for Amber ff03w and the SPC/E water model (Berendsen, Grigera & Straatsma, 1987) for OPLS. We also examined the effect of the water model by testing CHARMM22* with standard TIP3P.

Setup and Equilibration

All simulations were performed with Gromacs version 4.5 (van der Spoel et al., 2010), using the crystal structure of GB1 (PDB ID 1PGB; (Gallagher et al., 1994)) as starting structure. The protein was centred and solvated in a dodecahedral box with an edge length of 12Å, and the resulting system consisted of the protein (56 residues), ~5300 water molecules and 4 sodium ions to neutralize the total charge. The cutoff for the Van der Waals interactions was 9Å, while the long-range electrostatic interactions were calculated using the particle mesh Ewald method with a real-space cutoff of 12Å. For each force field and water model, we first subjected the system to 0.2 ns energy minimization, followed by a 1 ns solvent equilibration with position restraints on the protein backbone, followed by 1ns of protein equilibration. After equilibration, we performed a 10 ns simulation in the NPT ensemble (Parrinello & Rahman, 1981, 1982; Bussi, Donadio & Parrinello, 2007) at 298K and 1atm pressure and calculated the average volume of the system. Finally, we selected the frame with a volume closest to this average from the second half of these 10 ns, and used this as starting point for WT-MetaD simulations in the NVT ensemble (Bussi, Donadio & Parrinello, 2007). The lengths of all bonds to hydrogen atoms were kept fixed using the LINCS algorithm (Hess et al., 1997) , and simulations were performed using a 2 fs time step.

Molecular Dynamics Simulations

Following equilibration, we performed a standard 100 ns unbiased MD simulation with each force field. This was followed by three 100ns WT-MetaD simulations at 298K using the PLUMED 1.3 plugin (Bonomi et al., 2009), and in each simulation enhancing separately the sampling of one the three salt-bridges (K12–E23, K39–E35 and K58–D55). In metadynamics simulations, sampling is enhanced by adding a history-dependent biasing potential along one or more selected degree of freedom (collective variables, CVs). In order to sample the free energy of forming the salt bridges, we used two CVs for each of the three ion-pairs (in three different simulations). For the lysine-glutamate ion pairs we used the distance between the NZ–OE1 and NZ–OE2 atoms as CVs, and for the lysine-aspartate pair we used the distance between the NZ–OD1 and NZ–OD2 atoms. For obvious symmetry reasons, the free energy surface should be identical for the two NZ-oxygen pairs (e.g. NZ-OD1/OD2), and so the use of two CVs for each salt bridge allows us both to map the interactions with various combinations of distances to the two oxygens, but also to assess convergence via the similarity between the symmetry-related atoms. In the metadynamics simulations, the biasfactor, sigma parameters, initial height of the Gaussian hills and the deposition rate were set to 4, 0.05 Å, 0.12 kcal mol^{-1} and 2 ps, respectively. In order not to enhance sampling of unrealistically long distances and to cause unfolding of GB1, we introduced a restraining potential (a soft "wall") on the CVs, using the form $E_{restr}=k(d_i-d_{max})^4$. This potential acted whenever the distance, d_i, was above the longest distance observed in the unbiased simulations (d_{max}=13.5 Å) and used a force constant of k= 4.8 kcal mol^{-1} Å^{-4}. Finally, to test the robustness of the calculations, we also performed a 50-ns parallel-tempering metadynamics (PT-MetaD) (Hansmann, 1997; Bussi et al., 2006) simulation for the K39–E35 salt bridge using the CHARMM22* force field and TIPS3P water model. In particular, in the PT-MetaD simulations, we further enhanced the sampling with exchanges between different temperatures with a replica-exchange scheme. We used six replicas (at 294K, 298K, 308K, 322K, 337K and 353K) where the width of the energy distribution (of all but the "neutral" 298K replica) was increased in a preparatory step of 10 ns as previously described (Sutto & Gervasio, 2013). All replicas were also subjected to an additional biasing force through metadynamics with the same parameters (initial height of the Gaussian hills, deposition rate and biasfactor) used in the WT-MetaD simulations described above.

Analyses

We analysed the salt bridge formation by creating two-dimensional free energy profiles for each salt bridge and for each force field, and using the two nitrogen–oxygen distances as coordinates. To quantify the formation of each salt bridge, we calculated the fraction of frames (after removing the metadynamics bias) where one of the two distances were below 5 Å.

Results and Discussion

Molecular dynamics simulations of salt bridge formation

We performed MD simulations of GB1 using six different combinations of force fields and water models (Table 1) to examine the formation of salt bridges, and to benchmark against NMR experiments. While unbiased MD simulations showed reversible formation and breaking of the ionic interactions (Fig. 2A), we decided to use enhanced sampling metadynamics simulations (Fig. 2B) to ensure better convergence of the free energy of salt bridge formation. In both types of simulations, we find that the carboxylate and ammonium

group spend time at multiple, relatively distinct sets of distances, though with differences between force fields and salt bridges (see below). Thus, for each of the six different force field combinations we performed three metadynamics simulations, one for each salt bridge pair. In each of these 18 simulations, we simultaneously biased and monitored two CVs, corresponding to the distance between the ammonium nitrogen atom in the lysine side chain, and each of the two oxygen atoms in the carboxylate group in aspartate or glutamate. The metadynamics bias did not have any major effect on the protein stability over the course of the 100ns simulations. For example, the average backbone RMSD was below 1.5 Å for both the unbiased and biased simulations (Fig. S1).

