Force Field Limitations of All-Atom Continuous Constant pH Molecular Dynamics

Craig A. Peeples, Ruibin Liu, and Jana Shen*

Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, Baltimore, MD 21201

E-mail: jana.shen@rx.umaryland.edu

Abstract

All-atom constant pH molecular dynamics simulations offer a powerful tool for understanding pH-mediated and proton-coupled biological processes. As the protonation equilibria of protein sidechains are shifted by electrostatic interactions and desolvation energies, pK_{a} values calculated from the constant pH simulations may be sensitive to the underlying protein force field and water model. Here we investigated the force field dependence of the all-atom particle mesh Ewald (PME) continuous constant pH (PME-CpHMD) simulations of a mini-protein BBL. The replica-exchange titration simulations based on the Amber ff19sb and ff14sb force fields with the respective water models showed significantly overestimated pK_{a} downshifts for a buried histidine (His166) and for two glutamic acids (Glu141 and Glu161) that are involved in salt-bridge interactions. These errors (due to undersolvation of neutral histidines and over-stabilization of salt bridges) are consistent with the previously reported pK_{a}'s based on the CHARMM c22/CMAP force field, albeit in larger magnitudes. The pK_{a} calculations also demonstrated that ff19sb with OPC water is significantly more accurate than ff14sb with TIP3P water, and the salt-bridge related pK_{a} downshifts can be partially alleviated by the atom-pair specific Lennard-Jones corrections (NBFIX). Together, these data suggest that the accuracies of the protonation equilibria of proteins from constant pH simulations can significantly benefit from improvements of force fields.

Introduction

Solution pH mediates many important biological processes through coupling proton titration with conformational changes of proteins. Over the last two decades, constant pH methods have been developed to rigorously account for solution pH in condensed-phase molecular dynamics (MD) simulations. Instead of fixing the protonation states of protein sidechains to the initial conditions, e.g., histidines with one proton on either the δ or ϵ imidazole nitrogen, constant pH simulations allow protonation states to respond to changes in the electrostatic environment accompanying conformational dynamics. Currently, the main approaches to enable proton-coupled dynamics is through Monte-Carlo (MC) sampling of protonated and deprotonated states (discrete or hybrid MD/MC constant pH methods)[1–5] and an extended Hamiltonian description whereby an auxiliary set of fictitious particles are propagated to represent proton titration (continuous constant pH or λ dynamics based constant pH methods).[6–11] A more detailed discussion of the development of constant pH methods is given in a recent review[12] for more references.

In the early developments of constant

pH methods, various generalized Born (GB)[2,6,7,13] or Poisson-Boltzmann (PB) implicit solvent models are used for both conformational and protonation state sampling. The so-called hybrid-solvent scheme combines the conformational sampling in explicit solvent with implicit-solvent calculation of titration energies[4,14] or forces,[8] resulting in significantly improved accuracies of the calculated pK_{a} values. The more recent developments[11,15–18] utilize explicit solvent for both conformational and protonation state sampling, allowing constant pH simulations of systems that implicit-solvent representations are less accurate for, e.g., nucleic acids,[19,20] surfactant micelles,[21] and proteins in lipid bilayers,[22] although careful reparameterization of dielectric boundary has mitigated some of the issues. Perhaps the most significant drawbacks of using continuum models (either fully implicit-solvent simulations or with the hybrid-solvent scheme) are the failure to account for explicit interactions with ions[23] and the poor computational scaling behavior, n^2, where n is the number of solute atoms, although work of Onufriev and coworker demonstrated that $nlogn$ can be achieved through a Hierarchical Charge Partitioning (HCP) approximation.[24] The first all-atom λ dynamics based CPHMD$^{\mathrm{MS}\lambda\mathrm{D}}$ implementation in CHARMM was applied to calculate the pK_{a}'s of RNAs[25] and peptides inserted in the lipid bilayer.[26]

By removing the aforementioned deficiencies due to the implicit-solvent models, all-atom constant pH simulations can in principle (i.e., given converged simulations) offer more accurate pK_{a} values than implicit- or hybrid-solvent constant pH simulations. The benchmark simulations using the first all-atom particle mesh Ewald (PME) continuous constant pH MD (CpHMD) implementation in CHARMM[16] demonstrated improved correlation between the calculated and experimental pK_{a} shifts relative to the solution (also known as the model) values as compared to its "predecessor", the hybrid-solvent CpHMD in CHARMM,[8] even though the root-mean-square errors (rmse) with respect to the experimental data are similar.

The key physical determinants of protein pK_{a} shifts relative to the model values are electrostatic interactions and desolvation free energies, which have been shown to vary between different protein force fields for proteins and water.[27–32] As such, pK_{a} values derived from constant pH simulations can be used to probe the force field deficiencies.[33] So far, the benchmark simulations of all-atom constant pH methods[15–18,34] have been conducted using the CHARMM c22/CMAP,[35,36] c36,[37] or c36m[38] force fields, while other force fields such as the widely used Amber ff14sb[39] or ff19sb[40] force fields have not been tested. Recently, Sequeira et al. compared the pK_{a} calculations using the hybrid-solvent MD/MC constant pH simulations with the GROMOS 54A7 and CHARMM c36m force fields.[41] Their study[41] demonstrated significant difference between the two force fields in calculating the highly shifted pK_{a}'s of Asp26 in human thioredoxin and Cys32/Cys35 in *E. coli* thioredoxin. Specifically, the c36m calculated pK_{a}'s were very stable over the simulation time (up to 50 ns), whereas the 54A7 calculated pK_{a}'s converged much slower but ultimately achieved better agreement with experiment.

In this work, we investigate the force field dependence of the all-atom PME CpHMD simulations by comparing the titration simulations of a mini-protein BBL (Figure 1) based on the Amber ff19sb,[40] ff14sb,[39] and CHARMM c22/CMAP[35,36] protein force fields and their respective preferred water models, TIP3P,[42] OPC,[43] and CHARMM-style TIP3P.[35] The pK_{a} data based on the c22 force field are taken from our previous work.[34] BBL is a well suited benchmark protein for the constant pH based pK_{a} calculations, as it contains buried histidines and carboxylic acids involved in salt-bridge interactions, for which large pK_{a} shifts relative to solution (also called model) values are expected and challenging to predict accurately.[44] Also, BBL has been previously used to benchmark the performance of the all-atom PME CpHMD implementations[16,34] in CHARMM[45] and Amber programs.[46] Importantly, due to its small size,

conformational sampling is less likely an issue. We found that the largest error for all force fields is for a buried histidine, while glutamic acids involved in salt bridge interactions have overestimated pK_a shifts. Specific ion binding was observed in the ff19sb simulations; however, the use of atom-pair specific corrections to the Lennard Jones parameters (widely known as NBFIX)[47–51] did not reduce the errors of the calculated pK_a's.

