SUPPLEMENTAL MATERIAL

Introducing titratable water to all-atom molecular dynamics at constant pH

Wei Chen, Jason A. Wallace, Zhi Yue and Jana K. Shen*
Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD 21201
*Corresponding author. Phone: (410) 706-4187; Fax: (410) 706-5017; E-mail: jshen@rx.umaryland.edu.

Simulation Details

Structure preparation

Three proteins were studied in this work: the 36-residue subdomain of villin headpiece, HP36 (PDB code 1VII), 45-residue binding domain of 2-oxoglutarate dehydrogenase multi-enzyme complex, BBL (PDB code 1W4H), and 129-residue hen egg white lysozyme, HEWL (PDB code 2LZT). The HBUILD facility of CHARMM was used to add hydrogens to the proteins. The N-termini were left free while the C-termini were N-methylamidated. The proteins were solvated in a truncated octahedral water box. The distance between the protein and the edges of the water box was at least 10 Å. Titratable co-ions (hydroxide and hydronium) were added for all titration sites (Asp, Glu, and His) to level the total net charge (1): 4 OH^- for HP36; 2 H_3O^+ and 6 OH^- for BBL; and 1 H_3O^+ and 9 OH^- for HEWL. Na^+ and/or Cl^- ions were then added to neutralize the system and reach the experimental ionic strength: 150 mM for for HP36 (2), 200 mM for BBL (3, 4), and 50 mM for HEWL (5). The model compounds were single amino acids acetylated at the N-terminus (ACE) and N-methylamidated at C-terminus (CT3). The model compounds were solvated in a cubic water box. For the simulation of model Asp and Glu, one hydroxide (interconverting with water) and one Na^+ counter-ion were added to the system, while for model His one hydronium (interconverting with water) and one Cl^- counter-ion were used. The derived parameters for the PMF were used to titrate hydroxide and hydronium in water boxes. Small adjustments were done to achieve reference pK_a's of hydroxide and hydronium. The resulting parameters were then tested on titration of model compounds (Asp, Glu, and His).

Simulation protocol

All simulations were carried out using an in-house modified version of CHARMM (version c36b2) (6). The all-atom CpHMD method was implemented in the PHMD module (1) and the pH-based replica-exchange protocol was added to the REPDSTR module (7). The CHARMM22/CMAP additive force field for proteins (8) and the modified CHARMM TIP3P water model (9) were used. It was shown that CHARMM force field underestimates the dielectric constants of alkanes by almost a factor of 2 (10). A smaller dielectric constant for the protein interior results in an overestimation of the desolvation energies of charged sidechains, thereby contributing to a systematic overestimation of the pK_a shifts for buried residues.

The respective partial charges for water, hydroxide and hydronium were taken from TIP3P (9),

CHARMM22 (11), and the work by Sagnella and Voth (12) (see Table S1). The parameters for the bond and angle energy terms for hydronium were from Sagnella and Voth (12). Other force field parameters for hydroxide and hydronium were adapted from TIP3P. To prevent prolonged attraction between hydroxide and hydronium, the van der Waals interaction distance (R_{min}) between oxygen atoms of hydroxide and hydronium was increased to 4.5 Å. The modified R_{min} for sodium, and for sodium-chloride as well as sodium-carbonyl oxygen pairs by Roux and coworkers were used (13, 14). The SHAKE algorithm was applied to all bonds involving hydrogen to allow a 2-fs time step. Non-bonded interactions were truncated at 14 Å, beyond which the electrostatic interactions were treated by the generalized reaction field (GRF) method (1). In the GRF term, $\varepsilon_{\rm in}$ and $\varepsilon_{\rm out}$ were set to 1.0 and 80.0, respectively, and the ionic strength was set to the experimental values as mentioned above. All simulations employed periodic-boundary conditions, and ambient temperature (300 K) and pressure (1 atm) respectively controlled by the Hoover thermostat (15) and Langevin piston pressure-coupling algorithm (16). The titration coordinates were propagated using the Langevin algorithm with a collision frequency of 5 ps^{-1}. The mass of the fictitious λ particles was set to 10 atomic mass units.

For each protein, one set of CpHMD simulations with the pH replica-exchange sampling protocol (7) was performed. The pH range was 1.5–6.5 for HP36, 0.5–8.0 for BBL, and -1.0–9.5 for HEWL. The pH interval was 0.5. Exchanges between adjacent pH replicas were attempted every 500 MD steps or 1 ps. Simulation length was 10 ns per replica.

Supplemental Tables and Figures

Table S1: Atomic partial charges for water, hydroxide and hydronium

Atom	H_2O	OH^-	H_3O^+
OH_2	-0.834	-1.32	-0.755
H1	0.417	0.32	0.585
H2	0.417	0.0	0.585
H3	-	-	0.585

Table S2: Calculated and reference p$K_{\rm a}$ values for model compounds

Model	Calc p$K_{\rm a}$	Calc Hill	Ref p$K_{\rm a}$
Asp	3.5 $\pm$ 0.17	0.99 $\pm$ 0.08	4.0
Glu	4.1 $\pm$ 0.13	0.96 $\pm$ 0.06	4.4
His	6.8 $\pm$ 0.10 (7.3†/7.0‡)	0.97 $\pm$ 0.01	6.45 (6.6†/7.0‡)

The p$K_{\rm a}$ values were calculated based on five independent sets of pH-REX CpHMD simulations, each of which was run with 5 pH replicas and lasted 10 ns per replica. Reference p$K_{\rm a}$'s for Asp and Glu are taken from Nozaki and Tanford (17), while those for His are taken from Bashford (18). †Nδ site. ‡Nϵ site.