Free-energy surfaces and geometry of ionic interactions

We used the metadynamics simulations to reconstruct the free energy surface for salt bridge formation along the two CVs (Fig. 1 and Fig. S2). The resulting profiles show a number of notable features. First, we find that they are highly symmetric across the "diagonal", as expected because of the equivalence of the two carboxylate oxygen atoms. Thus, while this behaviour is expected based on physical grounds, it provides a useful and independent test for convergence. Secondly, the profiles reveal a number of distinct free energy minima with depths of ~1 kcal mol^{-1} and separated by small free energy barriers. For example, in the case of the K12–E23 interaction in simulations using CHARMM22* and the TIPS3P water model (Fig. 1B), three minima are visible at short distances between the two charged side chains. Two of these are symmetrically placed on either side of the diagonal and correspond to one oxygen–nitrogen distance around 3.0 Å and the other at ~4.5 Å. These minima thus correspond to a salt bridge where one oxygen is close to the nitrogen (with a proton from the ammonium group in between) and the other oxygen further away. A third minimum is also observed where both oxygen–nitrogen distances are short (~ 2.5 Å – 3.0 Å), corresponding to a different salt bridge geometry. These three minima correspond to so called contact ion pairs where the cation and anion are in direct contact. In the case of the K39–E35 and K58–D55 pairs, additional minima are observed where the shortest of the two oxygen–nitrogen distances is ~ 6 Å – 10 Å (Figs. 1C and 1D). These correspond to solvent-separated ion pairs, where one or more solvent molecules sit between the amino acid side chains between the pairs and effecting the electrostatic interacting of the cation and anion (Collins, 1997; Marcus & Hefter, 2006; Zhou & Pang, 2017).

Comparing salt bridge stability in simulations and experiments

The free energy profiles suggest that the three salt bridges are formed and are relatively stable in all force fields, though with variations both in geometry and the depths of the different minima. These observations appear to be in conflict with the experimental NMR data that suggest that the K12–E23 and K58–D55 pairs do not form any substantial ionic interactions, and with at most only modest interactions between the charges in K39–E35.
As it is difficult to calculate directly the chemical shifts and isotope effects from the simulations, we opted to compare experiments and simulations indirectly by calculating the populations of the salt bridges using the free energy surfaces. In order to be conservative and not overestimate the calculated values by including e.g. solvent mediated interactions (as it is unclear how much they contribute to the experiments), we included only conformations where the nitrogen–oxygen distance is < 5.0 Å. The resulting populations vary between force fields and salt bridges, and range from ~10% to almost 100%, with an average value of ~60% (Table 1). As detailed below, we estimate the error of the calculated free energies of salt bridge formation to be ~0.5 kcal mol^{-1}, corresponding to errors of the populations of ~20% (down to 10% for the most skewed populations). Surprisingly, the K12–E23 and K58–D55 pairs form the most stable salt bridges in the simulations (~75% on average across force

fields) with the K39–E35 pair being distinctly less stable (on average ~40%). These observations are in clear deviation from those expected from the experimental observations, in that the simulations appear both to overestimate the stability of the salt bridges, and that the order of salt bridge stability also appears to be wrongly predicted.
For the K12–E23 and K58–D55 pairs we observe relatively small differences between force fields. In contrast, we observe greater differences for intra-helical K39–E35 interaction. While we do not have a systematic explanation of these differences, we note that K39 also forms transient interactions with other residues (including E48 and N43), suggesting that the variations observed might be the cumulative effect of a number of differences between the force fields. We find that the ionic interactions are slightly stronger in TIP3PS than in standard TIP3P, in line with previous observations from model compounds (Debiec, Gronenborn & Chong, 2014). As also observed before (Piana, Lindorff-Larsen & Shaw, 2011; Debiec, Gronenborn & Chong, 2014), we find that the DER correction in CHARMM22* decreases the strength of salt bridges to be more in line with the experiment.

Assessing convergence of simulations and effect of simulation parameters

When MD simulations are in apparent disagreement to experiments, it is important to assess whether these differences are due to the force field and simulation parameters, or whether they can be explained e.g. by insufficient sampling. While it is difficult to prove that simulations are converged, we performed several tests that suggest that our observations are rather robust. First, as discussed above, the high degree of symmetry of the calculated free energy surfaces suggest that the individual nitrogen–oxygen distance distributions are relatively converged. Second, we calculated the free energy of salt bridge formation during the simulation (from the populations where the nitrogen–oxygen distance is < 5.0 Å) and monitored the time evolution of this (Fig. 3 shows the results from three WT-MetaD simulations with CHARMM22*/TIPS3P). While the values fluctuate for the different salt bridges, they are relatively stable in the second half of the simulations. Finally, for the K35–E39 salt bridge we also performed a parallel-tempering metadynamics simulation in CHARMM22*/TIPS3P, and compared the results to that from the WT-MetaD (Fig. 3). The two simulations converge to essentially the same free energy difference. Based on these observations we estimate an error of about 0.5 kcal mol^{-1}, corresponding roughly to an error of 0.2 for the population of the salt bridges when the free energy of salt bridge formation is close to zero.
Another key parameter in a simulation is the cutoff used to truncate the Lennard-Jones interactions and to switch between a direct calculation of electrostatic forces and the calculations of longer-range electrostatics using Ewald summation. For the K12–E23 and K35–E39 salt bridges and the CHARMM22*/TIPS3P force field we therefore repeated the WT-MetaD simulations varying this cutoff between 9.0 Å – 14.0 Å and calculated the free energy of salt bridge formation (Fig. 4). While the results vary slightly and in accordance with the estimated uncertainty of our simulations, we find no systematic dependency of the free energy differences on the cutoff used. Thus, these calculations suggest that the cutoff used in the simulations are sufficient to obtain reasonable results.