Methods and Protocols

Unless otherwise noted, the simulations were performed using the all-atom particle-mesh (PME) CpHMD implementation[34] in Amber24.[46]

Preparation of the protein system. The coordinates of BBL were retrieved from the Protein Data Bank (PDB) entry 1W4H.[52] The first NMR model was used. The NMR model shows no interactions between the terminal groups and any titratable residues. For simulations, the N- and C-termini were acetylated (ACE-) and amidated (-NH2), respectively. Following our previous protocol,[53] the positions of hydrogens were built using the HBUILD command in the CHARMM program,[45] and a custom CHARMM script was used to add dummy hydrogens to the *syn* positions of carboxylate oxygens on all Asp and Glu sidechains. Next, the CHARMM coordinate file was converted to the Amber format. The protein was solvated in a truncated octaheral water box with a minimum of 15 Å between the protein heavy atoms and the water oxygens at the box edges. To neutralize the simulation box at pH 7.5 and reach an ionic strength of 150 mM, 18 Na^+ and 19 Cl^- were added. Note, to accommodate the replica-exchange protocol, the number of counter ions was not adjusted for other pH conditions (see Concluding Discussion).

Minimization, heating, and equilibration of the protein. The *pmemd.cuda* engine of the Amber2024 program[46] was used for simulations. First, the protein system underwent 10,000 steps of energy minimization using the steepest descent (1000 steps) and conjugate gradient (9000 steps) algorithms, whereby the protein heavy atoms were harmonically restrained with a force constant of 100 kcal·mol^{-1}·Å^{-2}. Next, the PME-CpHMD module[34] was turned on and the restrained system was heated from 100 to 300 K over 1 ns at pH 7.5 with a 1 fs timestep in the NVT ensemble. Finally, a two-stage equilibration was performed in the NPT ensemble. In the first stage, the PME-CpHMD simulation was performed at pH 7.5 with a 1-fs time step. A harmonic restraint was placed on heavy atoms, and it was reduced from 100 in the first 1 ns to 10 kcal·mol^{-1}·Å^{-2} in the second 1 ns. In the second stage, 16 pH replicas were placed with increments of 0.5 units in the pH range 1.0–8.5 and simulated independently for 2 ns, employing a 2 fs timestep and harmonic restraints on the backbone heavy atoms. Here the restraints were gradually reduced in four 500-ps steps, from 10 to 5, 2.5, and 0 kcal·mol^{-1}·Å^{-2}.

Production pH replica-exchange CpHMD simulations. In the production run, the asynchronous pH replica-exchange (REX) protocol[54] was turned on using the same 16 pH replicas as in the equilibration stage. The exchange between adjacent pH was attempted every 1000 MD steps (2 ps) according to the Metropolis criterion.[8] The REX simulations were run for 32 ns per replica, with the aggregate simulation time of 512 ns. The pK_a's for all residues were converged (Supplementary Figure S1–S4). The first 10 ns per replica was discarded in the conformational analysis and pK_a calculations.

MD protocol. The SHAKE algorithm was employed to constrain the bonds involving hydrogens whenever a 2-fs step was used. The production runs were performed in the NPT ensemble, where the temperature was maintained at 300 K using the Langevin thermostat

with a collision of 1.0 ps^{-1}, and the pressure was maintained at 1 atm using the Berendsen barostat with a relaxation time of 1 ps. The PME method was used to calculate long-range electrostatics with a real-space cutoff of 12 Å and a 1-Å grid spacing in the reciprocal space calculation. Consistent with the our previous titration simulations with the c22 force field,[34] Lennard-Jones energies and forces were smoothly switched off over the range of 10 to 12 Å.

Model peptide pK_a values and protein pK_a corrections. The model titration parameters for ff19sb and ff14sb simulations of Asp, Glu, and His were taken from the previous work.[34] The validation simulations[34] were conducted using titration simulations of penta-peptides ACE-AlaAlaXAlaAla-NH_2 (X = Asp, Glu, or His) at independent pH conditions with an interval of 0.5 pH in the pH range 2–5.5 for Asp, 2.5–6 for Glu, and 4–8 for His. The simulations lasted 20 ns at each pH. The calculated pK_a's along with the NMR determined values are given in Table 1. Note, ff14sb and ff14sb^{fix} gave nearly identical pK_a's for Asp and Glu penta-peptides. Given the small deviations from the experimental pK_a's, the following post-corrections were made to the calculated protein pK_a's. ff19sb: Asp (+0.35), Glu (+0.33), His (+0.06). ff14sb and ff14sb^{fix}: Asp (+0.16), Glu (+0.18), His (-0.23). This is a caveat, as the corrections are different for individual residue types; the more rigorous way is to offset the model pK_a's before simulations. However, given the small magnitude of corrections (below 0.35 units), the resulting inaccuracies are negligible. Furthermore, for convenience, we retained the original pH scales in plotting the residue-based pH-dependent analysis.

We previously showed that protein pK_a's derived from the PME-CpHMD simulations are affected by the simulation box size, and the finite-size effect decreases with increasing number of water, becoming negligible with large water boxes.[34] Since the current simulations employed a very large water box (15 Å cushion between protein and box edges), the calculated finite-size[16,34] corrections were small (no larger than 0.1 unit). As such, no finite-size corrections were made to the calculated pK_a's.

Table 1: Calculated model peptide pK_a's based on Amber ff19sb and ff14sb force fields[a]

	Asp	Glu	His
ff19sb	3.32±0.07	3.92±0.05	6.48±0.06
ff14sb	3.51±0.11	4.07±0.07	6.71±0.10
NMR	3.67	4.25	6.54

[a]Titration simulations of penta-peptides ACE-AlaAlaXAlaAla-NH_2 (X = Asp, Glu, or His) were performed using the titration parameters from our previous work.[34] NMR derived pK_a's are taken from Ref.[55]

Force field parameters and CpHMD specific changes. The peptides and proteins were represented by the Amber ff19sb[40] or ff14sb[39] protein force field, with the respective OPC[43] or TIP3P[56] model for representing water. The ion parameters were taken Ref.[57] For the simulations with ff14sb, the NBFIX corrections developed by Yoo and Aksimentiev (CUFIX) were applied to the atom pair specific Lennard-Jones parameters to destabilize the attractive interactions between Na^+ and carboxylate (Asp^- or Glu^-),[49] between amine (Lys^+) or guanidinium (Arg^+) and carboxylate (Asp^- or Glu^-).[50,51] The all-atom PME CpHMD implementations in Amber include two force field modifications common to constant pH methods.[2,13,34] First, to allow the single reference concept in constant pH simulations, the partial charges on the backbone are fixed to the values of a single protonation state (Asp^-/Glu^- and Hid/Hie), and the residual charge (ranging from 0.10 to 0.14 e for Asp, Glu and His) is absorbed onto the C-β atom for titration dynamics.[13] Second, the rotation barrier around the carboxylate C-O bond is increased to 6 kcal/mol to prevent the dummy proton from rotating from the syn (initial position in the set up) to the anti position, in which case the proton would lose the ability

to titrate.[2,13]

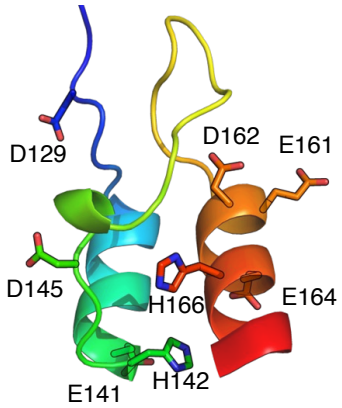

Figure 1: **Structure and titratable sites of BBL.** The NMR model of BBL (PDB ID: 1W4H,[52] first entry) The titratable sidechains are labeled and shown in the stick model.