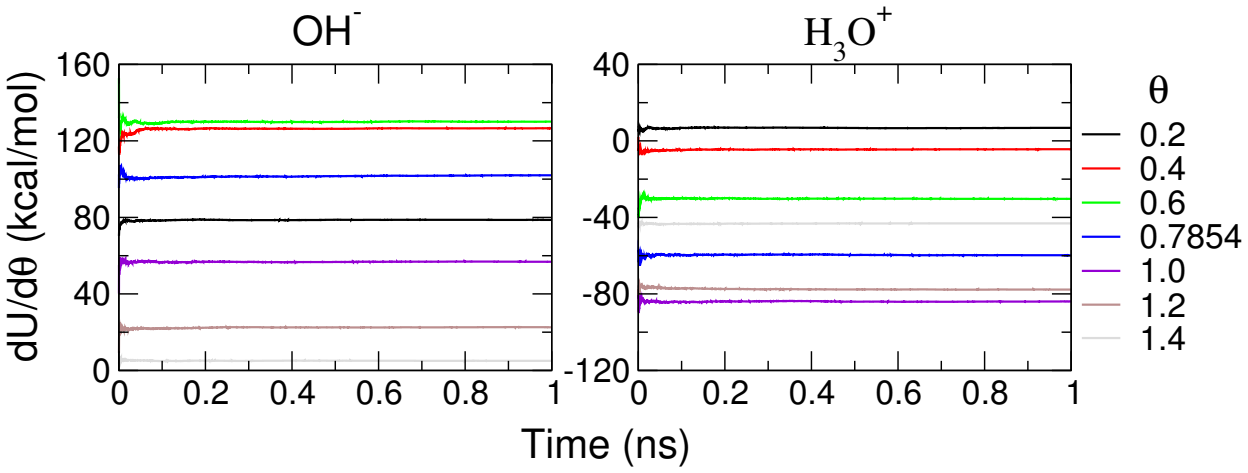

Figure S1: Convergence of the average forces at specified θ values. Calculation was performed cumulatively for hydroxide (*left*) and hydronium (*right*).

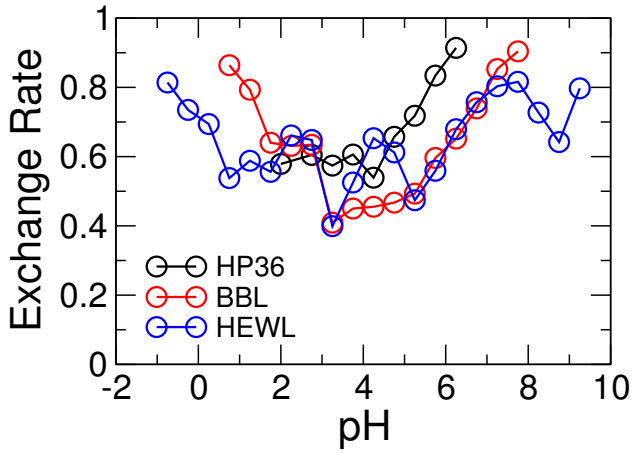

Figure S2: Exchange rates between the adjacent pH replicas in the 10-ns simulations of proteins.

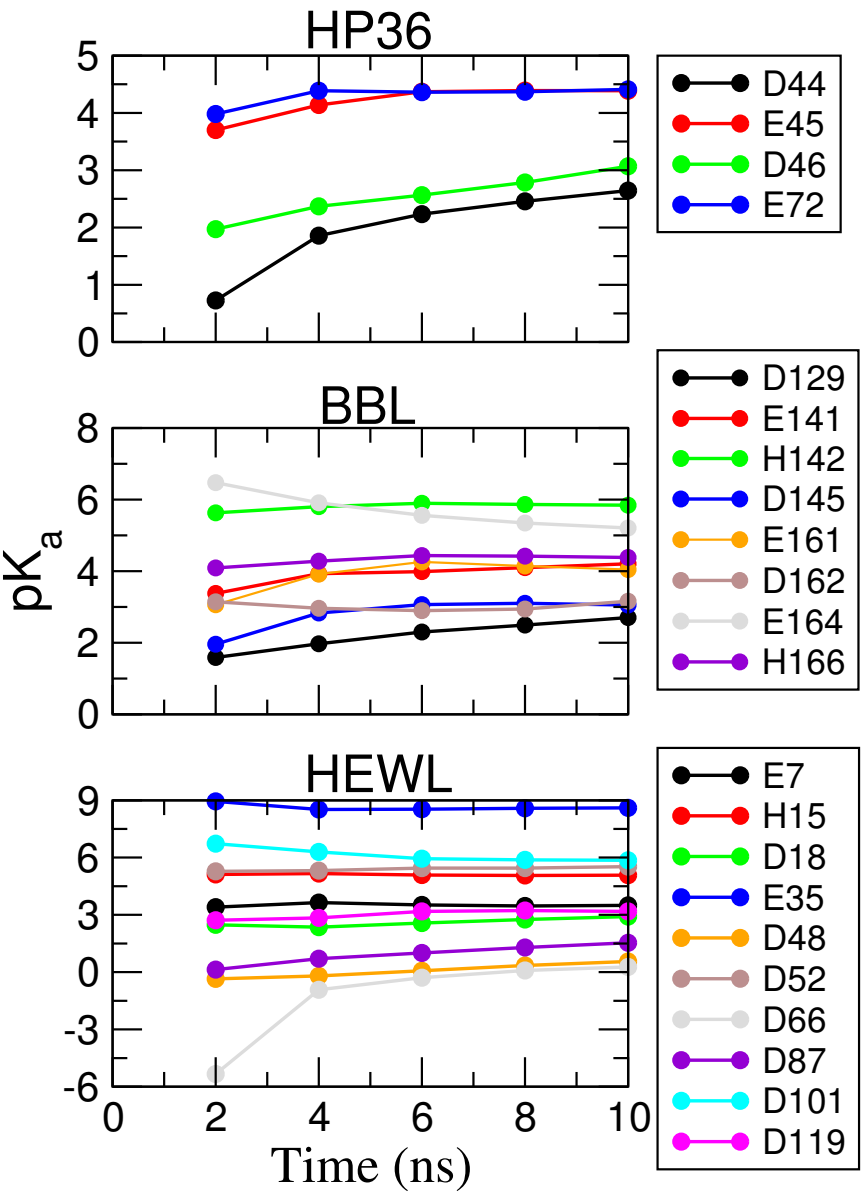

Figure S3: The pK_{a}'s of titratable residues of proteins calculated cumulatively every 2 ns.