Conclusions

Ionic salt bridge interactions are pervasive in experimental protein structures and have in certain cases been shown to contribute both to protein stability (Vaughan et al., 2002) and fast association kinetics (Schreiber, Haran & Zhou, 2009). Nevertheless, our understanding of the geometry and energetics of salt bridges in solution is limited by the difficulty in experimental and computational studies. While molecular dynamics simulations in explicit solvent may in principle be a quantitative and predictive model for salt bridge formation, the

accuracy of such simulations hinges upon the force fields used. Building upon earlier work on model compounds (Piana, Lindorff-Larsen & Shaw, 2011; Debiec, Gronenborn & Chong, 2014), we have here performed simulations of GB1 to determine the free energy landscape of lysine-carboxylate salt bridge formation. Comparison with experimental NMR data suggest that while the force fields recapitulate the transient and weak nature of these solvent exposed ionic interactions, they also appear to overestimate their stability slightly. This observation is in line with those from the model compounds, suggesting that together these kinds of calculations might also be used to improve the force fields. Indeed, a newly developed charge model for the Amber force field appears to provide a more balanced description of salt bridge interactions, though still with a slight over estimation (Debiec et al., 2016). We thus urge practitioners of MD simulations to take the small, but significant force field bias into account when interpreting the importance of salt bridges observed in simulations, unless such observations are supported by experimental data. Finally, we hope that experimentalists will continue to develop approaches to study electrostatic interactions in proteins, in particular experiments that can be compared directly to simulations. Recent examples include extension of the GB1 studies to salt bridges in barnase (Williamson et al., 2013), novel NMR methods for studying electrostatics (Hass & Mulder, 2015), NMR methods to study arginine side chains (Mackenzie & Hansen, 2017; Yoshimura et al., 2017), and the engineering of a protein without titratable side chains as a platform for studies of protein electrostatics (Højgaard et al., 2016).

Acknowledgements

We thank the members of the Lindorff-Larsen group for advice and discussions.

References

Abascal JL., Vega C. 2005. A general purpose model for the condensed phases of water: TIP4P/2005. *The Journal of chemical physics* 123:234505. DOI: 10.1063/1.2121687.

Banks JL., Beard HS., Cao Y., Cho AE., Damm W., Farid R., Felts AK., Halgren TA., Mainz DT., Maple JR., Murphy R., Philipp DM., Repasky MP., Zhang LY., Berne BJ., Friesner RA., Gallicchio E., Levy RM. 2005. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *Journal of Computational Chemistry* 26:1752–1780. DOI: 10.1002/jcc.20292.

Barducci A., Bussi G., Parrinello M. 2008. Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Physical Review Letters* 100. DOI: 10.1103/PhysRevLett.100.020603.

Barlow DJ., Thornton JM. 1983. Ion-pairs in proteins. *Journal of Molecular Biology* 168:867–885. DOI: 10.1016/S0022-2836(83)80079-5.

Berendsen HJC., Grigera JR., Straatsma TP. 1987. The missing term in effective pair potentials. *Journal of Physical Chemistry* 91:6269–6271. DOI: 10.1021/j100308a038.

Best RB., Hummer G. 2009. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *Journal of Physical Chemistry B* 113:9004–9015. DOI: 10.1021/jp901540t.

Best RB., Mittal J. 2010. Protein simulations with an optimized water model: Cooperative helix formation and temperature-induced unfolded state collapse. *Journal of Physical Chemistry B* 114:14916–14923. DOI: 10.1021/jp108618d.

Bonomi M., Branduardi D., Bussi G., Camilloni C., Provasi D., Raiteri P., Donadio D., Marinelli F., Pietrucci F., Broglia RA., Parrinello M. 2009. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications* 180:1961–1972. DOI: 10.1016/j.cpc.2009.05.011.

Bussi G., Donadio D., Parrinello M. 2007. Canonical sampling through velocity rescaling. *Journal of Chemical Physics* 126. DOI: 10.1063/1.2408420.

Bussi G., Gervasio FL., Laio A., Parrinello M. 2006. Free-energy landscape for β hairpin folding from combined parallel tempering and metadynamics. *Journal of the American Chemical Society* 128:13435–13441. DOI: 10.1021/ja062463w.

Casadio R., Vassura M., Tiwari S., Fariselli P., Luigi Martelli P. 2011. Correlating disease-related mutations to their effect on protein stability: A large-scale analysis of the human proteome. *Human Mutation* 32:1161–1170. DOI: 10.1002/humu.21555.

Collins KD. 1997. Charge density-dependent strength of hydration and biological structure. *Biophysical Journal* 72:65–76. DOI: 10.1016/S0006-3495(97)78647-8.

Debiec KT., Cerutti DS., Baker LR., Gronenborn AM., Case DA., Chong LT. 2016. Further along the Road Less Traveled: AMBER ff15ipq, an Original Protein Force Field Built on a Self-Consistent Physical Model. *Journal of Chemical Theory and Computation* 12:3926–3947. DOI: 10.1021/acs.jctc.6b00567.

Debiec KT., Gronenborn AM., Chong LT. 2014. Evaluating the strength of salt bridges: A comparison of current biomolecular force fields. *Journal of Physical Chemistry B* 118:6561–6569. DOI: 10.1021/jp500958r.

Dill KA. 1990. Dominant Forces in Protein Folding. *Biochemistry* 29:7133–7155. DOI: 10.1021/bi00483a001.

Foit L., Morgan GJ., Kern MJ., Steimer LR., von Hacht AA., Titchmarsh J., Warriner SL., Radford SE., Bardwell JCA. 2009. Optimizing Protein Stability In Vivo. *Molecular Cell* 36:861–871. DOI: 10.1016/j.molcel.2009.11.022.

Gallagher T., Alexander P., Bryan P., Gilliland GL. 1994. Two Crystal Structures of the B1 Immunoglobulin-Binding Domain of Streptococcal Protein G and Comparison with NMR. *Biochemistry* 33:4721–4729. DOI: 10.1021/bi00181a032.

Hansmann UHE. 1997. Parallel Tempering Algorithm for Conformational Studies of Biological Molecules. *Chemical Physics Letters* 281:140–150. DOI: 10.1016/S0009-2614(97)01198-6.

Hass MAS., Mulder FAA. 2015. Contemporary NMR Studies of Protein Electrostatics. *Annual Review of Biophysics* 44:53–75. DOI: 10.1146/annurev-biophys-083012-130351.