Results and Discussion

Table 2: Comparison of the calculated pK_a's of BBL using the all-atom PME CpHMD titration based on Amber and CHARMM force fields

Residue	Expt	ff19sb[a]	ff14sb[fix a]	c22[b]	ff14sb/GB[c]
D129	3.9	3.5	3.6	3.6	2.8
E141[d]	4.5	3.4	3.1	4.0	4.1
H142	6.5	6.6	6.5	5.8	6.9
D145	3.7	3.3	2.2	3.2	2.6
E161[d]	3.7	2.6	2.6	4.0	3.3
D162[d]	3.2	3.3	1.7	2.9	3.2
E164	4.5	3.6	3.4	4.0	4.0
H166[d]	5.4	2.4	2.1	4.2	6.0
maxe[e]		-3.0	-3.3	-1.2	-1.1
rmse[e]		1.26	1.58	0.62	0.66
rmse (w/o H166)		0.73	1.14	0.48	0.67

[a] This work. ff19sb refers to the ff19sb protein force field[40] with the OPC water model[43] and Amber default ion force field.[57] ff14sb[fix] refers to the ff14sb protein force field[39] with TIP3P water model[56] and the NBFIX corrected ion force fields.[49–51] [b] Taken from the previous work[34] using CHARMM c22 protein force field[35] with the CHARMM style TIP3P water model[35,56] and the NBFIX corrected ion force fields.[48,58,59] [c] Taken from the previous work[13] using the ff14sb protein force field and GBNeck2 implicit-solvent model.[60] [d] Residues discussed in the main text. [e] Maximal error (maxe) and root-mean-square error (rmse) are calculated with respect to the experimental data.[61,62]

Comparison of the calculated pK_a's based on the Amber and CHARMM force fields and comparison to experiment. BBL has 6 carboxylic acids and 2 histidines. In this work, the calculated pK_a's based on ff19sb[40] gave a root-mean-square error (rmse) of 1.26 pH units with respect to experiment, which is much larger than the rmse (0.62) from our previous PME-CpHMD simulations[34] based on CHARMM c22,[35,36] as well as the rmse (0.66) from our previous[13] GBNeck2[60] simulations with ff14sb[39] (Table 2). Convergence and titration curves of ff19sb simulations are given in Supplementary Figure S1–S4. Note, the rmse of the pK_a's calculated from the preliminary ff14sb (without NBfix) simulations is 1.79, similar to ff19sb (Supplementary Table S1). Curiously, for both ff19sb and c22 simulations, the largest pK_a calculation error is for His166, which has the respective calculated pK_a's of 2.4 and 4.2 (Table 2). Both ff19sb and c22 simulations correctly predicted the pK_a downshift; however, the magnitude is too large compared to experiment (Table 2). For the ff14sb force field, the downshift is also overestimated (Supplementary Table S1). Without His166, the rmse for the ff19sb and c22 simulations are reduced to 0.73 and 0.48, respectively (Table 2), while for ff14sb the rmse without His166 is 1.38 (Supplementary Table S1). By contrast, the GBNeck2 implicit-solvent simulations underestimated the pK_a downshift of His166 by 0.6.

Following His166, the largest pK_a calculation errors with the ff19sb force field are for Glu141 and Glu161, which have the calculated pK_a's of 3.4 and 2.6, compared to 4.0 and 4.0, respectively, from the c22 simulations (Table 2). Compared to the experimental pK_a's of 4.5 and 3.7 for Glu141 and Glu161, the ff19sb simulations underestimated both pK_a's by about 1.1 unit, whereas the c22 simulations underestimated the pK_a's by 0.5 and 0.3 units, respectively. Below we discuss the origins of the underestimated pK_a's for Glu141, Glu161, and His166.

pK_a downshifts of Glu141 and Glu161 are due to salt bridge formation. The pK_a's of Glu141 and Glu161 calculated by the ff19sb simulations are both downshifted with respect to the model value, suggesting that they are involved in attractive electrostatic interactions. Trajectory analysis shows that Glu141 occasionally forms a salt bridge with Arg137, with the occupancy (or probability) increasing from zero at pH 1 to a maximum of nearly 20% near pH 4 where deprotonation is completed (Figure 2a–c). Over the same pH range, the

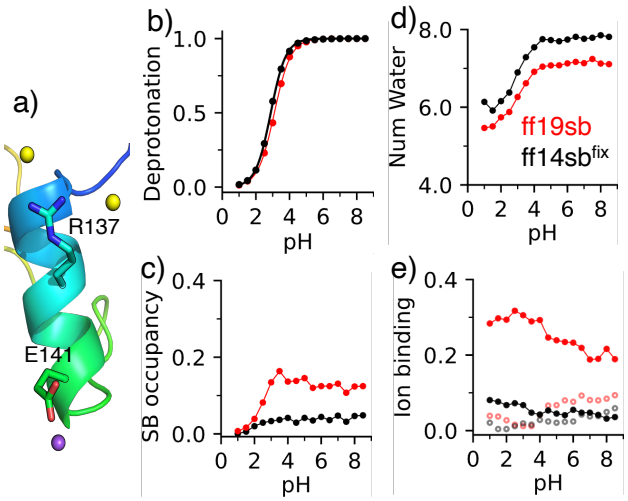

Figure 2: **Titration of Glu141 in BBL and the pH-dependent salt bridge formation, solvent exposure, and ion binding.** a) Snapshot of the bound ions from the ff19sb trajectory at pH 6.5. Na^+ and Cl^-ions are shown as purple and yellow spheres, respectively. b) pH-dependent deprotonated fraction of Glu141. c) pH-dependent occupancy of the salt bridge formation between Glu141 and Arg137. A salt bridge is considered formed if the distance between a carboxylate oxygen of the charged Glu141 and a guanidinium nitrogen is below 4.0 Å. d) Number of water molecules within 3.4 Å of a carboxylate oxygen of Glu141 at different pH. e) Occupancy of Cl^- binding of Arg137 (closed circles) and Na^+ binding of Glu141 (open circles) at different pH. An ion is considered bound if it is within a cutoff distance, 4.2 Å for Cl^- and 4.0 Å for Na^+, from a guanidinium nitrogen of Arg137 or a carboxylate oxygen of Glu141. The data based on the ff19sb and ff14sbfix force fields are shown in red and black, respectively, and do not include pK_a corrections.

solvent exposure of Glu141, defined by the number of water molecules in the first solvent shell of the carboxylate oxygens, also increases and plateaus to about 7 upon complete deprotonation (Figure 2d). Since the first solvent shell of carboxylate oxygens of the model pentapeptide (AAEAA) contains an average of 7 water molecules (SI), these analyses (Figure 2c,d) suggests that Glu141 becomes fully exposed to solvent once the salt bridge is disrupted.