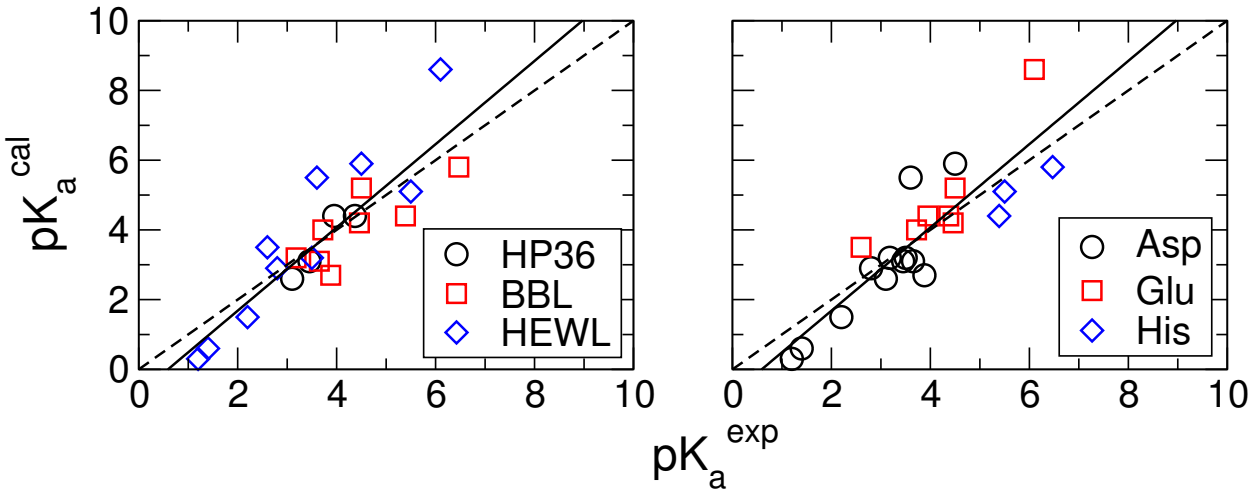

Figure S4: Comparison between calculated and experimental p$K_{\rm a}$'s for different proteins (*left*) and residues (*right*). Calculated p$K_{\rm a}$'s are from the entire 10-ns simulations. The solid line represents the linear regression to the data (slope=1.20 and R^2=0.76).

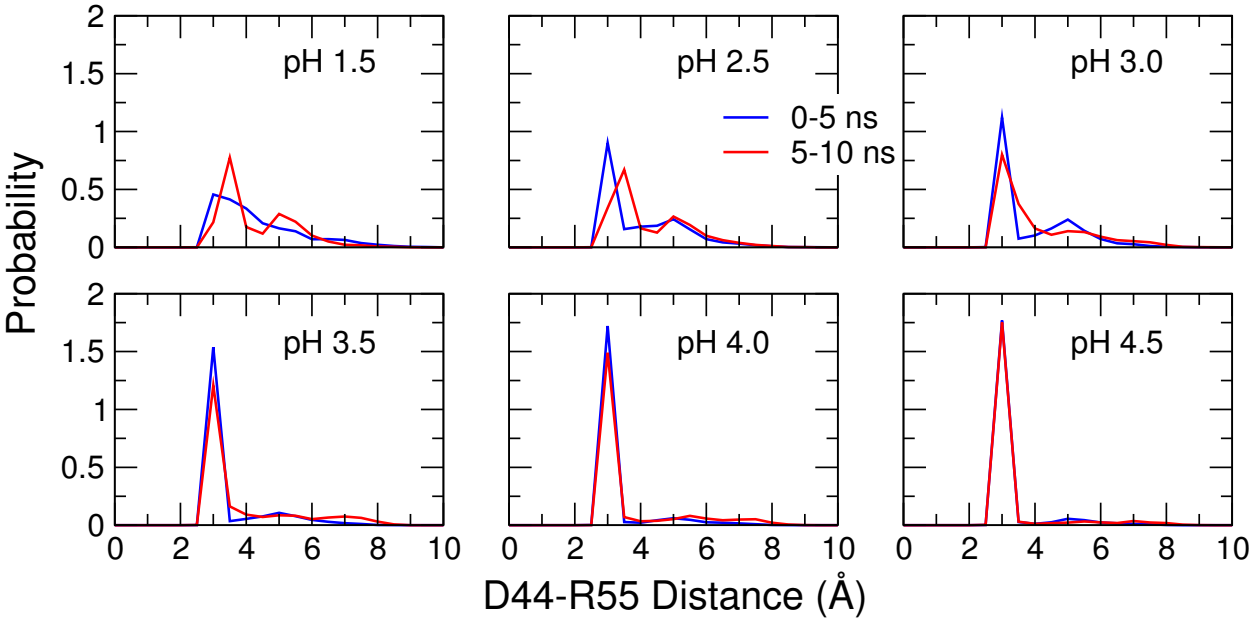

Figure S5: Probability distribution of the minimal distance between the carboxyl oxygen atoms of Asp44 and the guanidino nitrogen atoms of Arg55 in HP36. The first and second 5-ns data are shown in blue and red, respectively.

