Supporting Information for:

Infinitely Dilute Partial Molar Properties of Proteins from Computer Simulation

*Elizabeth A. Ploetz and Paul E. Smith**

Kansas State University

Department of Chemistry

213 CBC Building

Manhattan, KS 66506-0401, USA

Additional details of the simulation methods

BPTI and HEW lysozyme

All BPTI and HEW lysozyme simulations were performed using the GROMACS 4.0.5 simulation package.[1] Each system was minimized for over 1,000 steps using the steepest descent algorithm and then equilibrated for 100 ps at p = 1 bar and T = 100 K, 100 ps at 1 bar and 200 K, and 100 ps at 1 bar and 300 K, with position restraints on all the heavy atoms during each equilibration. 25 ns of production simulation (100 ns for pure water) without position restraints was performed for each system at 300 K and 1 bar in the isothermal-isobaric (*NpT*) ensemble. In an effort to ensure that the simulations would sample from the *NpT* ensemble and the system fluctuations would be correct, the Nosé-Hoover (chain length of one) and Parrinello-Rahman T and p baths were used, with relaxation times of 0.5 ps and 2.5 ps, respectively, as advised by the Gromacs 4.5 manual.[2-4] A 4.5×10^{-5} bar^{-1} compressibility was used for all systems. Periodic boundary conditions and the minimum image convention were employed. All bond lengths were constrained using the Settle[5] and LINear Constraint Solver (LINCS)[6] algorithms for water and non-water molecules, respectively. The use of bond constraints allowed for a two fs time step to be used for the integration of the equations of motion, which was performed using the Leap-Frog[7] algorithm. The particle-mesh Ewald technique was used to calculate electrostatic interactions with cutoff distances of 1.0 nm and 1.5 nm for the real space electrostatic and van der Waals interactions, respectively, a convergence parameter of 3.123 nm^{-1}, cubic interpolation, a maximum fast Fourier transform grid spacing of 0.12 nm for the reciprocal space sum, and tinfoil boundary conditions.[8] The grid-based neighbor list was updated every ten steps.

Trp-cage

We simulated two conformations of trp-cage using KBFF. We chose the first NMR structure in PDB ID 2jof, a stabilized mutant of the original sequence, as our initial structure for the native conformations. The starting denatured conformation for the KBFF simulation was selected from a 100 ns canonical (*NVT*) simulation at 500 K ran using the AMBER99sb force field, which had been previously generated for a separate project.[9] The conformations were clustered using the Gromacs program g_cluster, specifically, the algorithm of Daura *et al.*[10] The initial structure was taken from a cluster that was nearly fully extended and had little secondary structure.

The denatured trp-cage conformations were equilibrated for 100 ps at 300 K followed by 25 ns of production. The native conformations were equilibrated for 100 ps at p = 1 bar and T = 100 K, 100 ps at 1 bar and 200 K, and 5 ns at 1 bar and 300 K, with position restraints on all the heavy atoms during each equilibration. They were then equilibrated for 5 ns at 300 K and 1 bar with position restraints on the α-carbons and 5 ns at 300 K and 1 bar without position restraints, followed by 25 ns of production simulation without position restraints.

For the AMBER99sb + TIP3P trp-cage simulations, Gromacs version 4.5.3 was used,[9] the particle-mesh Ewald cutoff distance was 1.2 nm for both the real space electrostatic and van der Waals interactions and the convergence parameter was 2.603 nm^{-1}. The pure TIP3P production simulation was 100 ns. All other details were the same as those for the KBFF + SPC/E simulations.

SUPPORTING INFORMATION REFERENCES

1. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435-447.
2. Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845-854.
3. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695-1697.
4. Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182-7190.
5. Miyamoto, S.; Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992**, *13*, 952-962.
6. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463-1472.
7. Hockney, R. W.; Goel, S. P.; Eastwood, J. W. Quiet high resolution computer models of a plasma. *J. Comput. Phys.* **1974**, *14*, 148-158.
8. Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N.log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089-10092.
9. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712-725.
10. Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Peptide folding: When simulation meets experiment. *Angew. Chem. Int. Ed.* **1999**, *38*, 236-240.