On the use of interaction energies in QM/MM free energy simulations: Supporting Information

Phillip S. Hudson,[†,‡] H. Lee Woodcock,[*,†] and Stefan Boresch[*,¶]

†*Department of Chemistry, University of South Florida, 4202 E. Fowler Ave., CHE205, Tampa, FL 33620-5250, United States*

‡*Laboratory of Computational Biology, National Institutes of Health, National Heart, Lung and Blood Institute, 12 South Dr. Rm 3053, Bethesda, MD 20892-5690*

¶*University of Vienna, Faculty of Chemistry, Department of Computational Biological Chemistry, Währingerstraße 17, A-1090 Vienna, Austria*

E-mail: hlw@usf.edu; stefan@mdy.univie.ac.at

1 Simulation Details

Free energy differences between force fields. The Langevin dynamics simulations in the gas phase employed a timestep of 1 fs. Equilibration consisted of 1 million steps (1 ns), followed by a production simulation of 100 million steps (100 ns). 100,000 configurations were saved, i.e., 1 ps interval. The solute molecules were fully flexible. Temperature was maintained by Langevin dynamics around 300K, and a friction coefficient of 5 ps^{-1} was applied to all atoms. Neither cut-offs nor tapering functions were applied when computing non-bonded interactions.

Simulations in aqueous solution also employed a timestep of 1 fs and consisted of 1 million steps equilibration (1 ns), followed by a production simulation of 100 million steps (100 ns). The solutes (Ala, Ser) were solvated in a cubic box of 986 TIP3 waters. Simulations were carried out at fixed volume (side length $L = 30.915$ Å); this value was the average box size observed during initial constant pressure equilibration runs for solvated Ala and Ser. Temperature was maintained near 300 K by a Nosé-Hoover thermostat. In all simulations, Lennard-Jones interactions were truncated beyond 12Å, and force switching[1] was applied between 10–12Å. Electrostatic interactions were computed by the particle mesh Ewald method[2,3] with a real space cut-off radius of 12Å and $\kappa = 0.34$Å^{-1}; a 32×32×32 grid was used for the fast Fourier transform calculations. The solute was fully flexible, while SHAKE[4] was used to maintain rigid water geometries. As in the gas phase 100,000 coordinate sets were saved for post-processing.

Solvation free energy of *bis*-2-chloro diethyl ether (2CLE). For ease of reading, we reproduce Fig. 3 of the main manuscript below:

$$\begin{tikzcd}
(\mathbf{S}+W)_{aq}^{DFTB3/MM} \arrow[rrr, dotted] & & & \mathbf{S}_{gasp}^{DFTB3} \\
(\mathbf{S}+W)_{aq}^{MM} \arrow[u, "(i)"'] \arrow[r, "(ii)"] & (\mathbf{S}_{\text{-}chg}+W)_{aq}^{MM} \arrow[r, "(iii)"] & (\mathbf{S}_{\text{-}nbd}+W)_{aq}^{MM} \arrow[r, "(iv)"] & \mathbf{S}_{gasp}^{MM} \arrow[u, "(v)"']
\end{tikzcd}$$

Figure 1: Detailed thermodynamic cycle to compute the solvation free energy of 2CLE, including an overview of the purely classical steps. **S** denotes the solute, W water. The figure is identical to Fig. 3 of the main manuscript.

Gas phase simulations were carried out using the two force fields (C36, GAAMP), as well as the DFTB3 level of theory. In all gas phase simulations, neither nonbonded cut-offs nor switching or shifting functions were applied; the timestep was 0.5 fs. Average temperature was kept close to 300K by means of Langevin dynamics with a friction coefficient of 5 ps^{-1} applied to all atoms. The force field simulations were 400 ns long, during the course of which 40,000 coordinate frames were saved. In addition, to compute the free energy correction (iv) of Fig. 1 an additional, analogous simulation with all nonbonded interactions turned off was carried out. BAR based on the simulations with nonbonded interactions turned off and on was used to compute the free energy correction (iv) in a single step. We also carried out

Langevin dynamics simulations (friction coefficient 5 ps^{-1}, timestep 0.5 fs) at the DFTB3 level of theory. Ten simulations, consisting of 500 ps equilibration and 10 ns production were carried out, starting from different initial velocities. Restart files were saved every 10 ps, resulting in a total of 10,000 coordinate/velocity sets.