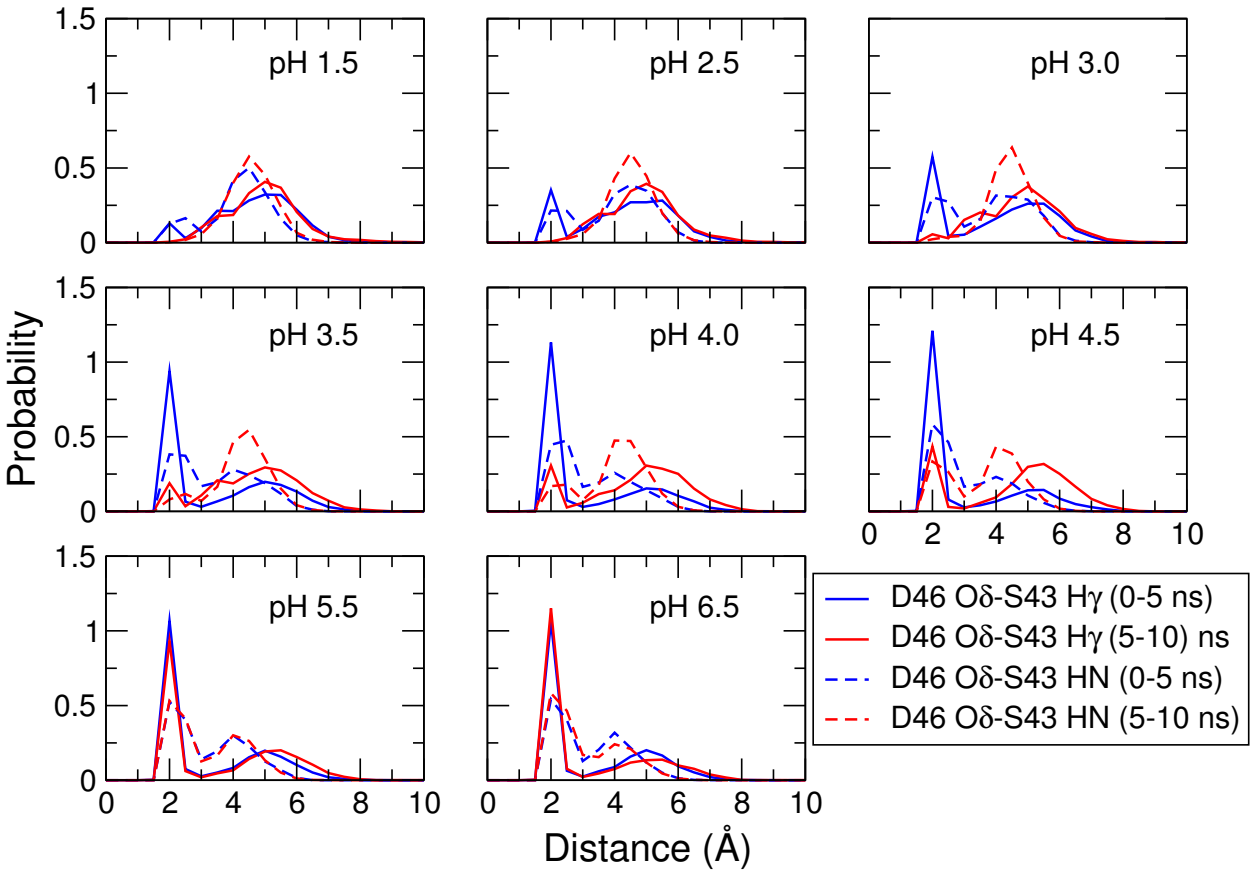

Figure S6: Probability distribution of the minimal distance from the carboxyl oxygen atom of Asp44 to the hydroxyl hydrogen atom (solid lines) and amide hydrogen (dashed lines) of Ser43 in HP36. The first and second 5-ns data are shown in blue and red, respectively.

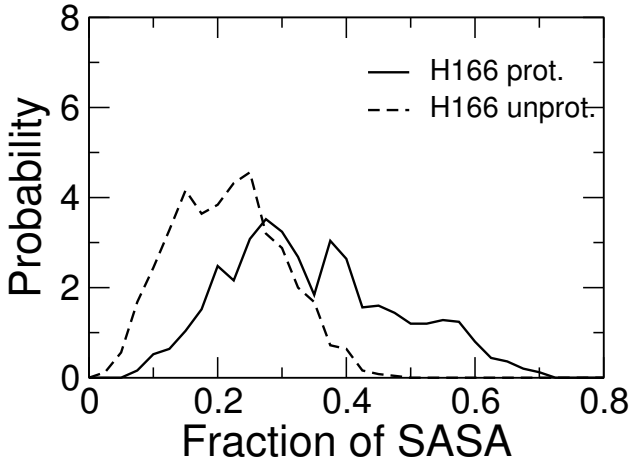

Figure S7: Probability distribution of the fractional SASA for the His166 side chain of BBL when His166 is fully protonated (pH 0.5) and unprotonated (pH 8). The fully solvent-exposed surface area (SASA) for His is 179 $Å^2$.