Hess B., Bekker H., Berendsen HJC., Fraaije JGEM. 1997. LINCS: A Linear Constraint Solver for molecular simulations. *Journal of Computational Chemistry* 18:1463–1472. DOI: 10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

Højgaard C., Kofoed C., Espersen R., Johansson KE., Villa M., Willemoës M., Lindorff-Larsen K., Teilum K., Winther JR. 2016. A Soluble, Folded Protein without Charged Amino Acid Residues. *Biochemistry* 55:3949–3956. DOI: 10.1021/acs.biochem.6b00269.

Hornak V., Abel R., Okur A., Strockbine B., Roitberg A., Simmerling C. 2006. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function and Genetics* 65:712–725. DOI: 10.1002/prot.21123.

Jacobsen CF., Linderstrøm-Lang K. 1949. Salt linkages in proteins. *Nature* 164:411–412. DOI: 10.1038/164411a0.

Jensen M., Jogini V., Borhani DW., Leffler AE., Dror RO., Shaw DE. 2012. Mechanism of voltage gating in potassium channels. *Science* 336:229–233. DOI: 10.1126/science.1216533.

Jorgensen WL., Chandrasekhar J., Madura JD., Impey RW., Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* 79:926–935. DOI: 10.1063/1.445869.

Kumar S., Nussinov R. 2002. Close-range electrostatic interactions in proteins. *ChemBioChem* 3:604–617. DOI: 10.1002/1439-7633(20020703)3:7<604::AID-CBIC604>3.0.CO;2-X.

Linderstrøm-Lang K. 1924. On the ionization of proteins. *CR Trav. Lab. Carlsberg* 15:70–95.

Lindorff-Larsen K., Piana S., Palmo K., Maragakis P., Klepeis JL., Dror RO., Shaw DE. 2010. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function and Bioinformatics* 78:1950–1958. DOI: 10.1002/prot.22711.

Mackenzie HW., Hansen DF. 2017. A13C-detected15N double-quantum NMR experiment to probe arginine side-chain guanidinium15Nηchemical shifts. *Journal of Biomolecular NMR* 69:123–132. DOI: 10.1007/s10858-017-0137-2.

MacKerell AD., Bashford D., Bellott M., Dunbrack RL., 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., Roux B., Schlenkrich M., Smith JC., Stote R., Straub J., Watanabe M., Wiórkiewicz-Kuczera J., Yin D., Karplus M. 1998. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *The Journal of Physical Chemistry B* 102:3586–3616. DOI: 10.1021/jp973084f.

Mackerell AD., Feig M., Brooks CL. 2004. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulation. *Journal of Computational Chemistry* 25:1400–1415. DOI: 10.1002/jcc.20065.

Marcus Y., Hefter G. 2006. Ion pairing. *Chemical Reviews* 106:4585–4621. DOI: 10.1021/cr040087x.

Nielsen S V., Stein A., Dinitzen AB., Papaleo E., Tatham MH., Poulsen EG., Kassem MM., Rasmussen LJ., Lindorff-Larsen K., Hartmann-Petersen R. 2017. Predicting the impact of Lynch syndrome-causing missense mutations from structural calculations. *PLoS Genetics* 13. DOI: 10.1371/journal.pgen.1006739.

Parrinello M., Rahman A. 1981. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* 52:7182–7190. DOI: 10.1063/1.328693.

Parrinello M., Rahman A. 1982. Strain fluctuations and elastic constants. *The Journal of Chemical Physics* 76:2662–2666. DOI: 10.1063/1.443248.

Piana S., Lindorff-Larsen K., Shaw DE. 2011. How robust are protein folding simulations with respect to force field parameterization? *Biophysical Journal* 100. DOI: 10.1016/j.bpj.2011.03.051.

Sarakatsannis JN., Duan Y. 2005. Statistical characterization of salt bridges in proteins. *Proteins: Structure, Function and Genetics* 60:732–739. DOI: 10.1002/prot.20549.

Schreiber G., Haran G., Zhou HX. 2009. Fundamental aspects of protein - Protein association kinetics. *Chemical Reviews* 109:839–860. DOI: 10.1021/cr800373w.

Schueler O., Margalit H. 1995. Conservation of salt bridges in protein families. *Journal of Molecular Biology* 248:125–135. DOI: 10.1006/jmbi.1995.0206.

Serrano L., Horovitz A., Avron B., Bycroft M., Fersht AR. 1990. Estimating the Contribution of Engineered Surface Electrostatic Interactions to Protein Stability by Using Double-Mutant Cycles. *Biochemistry* 29:9343–9352. DOI: 10.1021/bi00492a006.

Speakman, J. B., & Hirst MC. 1931. Constitution of the keratin molecule. *Nature* 128:1073.

van der Spoel D., Lindahl E., Hess B., van Buuren a. R., Apol E., Meulenhoff PJ., Tieleman DP., Sijbers a. LTM., Feenstra K a., van Drunen R., Berendsed HJC. 2010. GROMACS User Manual version 4.5.6. *Www.Gromacs.Org*:7778. DOI: 10.1007/SpringerReference_28001.

Steinbrecher T., Zhu C., Wang L., Abel R., Negron C., Pearlman D., Feyfant E., Duan J., Sherman W. 2017. Predicting the Effect of Amino Acid Single-Point Mutations on Protein Stability—Large-Scale Validation of MD-Based Relative Free Energy Calculations. *Journal of Molecular Biology* 429:948–963. DOI: 10.1016/j.jmb.2016.12.007.

Sutto L., Gervasio FL. 2013. Effects of oncogenic mutations on the conformational free-energy landscape of EGFR kinase. *Proceedings of the National Academy of Sciences* 110:10616–10621. DOI: 10.1073/pnas.1221953110.

Takano K., Tsuchimori K., Yamagata Y., Yutani K. 2000. Contribution of salt bridges near the surface of a protein to the conformational stability. *Biochemistry* 39:12375–12381. DOI: 10.1021/bi000849s.