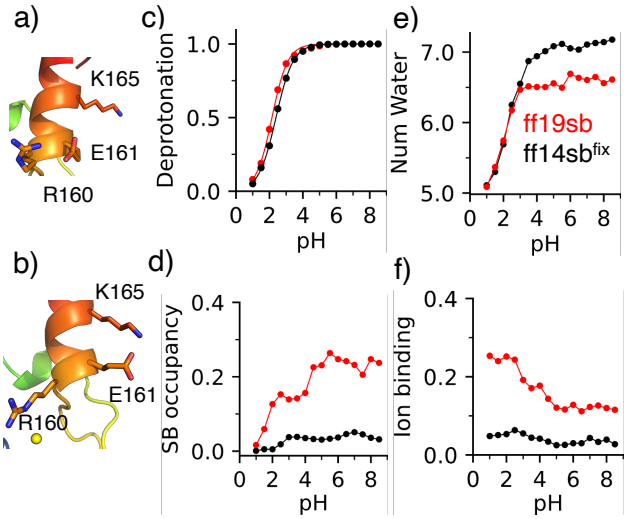

Figure 3: **Titration of Glu161 in BBL and the pH-dependent salt bridge formation, solvent exposure, and ion binding.** a) Snapshot showing a salt bridge between Glu161 and Arg160 taken from the ff19sb trajectory at pH 7. b) Snapshot of Arg160 bound to Cl^- and Glu161 forming a salt bridge with Lys165 from the ff19sb trajectory at pH 7. c) Deprotonated fraction of Glu161 at different pH. d) Salt bridge occupancy of Glu161 with either Arg160 or Lys165 as different pH. e) Number of water molecules within 3.4 Å of any carboxylate oxygen of Glu161 at different pH. f) Occupancy of Cl^- binding to Arg160 at different pH. The data based on ff19sb and ff14sbfix are shown in red and black, respectively, and do not include pK_a corrections.

The strong correlation between the sigmoidal shaped pH profiles of the deprotonation fraction, salt-bridge occupancy, and solvent exposure of Glu141 indicates that the salt-bridge interaction, which stabilizes the charged state, is the main determinant for the pK_a downshift. This phenomenon has been observed in the previous titration simulations using the CpHMD implementations in the CHARMM and AMBER programs with the c22 force field.[16,34] However, since the experimental pK_a of Glu is 4.5, which is 0.3 up shifted relative to the model value, the Glu141–Arg137 salt bridge interaction is likely

overly stabilized by the ff19sb force field, and the overstabilization is to a lesser extent with the c22/CMAP force field as the pK_a underestimation is 0.5 unit smaller.

We next examined Glu161, which shows a pK_a downshift that is 1.2 units too large compared to experiment (Table 2). Similar to Glu141, deprotonation of Glu161 is also strongly correlated with the salt-bridge formation (with either Lys165 or Arg160, Figure 3a–d). The salt-bridge occupancy of Glu161 increases to a maximum of 28% near pH 5 (Figure 3d), which is higher than Glu141 (Figure 2c). Over the same pH range, the number of water molecules in the first solvent shell of Glu161 increases to about 6.5 (Figure 3e), which is slightly lower than that for Glu141 (Figure 2d). The increased salt-bridge interactions and slightly decreased solvent exposure favor the charged state, which may explain the lower pK_a of Glu161 relative to Glu141.

Downshifted pK_a of Asp162 is due to hydrogen bonding. The calculated pK_a's of Asp162 based on the ff19sb and c22 force fields are similar and downshifted relative to the model value by 0.4 and 0.7, respectively. These pK_a downshifts are similar to the experimental downshift of 0.5 units. Trajectory analysis suggests that the pK_a downshift of Asp162 can be attributed to the formation of a hydrogen bond (h-bond) with Thr152 (Figure 4a). The deprotonation of Asp162 occurs over the pH range of 1 to 5 (Figure 4b), where the increased deprotonation is accompanied by the h-bond formation between the carboxylate of Asp162$^-$ and the hydroxyl group of Thr152 which increases from 5.3% to 83.8% (Figure 4c). Note, although protonated carboxylic can also serve as a h-bond acceptor to Thr152, the occupancy is very low (maximum ∼20% at pH 2.5, Supplemental Figure S5). This phenomenon has also been observed in the c22 based PME-CpHMD simulations[16,34] as well as the hybrid-solvent[63] and implicit-solvent CpHMD simulations of a variety of proteins.[64] The similarity between the experimental and calculated pK_a's of Asp162 based on the ff19sb and c22 force fields suggests that the h-bond interaction involving Asp162 may be represented appropriately by the force fields.

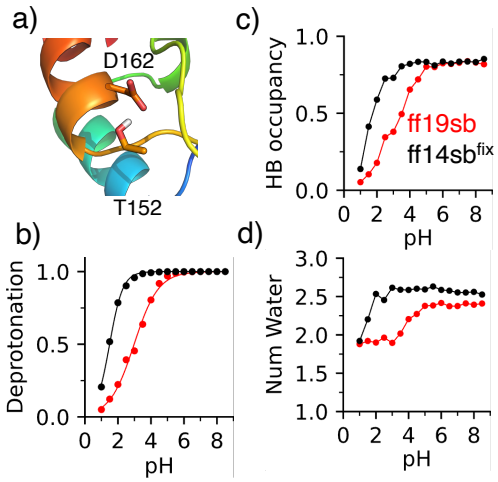

Figure 4: **Titration of Asp162 in BBL and the pH-dependent hydrogen bonding.** a) Snapshot of a hydrogen bond between Thr152 and Asp162 from the ff14sb^{fix} trajectory at pH 7. b) pH-dependent deprotonation fraction of Asp162. c) pH-dependent occupancy of the hydrogen bond (h-bond) formation between the carboxylate of Asp162 and the hydroxyl group of Thr152. A h-bond is considered formed if the distance between the donor and acceptor heavy atoms is below 3.5 Å and the donor−H···acceptor angle is greater than 150°. d) Number of water molecules within 3.4 Å of a carboxylate oxygen of Asp162 at different pH. The data based on the ff19sb and ff14sbfix force fields are shown in red and black, respectively, and do not include pK_a corrections.

pK_a downshift of His166 is due to solvent sequestration. To understand the significant pK_a downshift of His166, we examined its solvent exposure and hydrogen bond (h-bond) environment in the titration simulations (Figure 5a,b). Consistent with the analysis of the c22 titration simulations,[34] the pK_a downshift of His166 can be mainly attributed to solvent displacement. Above pH 3.5, the first solvent shell of His166 includes only 3 water molecules, as His166 becomes protonated with decreasing pH, the number of water molecules in the first solvation shell of His166 increases from about 3 at pH 4 to about 5 at pH 1 (Figure 5c,d). Consistent

with the c22 titration simulations,[34] the h-bond formation of His166 is minimal. At pH 1–2, His166 (at Nδ, Supplemental Figure S6) donates a h-bond to the backbone carbonyl oxygen of Asp162, with an occupancy of ∼11% (Figure 5e).