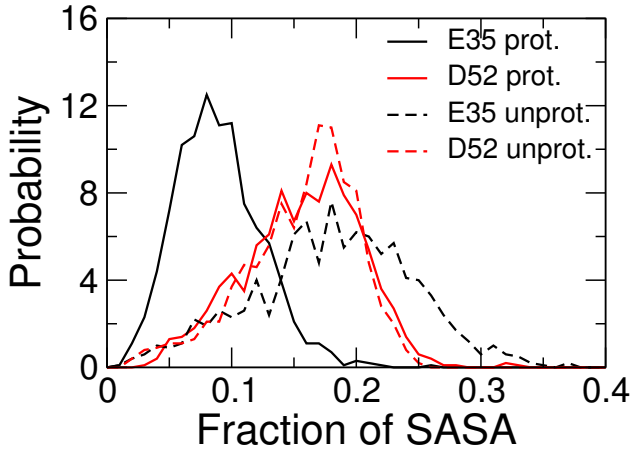

Figure S8: Probability distribution of the fractional SASA of the side chains of Glu35 and Asp52 in HEWL when they are fully protonated (pH 1) and unprotonated (pH 9.5). The fully solvent-exposed surface area (SASA) for Glu is 175 Å^2 and it is 145 Å^2 for Asp.

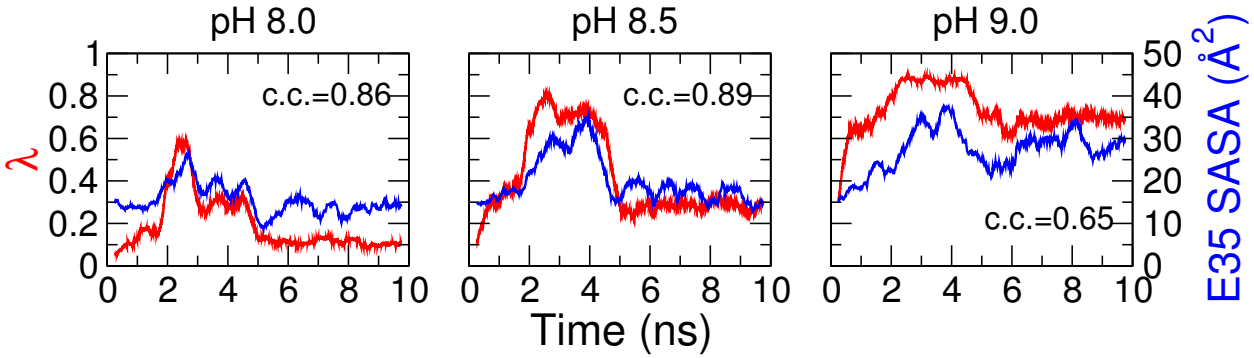

Figure S9: Correlation between the value of titration coordinate λ and the SASA of the side-chain of Glu35 of HEWL. Curves are running averages over 500-ps windows. Correlation coefficients are shown.

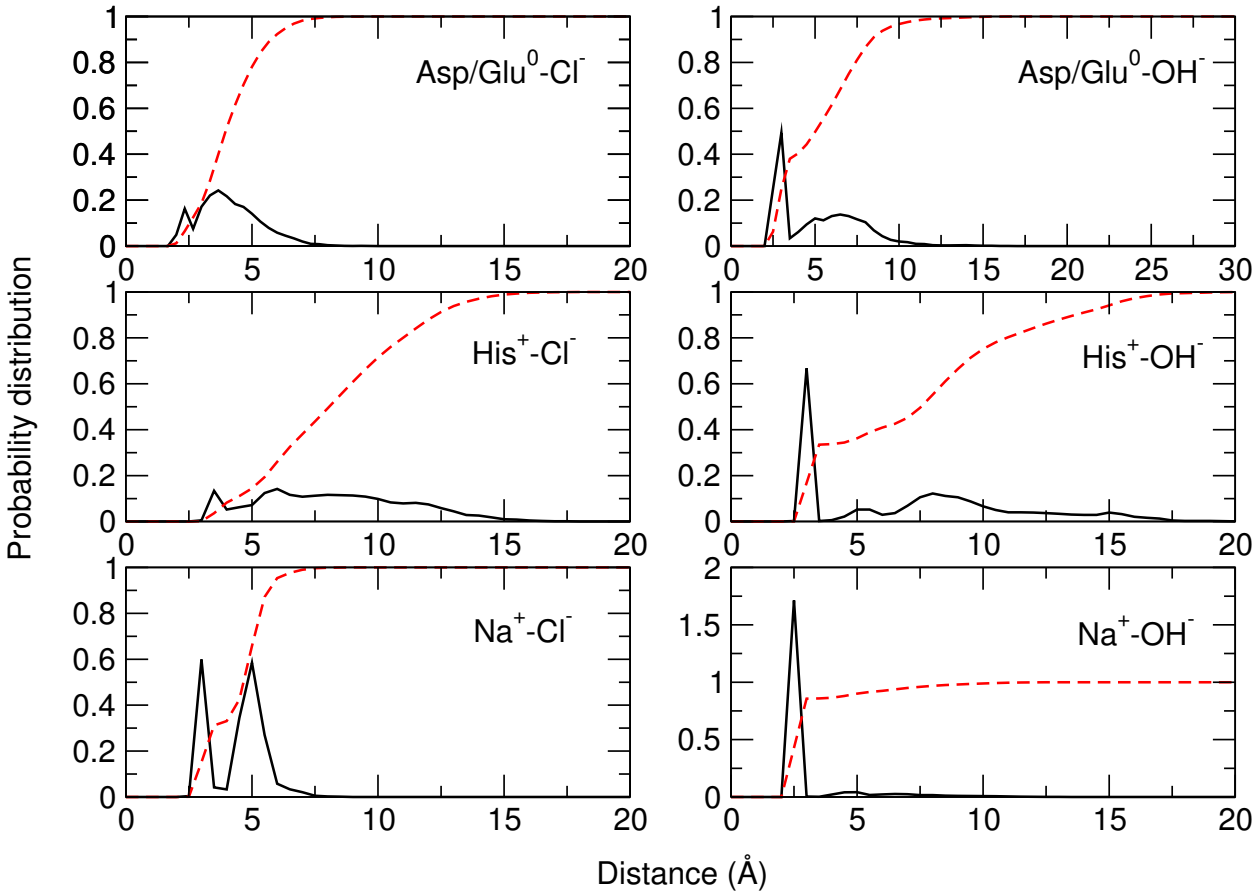

Figure S10: Interactions between the hydroxide and titratable sites/ions in BBL. Interactions in HP36 and HEWL are very similar. Left panel. Probability distribution of the minimum distance from the chloride ions to the carboxylic oxygens of Asp/Glu (top) or sodium ions (bottom). Right panel. Probability distribution of the minimum distance from the hydroxide to the carboxylic oxygens of Asp/Glu (top) or sodium ions (bottom). Dashed red curves are the integrated form (probability). We define the relative occupancy of the hydroxide as the probability when the minimum distance is within 5 Å. While the positions of the peaks are identical for the hydroxide and chloride ions, the major peak is much higher for the former.