Tomlinson JH., Ullah S., Hansen PE., Williamson MP. 2009. Characterization of salt bridges to lysines in the protein G B1 domain. *Journal of the American Chemical Society* 131:4674–4684. DOI: 10.1021/ja808223p.

Vaughan CK., Harryson P., Buckle AM., Fersht AR. 2002. A structural double-mutant cycle: Estimating the strength of a buried salt bridge in barnase. *Acta Crystallographica Section D: Biological Crystallography* 58:591–600. DOI: 10.1107/S0907444902001567.

Williamson MP., Hounslow AM., Ford J., Fowler K., Hebditch M., Hansen PE. 2013. Detection of salt bridges to lysines in solution in barnase. *Chem. Commun.* 49:9824–9826. DOI: 10.1039/C3CC45602A.

Yoshimura Y., Oktaviani NA., Yonezawa K., Kamikubo H., Mulder FAA. 2017. Unambiguous Determination of Protein Arginine Ionization States in Solution by NMR Spectroscopy. *Angewandte Chemie - International Edition* 56:239–242. DOI: 10.1002/anie.201609605.

Zhou H., Pang X. 2017. Electrostatic Interactions in Protein Structure, Folding,Binding, and Condensation. *Chemical Reviews*. DOI: 10.1021/acs.chemrev.7b00305.

Table 1

Force field	Water model	Population		
		K12–E23	K39–D35	K58–D55
CHARMM22*	TIPS3P	0.73	0.42	0.61
CHARMM22*	TIP3P	0.67	0.20	0.61
CHARMM27	TIPS3P	0.83	0.14	0.75
Amber ff03w	TIP4P/ 2005	0.60	0.48	0.81
Amber ff99SB*-ILDN	TIP3P	0.63	0.72	0.98
OPLS AA	SPC/E	0.65	0.39	0.98

Figure legends

Figure 1

Formation and stability of salt bridges in the B1 domain of Protein G (GB1). In panel A we show the structure of GB1 in cartoon representaion, highlighting the location of the three salt bridges that are found in crystal structures of GB1. In B–D we show free energy profiles of salt bridge formation as obtained from metadynamics simulations with the CHARMM 22* force field and TIPS3P water model. The different inserts illustrate representative structures and show both contact ion pairs where the two residues are in direct interactions, and solvent-separated ion pairs where one or more water molecules sit between the charged residues.

Figure 2

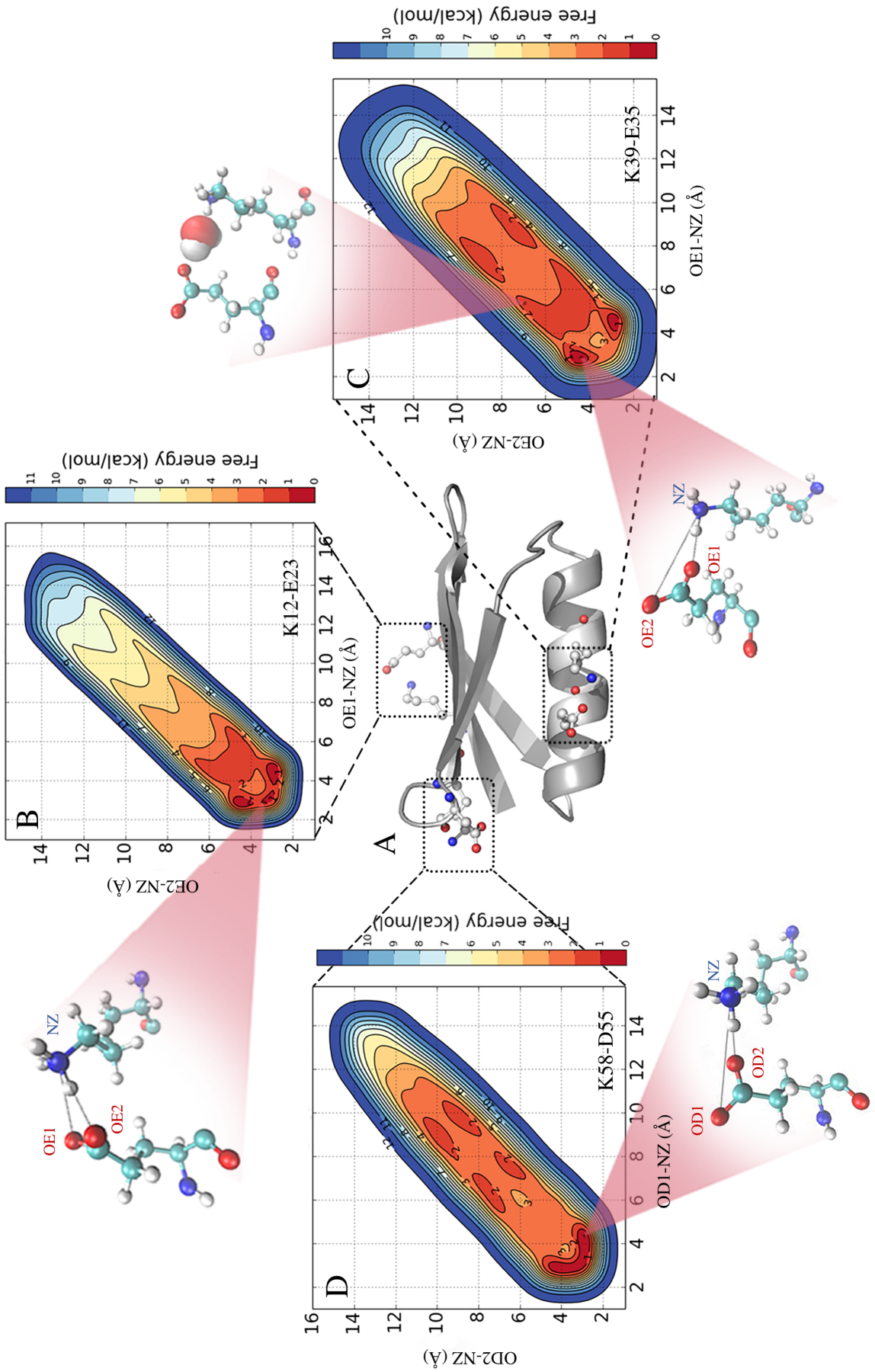
Salt bridge formation using both unbiased and enhanced sampling molecular dynamics simulations. Using the K12–E23 salt-bridge and the CHARMM22*/TIPS3P force field as an example, we compare (A) unbiased simulations with (B) metadynamics simulations. In panel B we also show how the restraint energy acts to avoid that the ion pairs to form excessively long distances. Note that since the metadynamics simulation is biased, the resulting distribution of distances is not expected to be the same until after this bias has been removed. This unbiasing was performed before calculating the free energy profiles in all other figures and analyses.