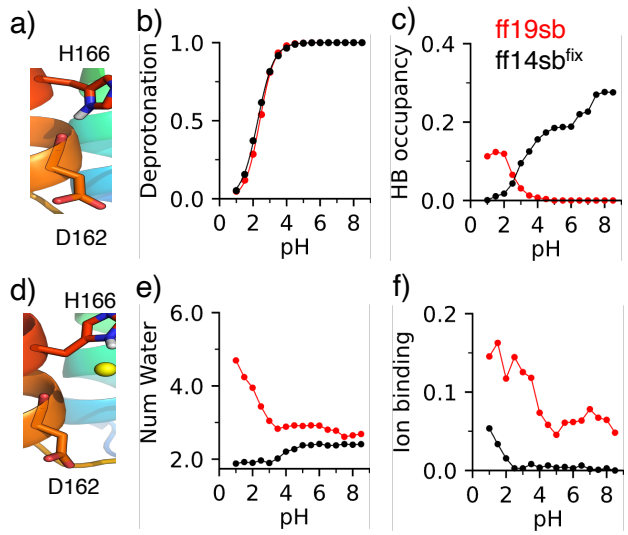

Figure 5: **Titration of His166 in BBL and the pH-dependent hydrogen bonding, solvent exposure, and ion binding.** a) Hydrogen bond formation between the imidazole nitrogen of His166 and the backbone carbonyl oxygen of Asp162 in a snapshot taken from the ff19sb trajectory at pH 2.5. b) The deprotonation fraction of His166 at different pH. c) Occupancy of the hydrogen bond (h-bond, shown in a) between His166 and Asp162. A h-bond is defined by a maximum distance of 3.5 Å between N and O and a minimum angle N−H···O of 150°. d) The imidazole nitrogen on His166 is bound to Cl^- (yellow sphere), shown in a snapshot taken from the ff19sb trajectory at pH 2.5. e) The number of water within 3.4 Å of an imidazole nitrogen on His166 at different pH. f) The occupancy of Cl^- binding of His166 at different pH. A binding event is defined as a Cl^- ion within 4.2 Å of an imidazole nitrogen on His166. The data based on ff19sb and ff14sbfix are shown in red and black, respectively, and do not include pK_a corrections.

NBFIX improves the pK_a calculations. To reduce the binding interactions between ions and proteins and between oppositely charged ions, atom-pair specific corrections to the Lennard-Jones parameters (commonly known as NBFIX) have been introduced to both CHARMM[47,48] and Amber[49–51] force fields. The ff19sb titration simulations showed specific binding of Cl^- with arginine and histidine and Na^+ with glutamic acid (see below), and such interactions were not present in the c22 simulations[34] with the NBFIX corrections.[48]

Since the NBFIX corrections for ff19sb have not been developed, we repeated the simulations of BBL using the NBFIX corrected ff14sb force field (ff14sb^{fix})[49–51] and compared to the pK_a's calculated in the previous work[34] based on the ff14sb force field, which gave a rmse of 1.78 or 1.38 without His166 (Supplemental Table S1). These errors are larger than those based on the ff14sb^{fix} simulations, which gave a rmse of 1.58 or 1.14 without His166 (Supplementary Table S1). It is worth noting that even with the NBFIX corrections, the pK_a errors based on ff14sb are still larger than the ff19sb simulations without the NBFIX corrections (rmse of 1.26 or 0.68 without His166). Below we discuss ion binding to the aforementioned residues, Glu141, Glu161, and Asp162, and His166 in the ff19sb simulations and the effects of using the NBFIX corrections.

NBFIX abolishes ion binding and weakens salt bridges involving Glu141 and Glu161. In the case of Glu141, the ff19sb simulations showed a considerable amount of Cl^- binding to Arg137, which decreases from about 30% at pH below 4 to about 19% at pH 7.0 (Figure 2a and 2e). There is also occasional binding of a Na^+ ion with Glu141, with a maximum occupancy of about 9% (Figure 2a and 2e). The introduction of NBFIX significantly weakened the interactions between Cl^- and Arg137 and between Na^+ and Glu141, with the highest occupancy below 10% for either ion binding in the entire simulation pH range (Figure 2e). At the same time, the salt-bridge interaction between Glu141 and Arg137 is nearly abolished (Figure 2c), resulting in the increased solvent accessibility of Glu141 (Figure 2d). Comparing the pK_a of Glu141 based on ff14sb^{fix} and ff14sb, the 0.2-unit upshift can be attributed to the weakened salt bridge (Supplementary Table S1).

In the case of Glu161, the ff19sb simulations showed significant occupancies of chloride binding to Arg160 in the entire simulation pH range, which were abolished in the ff14sb^{fix} simulations (Figure 3f). At the same time, the salt-bridge interactions of Glu141 with Lys165 and Arg160 were also disrupted (Figure 3d), while the solvent accessibility was slightly increased (by half of a water, Figure 3e). We suggest that the abolishment of the salt-bridge interactions is the major reason for the 0.4-unit reduction in the pK_a downshift for Glu161 in the ff14sb^{fix} as compared to the ff14sb simulations (Supplementary Table S1).

The exaggerated pK_a downshift of Asp162 is reduced by the ff19sb relative to the ff14sb force field. As to Asp162, the ff19sb and ff14sb^{fix} simulations showed no ion binding or salt bridge formation. This explains why the calculated pK_a's based on the ff14sb^{fix} and ff14sb force fields are nearly identical. However, they are respectively 1.6 and 1.7 units lower than the value from the ff19sb simulations (Table 2 and Supplementary Table S1). The lower pK_a may be attributed to the stronger h-bond interaction with the hydroxyl group of Thr152, which stabilizes the deprotonated state of Asp162 (Figure 4c). Consequently, the pH profile of the h-bond occupancy is shifted to a lower pH range, coinciding with the pH range of the aspartic acid deprotonation (Figure 4b).

NBFIX removes chloride binding with His166 and backbone h-bonding may be overstabilized by ff14sb. In the ff19sb simulations, a chloride ion occasionally binds His166 (mostly at the imidazole Nδ), with an occupancy increasing from about 5% above pH 4 to over 15% below pH 2 (Supplementary Figure S5). The pH dependence of ion binding is inversely correlated with the histidine deprotonation, which is consistent with the electrostatic attraction between the positively charged histidine and Cl^- at low pH. The ff14sb^{fix} simulations abolished ion binding (Figure 5f), and increased the pK_a of His166 by 0.2 units relative to the ff14sb simulations (Supplementary Table S1). Interestingly, the pK_a based on ff19sb is 0.3 units higher than ff14sb^{fix} (Table 2). This may be attributed to the increased solvent accessibility, particularly below pH 4 where the number of water molecules is 5 in the ff19sb as compared to 2 in the ff14sb^{fix} simulations; this increased solvation stabilizes the charged state of histidine resulting in a higher pK_a. Furthermore, the h-bonding interaction between His166 (exclusive to Nδ) and the backbone carbonyl of Asp162 is negligible in the ff19sb simulations, whereas its occupancy increases with pH to nearly 30% at pH 8 in the ff14sb^{fix} simulations (Figure 5e). It is conceivable that the increased solvent exposure is due to the disruption of the sidechain-to-backbone h-bond.

Concluding Discussion

In this work, we investigated the force field dependence of the all-atom PME CpHMD simulations using the pK_a calculations for a miniprotein which contains downshifted pK_a's of several carboxylic acids and one histidine. The CHARMM c22 and Amber ff19sb and ff14sb force fields were considered. Our data showed that the ff19sb and ff14sb force fields overestimate the pK_a downshifts of Glu141 and Glu161 involved in salt-bridge interactions and the pK_a downshift of the buried His166. These trends are consistent with the c22 force field but the overestimation by the ff19sb and ff14sb force fields is to a greater extent.