2CLE was solvated in a box of length 30.9120 Å containing 989 TIP3 waters. SHAKE was applied to the waters; the solute was flexible. Lennard-Jones interactions were switched off between 10 to 12 Å. Electrostatic interactions were computed by the particle mesh Ewald method[2,3] with a real space cut-off radius of 12Å and $\kappa = 0.34\text{Å}^{-1}$; a 32×32×32 grid was used for the fast Fourier transform calculations. We carried out NVT MD simulations using a Nosé-Hoover thermostat with the solute described with the CGenFF and GAAMP force fields, as well as using a SQM/MM setup in which waters were treated classically and the solute was described by DFTB3. The timestep in the purely classical simulations was 1 fs; 300 ps of equilibration were followed by production simulations of 20 ns, during which 20,000 restart files were saved as the starting point for non-equilibrium work switching simulations. Each simulation was repeated twice, starting from different random velocities. For the SQM/MM case, we carried out five 4 ns simulations starting from different random velocities. The timestep of each simulation was 0.5 fs. An additional equilibration period of 100 ps was discarded, then restart files were saved every 0.5 ps (i.e., 40,000 restart files in total).

The calculation of the free energy difference for turning off solute charges (step (*ii*) in Fig. 1) employed 5 alchemical states (λ = 0.00, 0.25, 0.50, 0.75, 1.00) where $U(\lambda) = U^{MM} - \lambda U^{MM}_{elec}$. Each λ state was simulated for 50 ns at constant volume and temperature, with a timestep of 0.5 fs, and pre-equilibrated for 500 ps. Coordinates were saved to disk every 1 ps during production phase (i.e., 50,000 snapshots) and BAR was used to compute the free energy differences for each of the λ intervals. The Lennard-Jones interactions of the solute were turned off over 6 alchemical states (λ=0.0, 0.2, 0.4, 0.6, 0.8, 1.0) where $U(\lambda) = U^{MM} - U^{MM}_{elec} - \lambda U^{MM}_{LJ}$ (step (iii) of Fig. 1). The first λ state (λ=0.0) is directly equivalent to the λ=1.0 state of the discharging step, and thus did not need to be explicitly run. For λ=0.2 to λ=0.6, MD simulations were performed, whereas for λ=0.8 and λ=1.0, Langevin dynamics (LD) simulations with a friction coefficient of 5 ps^{-1} (as the solute began to take on more gas phase like behavior). Simulations at λ = 0.2–1.0 as well as all post-processing steps (recalculations of energies needed for BAR) were carried out with the PERT module of CHARMM. Van der Waals endpoint problems were avoided by the PERT/PSSP soft-core potential.[5] Each simulation between λ=0.2 and λ=1.0 was initiated with 5 random initial velocities, performed for 10 ns with a timestep of 1 fs, and preceded by a thermal equilibration of 200 ps. Coordinate snapshots were saved every 1 ps, and BAR was used to compute the free energy differences between the individual λ-intervals.

The calculation of the free energy correction steps (i) and (v) is described in the main manuscript.

References

(1) Steinbach, P. J.; Brooks, B. R. New spherical-cutoff methods for long-range forces in macromolecular simulation. *J. Comput. Chem.* **1994**, *15*, 667–683.

(2) 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.

(3) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(4) Kräutler, V.; van Gunsteren, W. F.; Hünenberger, P. H. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J. Comput. Chem.* **2001**, *22*, 501–508.

(5) Brooks, B. R.; Brooks III, C. L.; MacKerell Jr., A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.;

Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545.