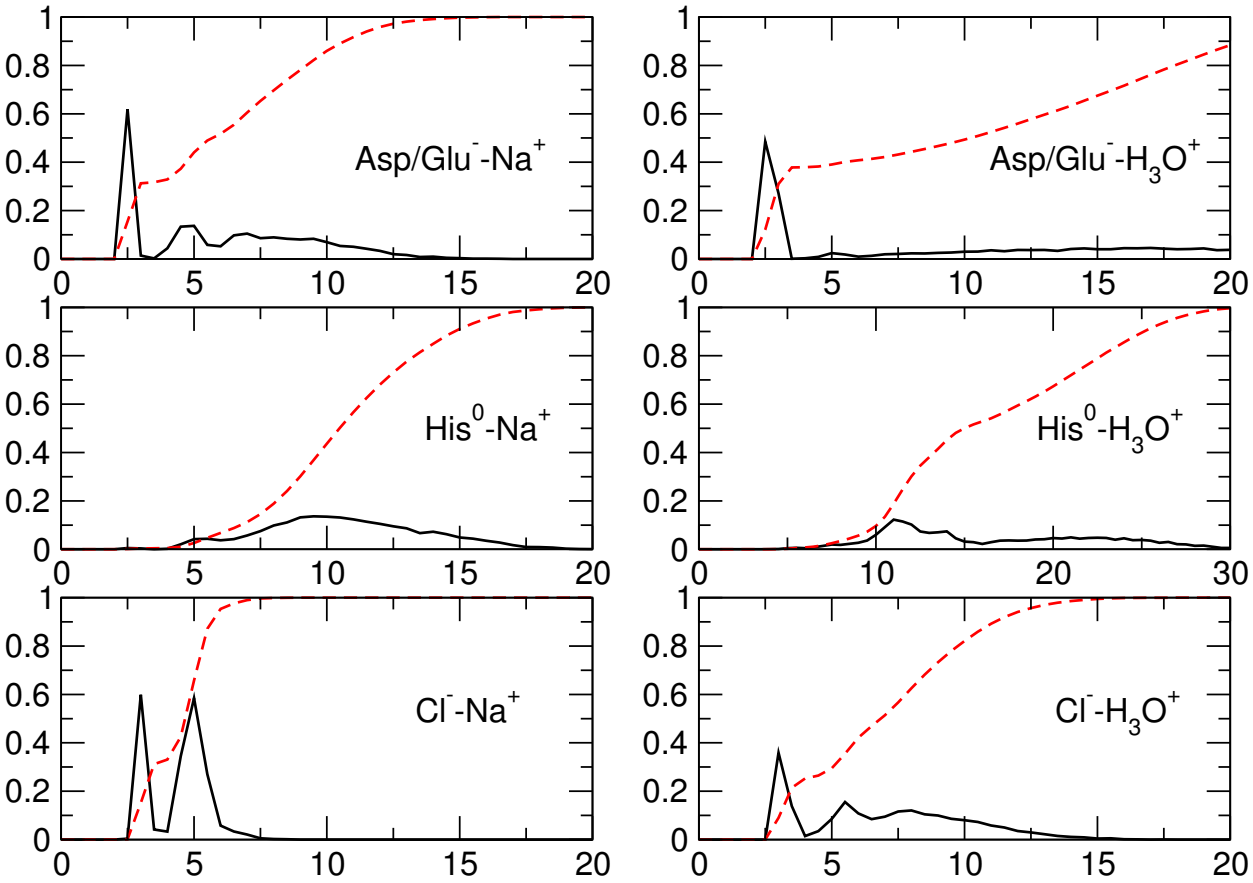

Figure S11: Interactions between the hydronium and titratable sites/ions in BBL. Interactions in HP36 and HEWL are very similar. Left panel. Probability distribution of the minimum distance from the sodium ions to the carboxylic oxygens of Asp/Glu (top) or sodium ions (bottom). Right panel. Probability distribution of the minimum distance from the hydronium to the carboxylic oxygens of Asp/Glu (top) or sodium ions (bottom). Dashed red curves are the integrated form (probability). We define the relative occupancy of the hydroxide as the probability when the minimum distance is within 5 Å.