The pK_a data and pH-dependent conformational analysis suggest that the salt-bridge interactions involving Glu may be overly stabilized, which is a known deficiency of protein force fields.[29,32,51,65] Application of the NBFIX corrections that significantly weaken ion binding and salt-bridge interactions in the ff14sb simulations[49–51] demonstrated small upshifts of the calculated pK_a's; however, the reduction in the pK_a downshift remains insufficient. This is a topic that deserves fur-

ther investigation in the future. Compared to the all-atom simulations, the pK_a's of Glu141 and Glu145 from the GBNeck2 simulations are much closer to experiment, although the pK_a downshifts remain slightly overestimated. This improvement may be in part explained by the larger Lys$^+$–Glu$^-$ and Arg$^+$–Glu$^-$ salt-bridge distances due to the difference in the geometry between GB and TIP3P.[60]

The overestimation of the pK_a downshift of His166 from the ff19sb and ff14sb simulations can be explained by the underestimated hydration free energy of the neutral histidine[31] by the ff14sb force field and TIP3P or OPC3[66] (similar to OPC[43]) water model, which suggests that the desolvation free energy of a buried histidine is too large. Interestingly, the pK_a of His166 is improved based on ff19sb relative to ff14sb, which may be attributed to the weakened intramolecular h-bonding and increased solvent accessibility (stronger solute-water interactions) in the OPC water.[43] Compared to the Amber force fields, the hydration free energy of the neutral histidine based on the c22 force field and TIP3P model is closer to experiment, although that of the Hid tautomer is also underestimated.[67] The difference in the hydration free energies between the ff14sb and c22 force fields explains why the pK_a of His166 is shifted lower based on the ff19sb or ff14sb relative to the c22 simulations. In stark contrast to the all-atom CpHMD simulations, the pK_a of His166 is somewhat overestimated in the GBNeck2 simulations, which may be due to the increased solvent exposure as a result of larger conformational fluctuation in the GBNeck2 solvent.[68]

The largest improvement between the ff19sb and ff14sb calculated pK_a's is for Asp162. While ff19sb gives a pK_a in agreement with experiment, the ff14sb or ff14sb^{fix} overestimates the pK_a downshift by 1.5 or 1.6 units. Similar to His166, the h-bond involving Asp162 is weakened in the ff19sb simulations, which may be attributed to solute-water interactions in the OPC water[43]

One caveat of this work is worth mentioning. In the past, we developed titratable ions[11] or water[69] to rigorously account for the fluctuation in system net charge during titration; however, such a scheme was not applied due to slowed convergence and instead PME plasma (i.e., a uniform background charge) was employed for neutralization. To accommodate pH replica exchange, the number of ions was determined at pH 7.5 and not adjusted for low pH conditions, where Asp and Glu become protonated resulting in a net positive charge. Hub et al. demonstrated[70] that invoking plasma stabilizes a test charge in a hexadecane slab due to the artificial background charge density in the low-dielectric region. In the BBL simulations, negatively charged plasma (at low pH) would stabilize positively charged His166 (the only buried residue), shifting its pK_a higher. However, this effect is not observed, likely because the local plasma density is low and the local dielectric constant is substantially higher than typically assumed ($\epsilon = 4$) for the protein interior. Note, the same observation was made in our previous work.[34] Nonetheless, we demonstrated previously[16,34] that a box size dependent systematic shift in pK_a values is present in PME-CpHMD simulations and can be corrected by accounting for the plasma offset potential for water.[71] In this work, no finite-size corrections[16,34] were made due to the use of very large simulation boxes, which was shown[34] to make corrections unnecessary.

Taken together, this study confirms that the accuracy of pK_a calculations using constant pH simulations is dependent on the underlying protein force field and water model. Given sufficient sampling, for example, in the case of the mini-protein BBL, deviations between the calculated and experimental pK_a's are largely reflective of the limitations of the force fields/water models. Salt-bridge overstabilization and underestimation of the histidine neutral state hydration are two common deficiencies of the ff19sb, ff14sb and c22 force fields (combined with their respective water models), although the magnitude of such errors appears to be larger for the ff19sb or ff14sb force fields. The BBL data also confirms that ff19sb with OPC is more accurate

than ff14sb with TIP3P in terms of the balance between intramolecular and intermolecular protein-water interactions. Our work represents an initial effort to understand the force field limitations with the goal of improving the accuracy of all-atom PME CpHMD simulations.

Data Availability

Parameter and simulation input files are freely available on `github.com/JanaShenLab/CpHMD_ff_comparison`.

Acknowledgement

Financial support is provided by the National Institutes of Health (1R35GM148261 and R01CA256557).

References

(1) Baptista, A. M.; Martel, P. J.; Petersen, S. B. Simulation of Protein Conformational Freedom as a Function of pH: Constant-pH Molecular Dynamics Using Implicit Titration. *Proteins* **1997**, *27*, 523–544.

(2) Mongan, J.; Case, D. A.; McCammon, J. A. Constant pH Molecular Dynamics in Generalized Born Implicit Solvent. *J. Comput. Chem.* **2004**, *25*, 2038–2048.

(3) Stern, H. A. Molecular Simulation with Variable Protonation States at Constant pH. *J. Chem. Phys.* **2007**, *126*, 164112.

(4) Swails, J. M.; York, D. M.; Roitberg, A. E. Constant pH Replica Exchange Molecular Dynamics in Explicit Solvent Using Discrete Protonation States: Implementation, Testing, and Validation. *J. Chem. Theory Comput.* **2014**, *10*, 1341–1352.

(5) Chen, Y.; Roux, B. Constant-pH Hybrid Nonequilibrium Molecular Dynamics–Monte Carlo Simulation Method. *J. Chem. Theory Comput.* **2015**, *11*, 3919–3931.

(6) Lee, M. S.; Salsbury, F. R.; Brooks, C. L., III Constant-pH Molecular Dynamics Using Continuous Titration Coordinates. *Proteins* **2004**, *56*, 738–752.

(7) Khandogin, J.; Brooks, C. L., III Constant pH Molecular Dynamics with Proton Tautomerism. *Biophys. J.* **2005**, *89*, 141–157.

(8) Wallace, J. A.; Shen, J. K. Continuous Constant pH Molecular Dynamics in Explicit Solvent with pH-Based Replica Exchange. *J. Chem. Theory Comput.* **2011**, *7*, 2617–2629.

(9) Donnini, S.; Tegeler, F.; Groenhof, G.; Grubmüller, H. Constant pH Molecular Dynamics in Explicit Solvent with λ-Dynamics. *J. Chem. Theory Comput.* **2011**, *7*, 1962–1978.

(10) Goh, G. B.; Knight, J. L.; Brooks, C. L. Constant pH Molecular Dynamics Simulations of Nucleic Acids in Explicit Solvent. *J. Chem. Theory Comput.* **2012**, *8*, 36–46.

(11) Wallace, J. A.; Shen, J. K. Charge-Leveling and Proper Treatment of Long-Range Electrostatics in All-Atom Molecular Dynamics at Constant pH. *J. Chem. Phys.* **2012**, *137*, 184105.

(12) Martins de Oliveira, V.; Liu, R.; Shen, J. Constant pH Molecular Dynamics Simulations: Current Status and Recent Applications. *Curr. Opin. Struct. Biol.* **2022**, *77*, 102498.

(13) Huang, Y.; Harris, R. C.; Shen, J. Generalized Born Based Continuous Constant

pH Molecular Dynamics in Amber: Implementation, Benchmarking and Analysis. *J. Chem. Inf. Model.* **2018**, *58*, 1372–1383.

(14) Baptista, A. M.; Teixeira, V. H.; Soares, C. M. Constant-pH Molecular Dynamics Using Stochastic Titration. *J. Chem. Phys.* **2002**, *117*, 4184–4200.