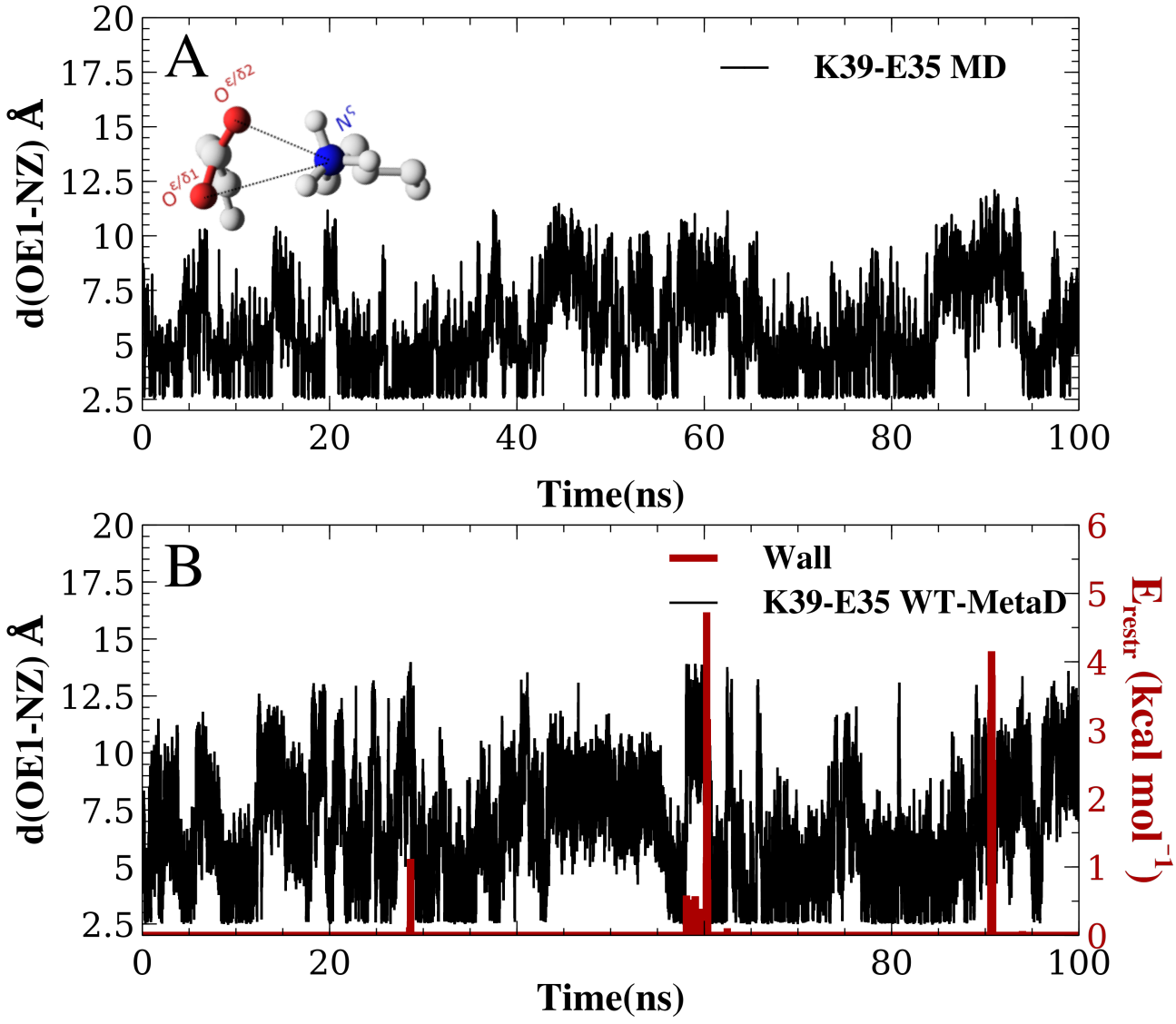
Figure 3

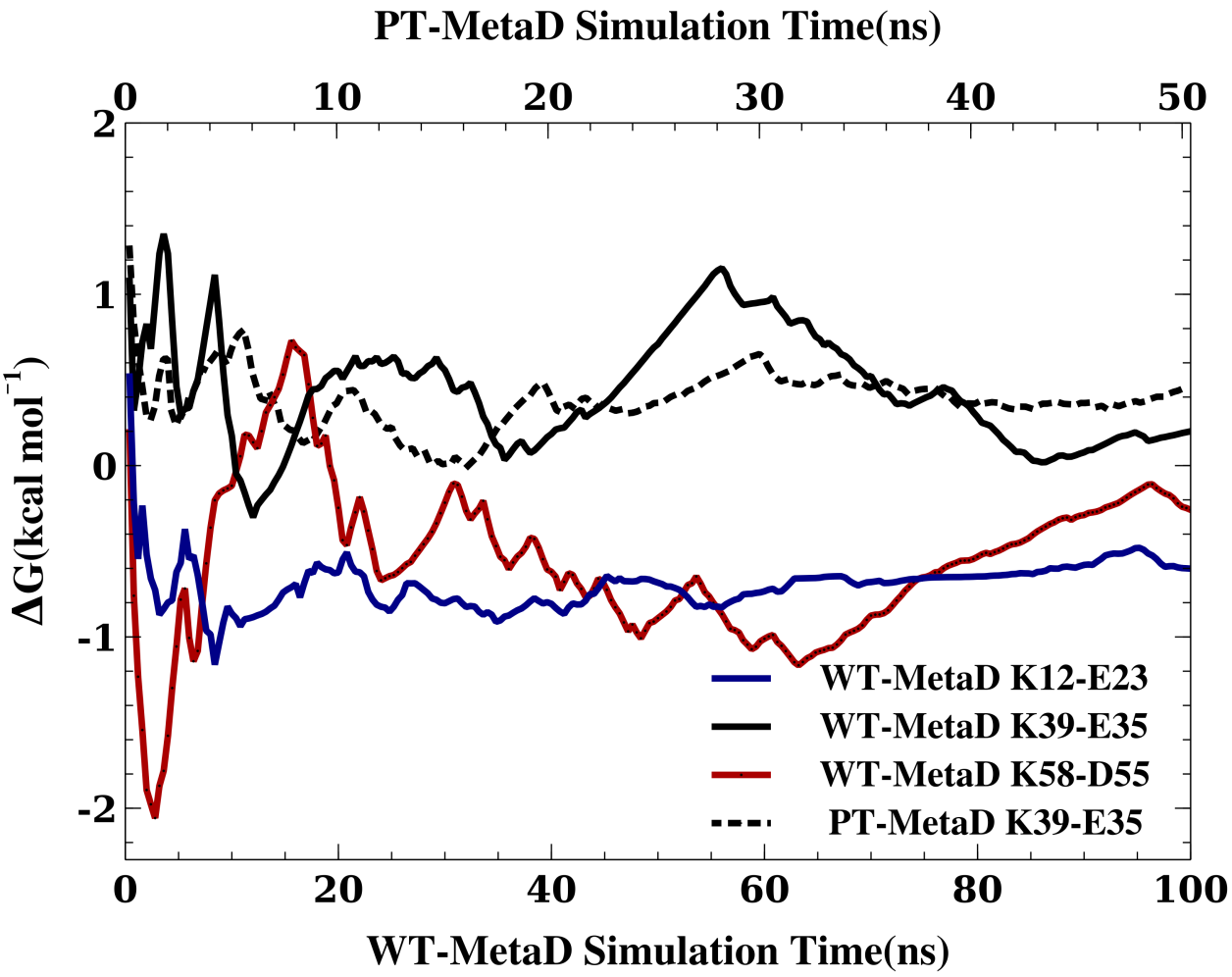
Assessing the convergence of the free energy differences. In the figure we show how the free energy of salt bridge formation as a function of time, using the CHARMM22*/TIPS3P force field as an example. Initially, the values for the three different salt bridges fluctuate, but eventually converge after 100 ns of simulation (solid lines, bottom axis). We also performed a PT-MetaD simulation for the of K39–E35 salt bridge as an alternative approach to determine the free energy landscape (dashed line). After initial fluctuations, the free energy difference converges after 50 ns (top axis) to a value close to that obtained using WT-MetaD.

Figure 4

Assessing the effect of varying the cutoff distance. We repeated the WT-MetaD for two of the salt bridges (K12–E23 and K39–E35) but varying the cutoff for the Lennard-Jones interactions and for switching from the direct calculation of electrostatic interactions to Ewald summation. The figure shows the effect of varying this cutoff on the stability of these two salt bridges.