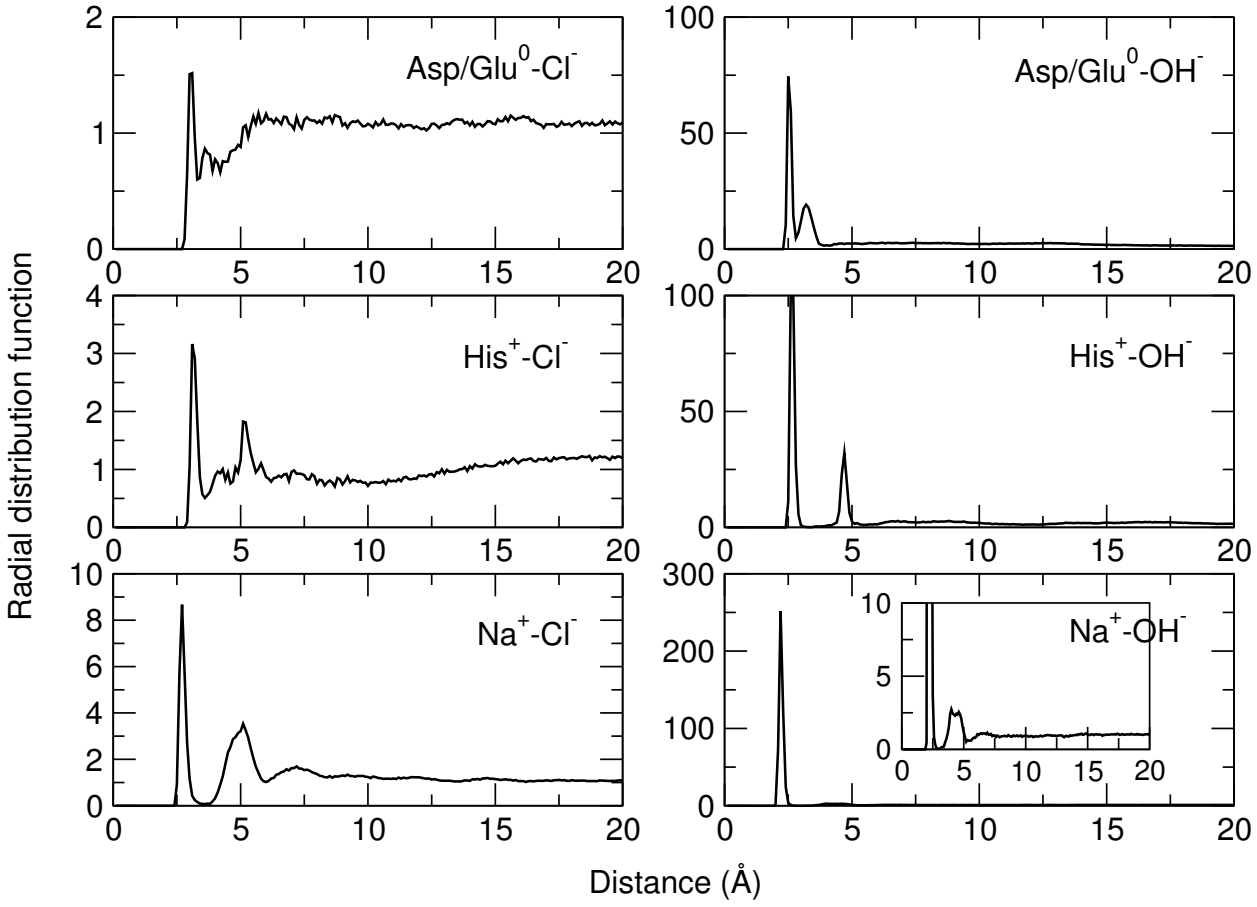

Figure S12: Interactions between hydroxide ions and titratable sites/ions in BBL. Left panel. Radial distribution function (RDF) from the chloride ions to the carboxylic oxygens of Asp/Glu (top) or sodium ions (bottom). Right panel. RDF from the hydroxide to the carboxylic oxygens of Asp/Glu (top) or sodium ions (bottom). RDFs for HP36 and HEWL are very similar. Note a second peak in the sodium-chloride RDF, which is due to the water-bridged interaction. We note that the RDF of Na^+–Cl^- is very similar to that obtained by Luo and Roux using the same force field (14).