(15) Goh, G. B.; Hulbert, B. S.; Zhou, H.; Brooks, C. L., III Constant pH Molecular Dynamics of Proteins in Explicit Solvent with Proton Tautomerism: Explicit Solvent CPHMD of Proteins. *Proteins* **2014**, *82*, 1319–1331.

(16) Huang, Y.; Chen, W.; Wallace, J. A.; Shen, J. All-Atom Continuous Constant pH Molecular Dynamics With Particle Mesh Ewald and Titratable Water. *J. Chem. Theory Comput.* **2016**, *12*, 5411–5421.

(17) Radak, B. K.; Chipot, C.; Suh, D.; Jo, S.; Jiang, W.; Phillips, J. C.; Schulten, K.; Roux, B. Constant-pH Molecular Dynamics Simulations for Large Biomolecular Systems. *J. Chem. Theory Comput.* **2017**, *13*, 5933–5944.

(18) Aho, N.; Buslaev, P.; Jansen, A.; Bauer, P.; Groenhof, G.; Hess, B. Scalable Constant pH Molecular Dynamics in GROMACS. *J. Chem. Theory Comput.* **2022**, *18*, 6148–6160.

(19) Chocholoušová, J.; Feig, M. Implicit Solvent Simulations of DNA and DNA-Protein Complexes: Agreement with Explicit Solvent vs Experiment. *J. Phys. Chem. B* **2006**, *110*, 17240–17251.

(20) Nguyen, H.; Pérez, A.; Bermeo, S.; Simmerling, C. Refinement of Generalized Born Implicit Solvation Parameters for Nucleic Acids and Their Complexes with Proteins. *J. Chem. Theory Comput.* **2015**, *11*, 3714.

(21) Wang, Y.; Wallace, J. A.; Koenig, P. H.; Shen, J. K. Molecular Dynamics Simulations of Ionic and Nonionic Surfactant Micelles with a Generalized Born Implicit-Solvent Model. *J. Comput. Chem.* **2011**, *32*, 2348–2358.

(22) Im, W.; Feig, M.; Brooks, C. L., III An Implicit Membrane Generalized Born Theory for the Study of Structure, Stability, and Interactions of Membrane Proteins. *Biophys. J.* **2003**, *85*, 2900–2918.

(23) Huang, Y.; Chen, W.; Dotson, D. L.; Beckstein, O.; Shen, J. Mechanism of pH-dependent Activation of the Sodium-Proton Antiporter NhaA. *Nat. Commun.* **2016**, *7*, 12940.

(24) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Implicit Solvent Model for Million-Atom Atomistic Simulations: Insights into the Organization of 30-Nm Chromatin Fiber. *J. Chem. Theory Comput.* **2016**, *12*, 5946.

(25) Goh, G. B.; Knight, J. L.; Brooks III, C. L. Constant pH Molecular Dynamics Simulations of Nucleic Acids in Explicit Solvent. *J. Chem. Theory Comput.* **2012**, *8*, 36–46.

(26) Panahi, A.; Brooks, C. L. Membrane Environment Modulates the p K_a Values of Transmembrane Helices. *J. Phys. Chem. B* **2015**, *119*, 4601–4607.

(27) Shirts, M. R.; Pande, V. S. Solvation Free Energies of Amino Acid Side Chain Analogs for Common Molecular Mechanics Water Models. *J. Chem. Phys.* **2005**, *122*, 134508.

(28) Hess, B.; van der Vegt, N. F. A. Hydration Thermodynamic Properties of Amino Acid Analogues: A Systematic Comparison of Biomolecular Force Fields and Water Models. *J. Phys. Chem. B* **2006**, *110*, 17616–17626.

(29) Debiec, K. T.; Gronenborn, A. M.; Chong, L. T. Evaluating the Strength of

Salt Bridges: A Comparison of Current Biomolecular Force Fields. *J. Phys. Chem. B* **2014**, *118*, 6561–6569.

(30) Zhang, H.; Jiang, Y.; Cui, Z.; Yin, C. Force Field Benchmark of Amino Acids. 2. Partition Coefficients between Water and Organic Solvents. *J. Chem. Inf. Model.* **2018**, *58*, 1669–1681.

(31) Zhang, H.; Yin, C.; Jiang, Y.; van der Spoel, D. Force Field Benchmark of Amino Acids: I. Hydration and Diffusion in Different Water Models. *J. Chem. Inf. Model.* **2018**, *58*, 1037–1052.

(32) Ahmed, M. C.; Papaleo, E.; Lindorff-Larsen, K. How Well Do Force Fields Capture the Strength of Salt Bridges in Proteins? *PeerJ* **2018**, *6*, e4967.

(33) Shen, J. K. Uncovering Specific Electrostatic Interactions in the Denatured States of Proteins. *Biophys. J.* **2010**, *99*, 924–932.

(34) Harris, J. A.; Liu, R.; Martins de Oliveira, V.; Vázquez-Montelongo, E. A.; Henderson, J. A.; Shen, J. GPU-Accelerated All-Atom Particle-Mesh Ewald Continuous Constant pH Molecular Dynamics in Amber. *J. Chem. Theory Comput.* **2022**, *18*, 7510–7527.

(35) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(36) Mackerell, A. D.; Feig, M.; Brooks, C. L. Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. *J. Comput. Chem.* **2004**, *25*, 1400–1415.

(37) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain χ_1 and χ_2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.

(38) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14*, 71–73.

(39) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.

(40) Tian, C.; Kasavajhala, K.; Belfon, K. A. A.; Raguette, L.; Huang, H.; Migues, A. N.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; Simmerling, C. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* **2020**, *16*, 528–552.

(41) Sequeira, J. G. N.; Rodrigues, F. E. P.; Silva, T. G. D.; Reis, P. B. P. S.; Machuqueiro, M. Extending the Stochastic Titration CpHMD to CHARMM36m. *J. Phys. Chem. B* **2022**, *126*, 7870–7882.

(42) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L.

Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(43) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* **2014**, *5*, 3863–3871.

(44) Alexov, E.; Mehler, E. L.; Baker, N.; M. Baptista, A.; Huang, Y.; Milletti, F.; Erik Nielsen, J.; Farrell, D.; Carstensen, T.; Olsson, M. H. M.; Shen, J. K.; Warwicker, J.; Williams, S.; Word, J. M. Progress in the Prediction of p K_a Values in Proteins: Prediction of p K_a Values in Proteins. *Proteins* **2011**, *79*, 3260–3275.

(45) Hwang, W.; Austin, S. L.; Blondel, A.; Boittier, E. D.; Boresch, S.; Buck, M.; Buckner, J.; Caflisch, A.; Chang, H.-T.; Cheng, X.; Choi, Y. K.; Chu, J.-W.; Crowley, M.; Cui, Q.; Deng, Y.; Devereux, M.; Ding, X.; Feig, M.; Gao, J.; Glowacki, D. R.; Ii, G.; Hamaneh, M. B.; Harder, E. D.; Hayes, R. L.; Huang, J.; Huang, Y.; Im, W.; Islam, S. M.; Jiang, W.; Jones, M. R.; Kaser, S.; Kearns, F. L.; Kern, N. R.; Klauda, J. B.; Lazaridis, T.; Lee, J.; Lemkul, J. A.; Liu, X.; Luo, Y.; MacKerell, A. D.; Major, D. T.; Meuwly, M.; Nam, K.; Nilsson, L.; Ovchinnikov, V.; Paci, E.; Park, S.; Pastor, R. W.; Post, C. B.; Prasad, S.; Pu, J.; Qi, Y.; Rathinavelan, T.; Roe, D. R.; Roux, B.; Rowley, C. N.; Shen, J.; Simmonett, A. C.; Sodt, A. J.; Topfer, K.; Upadhyay, M. CHARMM at 45: Comprehensive, Fast, and Free, a 15-Year Update. *J. Phys. Chem.* **2024**, xxxx.