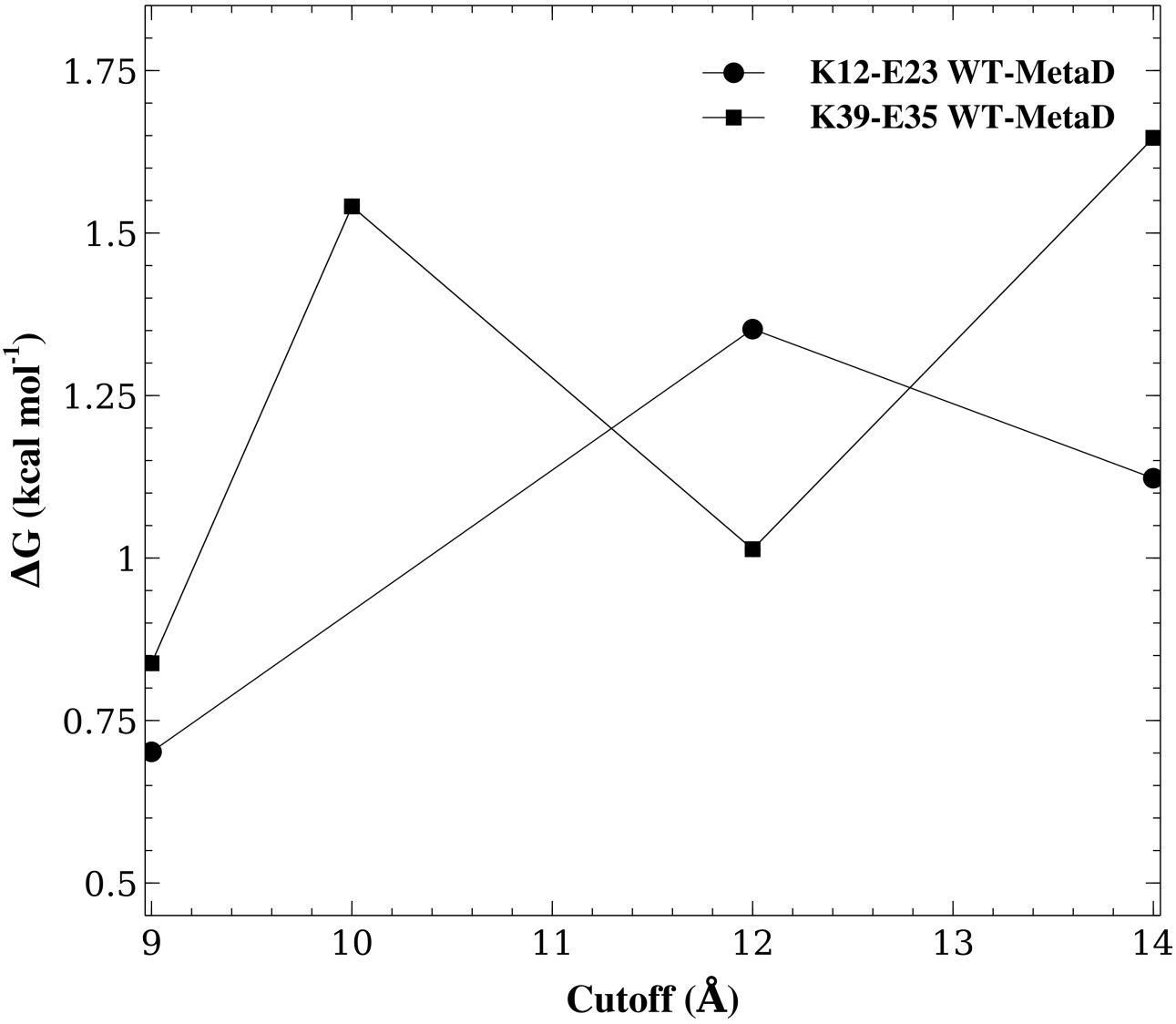
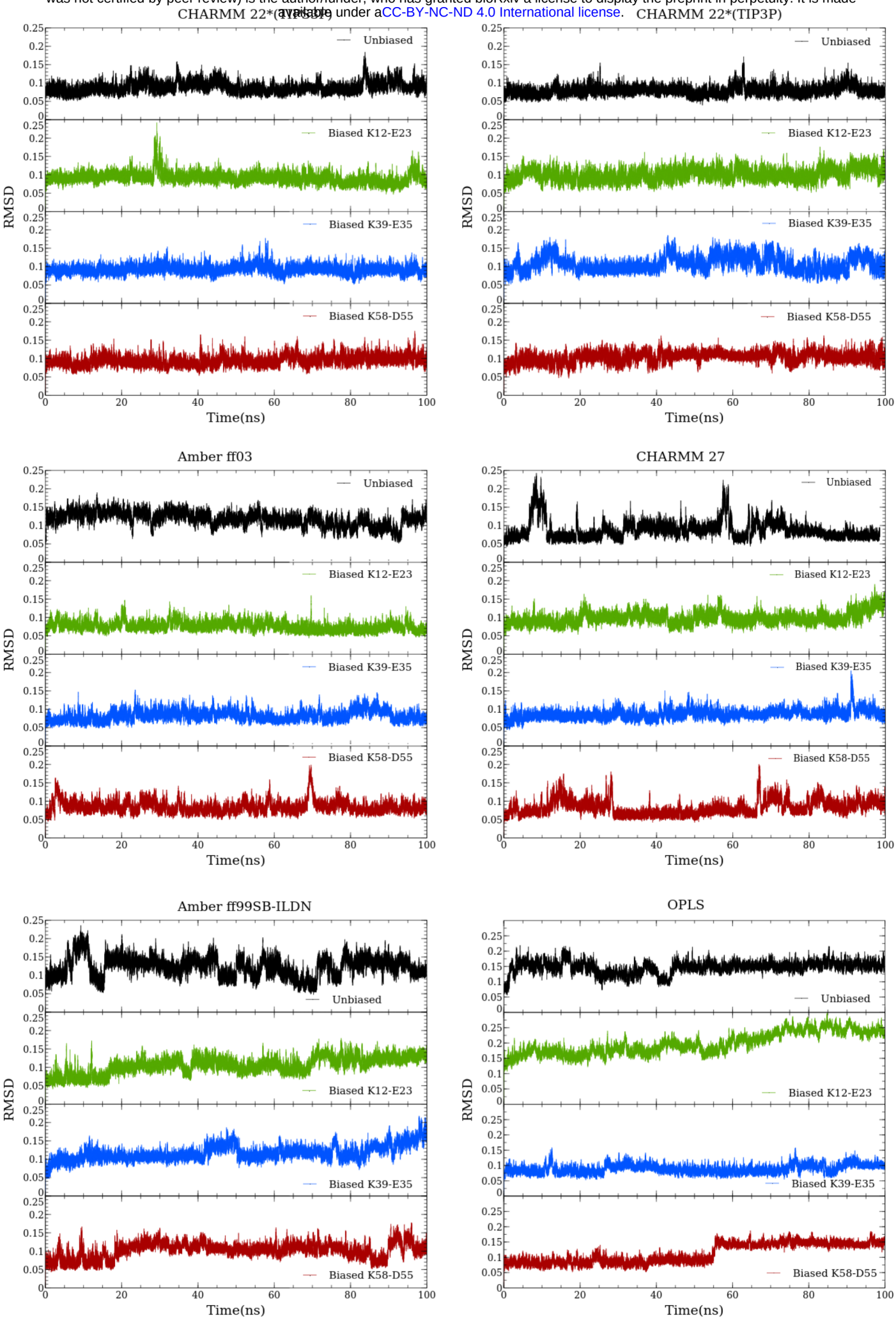


Figure S1

Reconstructed free energy surface from the metadynamic simulations for each force field setup. Salt bridges K12-E23(left), K39-E35(middle) and K58-D55(right).

Figure S2

The backbone RMSD for GB1 for each force field setup. From top to bottom; unbiased, K12-E23, K39-E35 and K58-D55 biased simulation.



K12-E23 K39-E35 K58-D55