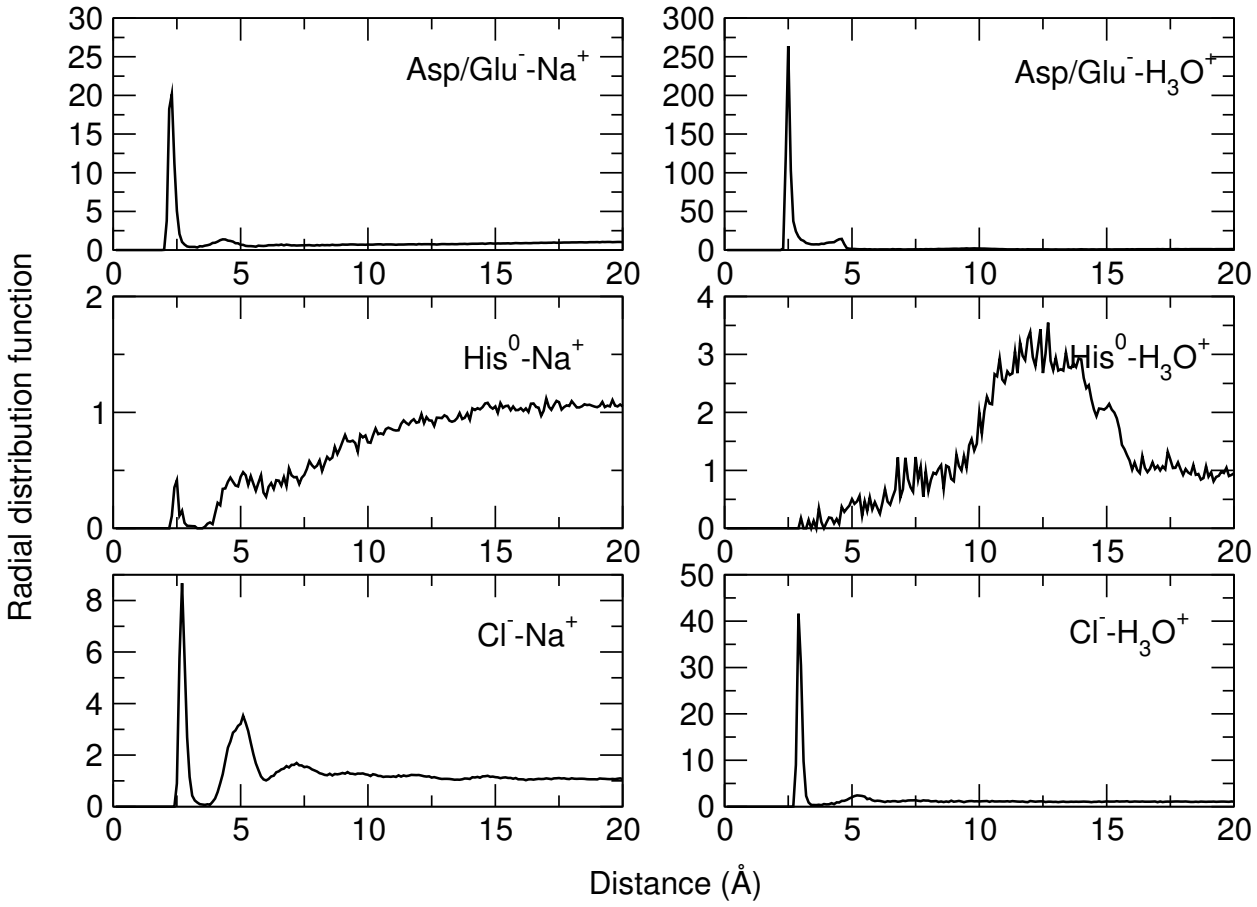

Figure S13: Interactions between hydronium and titratable sites/ions in BBL. Left panel. Radial distribution function (RDF) from chlorides to the carboxylic oxygens of Asp/Glu (top) or sodium ions (bottom). Right panel. RDF from the hydronium to the carboxylic oxygens of Asp/Glu (top) or sodium ions (bottom). RDFs for HP36 and HEWL are very similar. Note a second peak in the sodium-chloride RDF, which is due to the water-bridged interaction. The RDF of acid-Na^+ has two peaks: a major peak at 2.3 Å and a minor peak at 4.3 Å, which is in agreement with the RDF for the formate–Na^+ interaction (19).

References

1. Wallace, J. A., and J. K. Shen, 2012. Charge-leveling and proper treatment of long-range electrostatics in all-atom molecular dynamics at constant pH. *J. Chem. Phys.* 137:184105.

2. Bi, Y., 2008. Studies of the folding and stability of the villin headpiece subdomain. Ph.D. thesis, Stony Brook University.

3. Arbely, E., T. J. Rutherford, T. D. Sharpe, N. Ferguson, and A. R. Fersht, 2009. Downhill versus barrier-limited folding of BBL 1: energetic and structural pertubation effects upon protonation of a histidine of unusually low pKa. *J. Mol. Biol.* 387:986–992.

4. Arbely, E., T. J. Rutherford, H. Neuweiler, T. D. Sharpe, N. Ferguson, and A. R. Fersht, 2010. Carboxyl pK_a values and acid denaturation of BBL. *J. Mol. Biol.* 403:313–327.

5. Webb, H., B. M. Tynan-Connolly, G. M. Lee, D. Farrell, F. O'Meara, C. R. Søndergaard, K. Teilum, C. Hewage, L. P. McIntosh, and J. E. Nielsen, 2011. Remeasuring HEWL pK_a values by NMR spectroscopy: methods, analysis, accuracy, and implications for theoretical pK_a calculations. *Proteins* 79:685–702.

6. Brooks, B. R., C. L. Brooks III, A. D. Mackerell Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartles, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. K. T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, 2009. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* 30:1545–1614.

7. Wallace, J. A., and J. K. Shen, 2011. Continuous constant pH molecular dynamics in explicit solvent with pH-based replica exchange. *J. Chem. Theory Comput.* 7:2617–2629.

8. Mackerell, Jr., A. D., M. Feig, and C. L. Brooks, III, 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 simulations. *J. Comput. Chem.* 25:1400–1415.

9. Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

10. Vorobyov, I. V., V. M. Anisimov, and A. D. M. Jr., 2005. Polarizable Empirical Force Field for Alkanes Based on the Classical Drude Oscillator Model. *J. Phys. Chem. B* 109:18988–18999.

11. MacKerell Jr., A. D., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus, 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* 102:3586–3616.

12. Panahi, A., and M. Feig, 1996. Structure and Dynamics of Hydronium in the Ion Channel Gramicidin A. *Biophys. J.* 70:2043–2051.

13. Noskov, S. Y., and B. Roux, 2008. Control of ion selectivity in LeuT: two Na^+ binding sites with two different mechanisms. *J. Mol. Biol.* 377:804–818.

14. Luo, Y., and B. Roux, 2010. Simulation of osmotic pressure in concentrated aqueous salt solutions. *J. Phys. Chem. Lett.* 1:183–189.

15. Hoover, W. G., 1985. Canonical dynamics: Equilibration phase-space distributions. *Phys. Rev. A* 31:1695–1697.

16. Feller, S. E., Y. Zhang, R. W. Pastor, and B. R. Brooks, 1995. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* 103:4613–4621.

17. Nozaki, Y., and C. Tanford, 1967. Examination of titration behavior. *Methods Enzymol.* 11:715–734.

18. Bashford, D., D. A. Case, C. Dalvit, L. Tennant, and P. E. Wright, 1993. Electrostatic calculations of side-chain pK_a a values in myoglobin and comparison with NMR Data for histidines? *Biochemistry* 32:8045–8056.

19. Jagoda-Cwiklik, B., R. Vacha, M. Lund, M. Srebro, and P. Jungwirth, 2007. Ion pairing as a possible clue for discriminating between sodium and potassium in biological and other complex environments. *J. Phys. Chem. B* 111:14077–14079.