(46) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T., III; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Huang, Y.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. AMBER 2024. 2024.

(47) Luo, Y.; Roux, B. Simulation of Osmotic Pressure in Concentrated Aqueous Salt Solutions. *J. Phys. Chem. Lett.* **2010**, *1*, 183–189.

(48) Venable, R. M.; Luo, Y.; Gawrisch, K.; Roux, B.; Pastor, R. W. Simulations of Anionic Lipid Membranes: Development of Interaction-Specific Ion Parameters and Validation Using NMR Data. *J. Phys. Chem. B* **2013**, *117*, 10183–10192.

(49) Yoo, J.; Aksimentiev, A. Improved Parametrization of Li^+, Na^+, K^+, and Mg^{2+} Ions for All-Atom Molecular Dynamics Simulations of Nucleic Acid Systems. *J. Phys. Chem. Lett.* **2012**, *3*, 45–50.

(50) Yoo, J.; Aksimentiev, A. Improved Parameterization of Amine–Carboxylate and Amine–Phosphate Interactions for Molecular Dynamics Simulations Using the CHARMM and AMBER Force Fields. *J. Chem. Theory Comput.* **2016**, *12*, 430–443.

(51) Yoo, J.; Aksimentiev, A. New Tricks for Old Dogs: Improving the Accuracy of Biomolecular Force Fields by Pair-Specific Corrections to Non-Bonded Interactions. *Phys. Chem. Chem. Phys.* **2018**, *20*, 8432–8449.

(52) Ferguson, N.; Sharpe, T. D.; Schartau, P. J.; Sato, S.; Allen, M. D.; Johnson, C. M.; Rutherford, T. J.; Fersht, A. R. Ultra-Fast Barrier-limited Fold-

ing in the Peripheral Subunit-binding Domain Family. *J. Mol. Biol.* **2005**, *353*, 427–446.

(53) Henderson, J. A.; Lui, R.; Harris, J. A.; Huang, Y.; de Oliviera, V. M.; Shen, J. A Guide to the Continuous Constant pH Molecular Dynamics Methods in Amber and CHARMM [Article v1.0]. *Liv. J. Comput. Mol. Sci.* **2022**, *4*, 1563.

(54) Henderson, J. A.; Verma, N.; Harris, R. C.; Liu, R.; Shen, J. Assessment of Proton-Coupled Conformational Dynamics of SARS and MERS Coronavirus Papain-like Proteases: Implication for Designing Broad-Spectrum Antiviral Inhibitors. *J. Chem. Phys.* **2020**, *153*, 115101.

(55) Thurlkill, R. L.; Grimsley, G. R.; Scholtz, J. M.; Pace, C. N. pK Values of the Ionizable Groups of Proteins. *Protein Sci.* **2006**, *15*, 1214–1218.

(56) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926.

(57) Joung, I. S.; Cheatham, T. E. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, *112*, 9020–9041.

(58) Beglov, D.; Roux, B. Finite Representation of an Infinite Bulk System: Solvent Boundary Potential for Computer Simulations. *J. Chem. Phys.* **1994**, *100*, 9050–9063.

(59) Lev, B.; Roux, B.; Noskov, S. Y. Relative Free Energies for Hydration of Monovalent Ions from QM and QM/MM Simulations. *J. Chem. Theory Comput.* **2013**, *9*, 4165–4175.

(60) Nguyen, H.; Roe, D. R.; Simmerling, C. Improved Generalized Born Solvent Model Parameters for Protein Simulations. *J. Chem. Theory Comput.* **2013**, *9*, 2020–2034.

(61) Arbely, E.; Rutherford, T. J.; Sharpe, T. D.; Ferguson, N.; Fersht, A. R. Downhill versus Barrier-Limited Folding of BBL 1: Energetic and Structural Perturbation Effects upon Protonation of a Histidine of Unusually Low pKa. *J. Mol. Biol.* **2009**, *387*, 986–992.

(62) Arbely, E.; Rutherford, T. J.; Neuweiler, H.; Sharpe, T. D.; Ferguson, N.; Fersht, A. R. Carboxyl pKa Values and Acid Denaturation of BBL. *J. Mol. Biol.* **2010**, *403*, 313–327.

(63) Huang, Y.; Yue, Z.; Tsai, C.-C.; Henderson, J. A.; Shen, J. Predicting Catalytic Proton Donors and Nucleophiles in Enzymes: How Adding Dynamics Helps Elucidate the Structure-Function Relationships. *J. Phys. Chem. Lett.* **2018**, *9*, 1179–1184.

(64) Liu, R.; Zhan, S.; Che, Y.; Shen, J. Reactivities of the Front Pocket N-Terminal Cap Cysteines in Human Kinases. *J. Med. Chem.* **2022**, *65*, 1525–1535.

(65) Piana, S.; Robustelli, P.; Tan, D.; Chen, S.; Shaw, D. E. Development of a Force Field for the Simulation of Single-Chain Proteins and Protein–Protein Complexes. *J. Chem. Theory Comput.* **2020**, *16*, 2494–2507.

(66) Izadi, S.; Onufriev, A. V. Accuracy Limit of Rigid 3-Point Water Models. *J. Chem. Phys.* **2016**, *145*, 074501.

(67) Zhang, H.; Jiang, Y.; Yan, H.; Yin, C.; Tan, T.; Van Der Spoel, D. Free-Energy Calculations of Ionic Hydration Consistent with the Experimental Hydration Free Energy of the Proton. *J. Phys. Chem. Lett.* **2017**, *8*, 2705–2712.

(68) Nguyen, H.; Maier, J.; Huang, H.; Perrone, V.; Simmerling, C. Folding Simulations for Proteins with Diverse Topolo-

gies Are Accessible in Days with a Physics-Based Force Field and Implicit Solvent. *J. Am. Chem. Soc.* **2014**, *136*, 13959–13962.

(69) Chen, W.; Wallace, J. A.; Yue, Z.; Shen, J. K. Introducing Titratable Water to All-Atom Molecular Dynamics at Constant pH. *Biophys. J.* **2013**, *105*, L15–L17.

(70) Hub, J. S.; de Groot, B. L.; Grubmüller, H.; Groenhof, G. Quantifying Artifacts in Ewald Simulations of Inhomogeneous Systems with a Net Charge. *J. Chem. Theory Comput.* **2014**, *10*, 381–390.

(71) Rocklin, G. J.; Mobley, D. L.; Dill, K. A.; Hünenberger, P. H. Calculating the Binding Free Energies of Charged Species Based on Explicit-Solvent Simulations Employing Lattice-Sum Methods: An Accurate Correction Scheme for Electrostatic Finite-Size Effects. *J. Chem. Phys.* **2013**, *139*, 184103.