Accurate and Efficient Corrections for Missing Dispersion Interactions in Molecular Simulations

Michael R. Shirts*,†

Department of Chemistry, Columbia University, New York, New York 10027

David L. Mobley†

Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143

John D. Chodera and Vijay S. Pande

Department of Chemistry, Stanford University, California 94305

Received: May 10, 2007; In Final Form: August 22, 2007

In simulations, molecular dispersion interactions are frequently neglected beyond a cutoff of around 1 nm. In some cases, analytical corrections appropriate for isotropic systems are applied to the pressure and/or the potential energy. Here, we show that in systems containing macromolecules, either of these approaches introduce statistically significant errors in some observed properties; for example, the choice of cutoff can affect computed free energies of ligand binding to proteins by 1 to 2 kcal/mol. We review current methods for eliminating this cutoff-dependent behavior of the dispersion energy and identify some situations where they fail. We introduce two new formalisms, appropriate for binding free energy calculations, which overcome these failings, requiring minimal computational effort beyond the time required to run the original simulation. When these cutoff approximations are applied, which can be done after all simulations are completed, results are consistent across simulations run with different cutoffs. In many situations, simulations can be run with even shorter cutoffs than typically used, resulting in increased computational efficiency.

I. Introduction

The simulation of molecular systems is computationally intensive, and significant effort has gone into increasing computational efficiency in these simulations.[1−6] One common optimization in simulations with pairwise potentials is to divide the interactions into long and short range, which are treated differently. This can be done by evaluating long-range interactions via alternate summation methods like particle-mesh Ewald,[1] by evaluating long-range interactions less frequently,[2] or in the simplest case, simply by neglecting certain types of interactions beyond a given cutoff. By cutoffs, we mean any sort of truncation or tapering scheme by which the interaction energy is reduced to zero as a function of distance; it is well-known that abrupt cutoffs can lead to significant unphysical artifacts.[7,8]

Significant effort has gone especially into improving the efficiency of calculating long-range electrostatic interactions. It is well-established that simply truncating electrostatic interactions (either abruptly or with a tapered cutoff) can introduce significant inaccuracies.[4,9−11] Methods such as particle-mesh Ewald (PME),[1] reaction field (RF),[12] and particle−particle particle−mesh (P3M)[6] have been developed to increase the efficiency by calculating the potential energy and forces of an infinite periodic system. These approaches introduce fewer artifacts than even very long direct electrostatic cutoffs.[5,9]

In comparison, very little effort has gone into the treatment of dispersion interactions due to van der Waals forces, which are modeled in most force fields by the attractive part of the Lennard-Jones 12-6 potential. These interactions are much smaller in magnitude than electrostatic interactions beyond very short distances, since they decay as r^{-6}. Thus, they are usually not computed between atoms separated by more than some cutoff distance, typically between 0.8 and 1.2 nm. There are well-known, robust, and efficient corrections for the effect of these cutoffs on the potential energy and pressure for isotropic liquids.[13] These corrections are usually also used for nonisotropic systems, like solvated proteins, with the assumption that any deviations are negligible.

However, the validity of this isotropic assumption has not been thoroughly tested for many observables, and here, we show that it can introduce significant errors in important properties in several types of systems. For example, even when using such corrections, discrepancies in binding affinity between simulations run with different cutoffs can be up to 2 kcal/mol unless prohibitively expensive long-range cutoffs are used.[14] In this paper, we present efficient methods to correct the binding energies for cutoffs in the van der Waals term, which remain robust even down to cutoffs of 0.7 or 0.8 nm. Here, we provide a brief overview of how to apply corrections to general observables using reweighting methods but focus mainly on developing accurate corrections for binding free energy calculations.

II. Background: Analytical Dispersion Corrections in Isotropic Systems

A. Analytical Corrections for Truncation of Dispersion Interactions. Dispersion is a purely attractive interaction. Although the interaction energy between any two particles

* Corresponding author. E-mail: michael.shirts@columbia.edu
† These authors contributed equally to this work.

10.1021/jp0735987 CCC: $37.00
Published on Web 10/19/2007

diminishes as r^{-6}, a single particle interacts with all of the other particles in a system. Thus, the neglected portion of this interaction energy involves an integral over all space where $r > r_c$, and hence (with uniform particle density) diminishes only as r_c^{-3}. Standard cutoffs of 0.8−1.2 nm typically employed in molecular simulation protocols can therefore neglect a significant fraction of the total potential energy.

Since a major goal of simulation is to be able to provide quantitative insight and predictions when compared with experiment, treating cutoffs as an adjustable parameter is undesirable. A cutoff is an unphysical parameter, in that it is not something that can be defined experimentally. Using the same simulation parameters, two groups could obtain two different sets of predictions for exactly the same set of simulations by virtue of using two different sets of cutoffs. Which set of predictions is "correct"? One could argue that the cutoff parameters that were used to develop the parameters should be used. However, this is problematic if the cutoff method used in the original parametrization is not suitable for molecular simulations as, for example, with abrupt cutoffs.[7,8] Thus, we believe the best solution, short of parametrizing new force fields using extremely long cutoffs, is to make sure computed observables are independent of the choice of cutoff method or cutoff distance.

1. Isotropic Fluids with Homogeneous Lennard-Jones Sites. It is well-known how to correct for a radial cutoff of dispersion interactions in a homogeneous, isotropic fluid.[13] There, the radial and angular pair distribution functions, $g(r)$ and $g(\theta, \phi)$, are both assumed to be equal to unity outside the chosen cutoff distances. This is well-satisfied beyond 0.8 nm by a variety of common water models, as well as by real water.[15−17] However, this standard correction also assumes that all Lennard-Jones sites outside the cutoff are identical. This is true for solvent models which only have one Lennard-Jones site per molecule (such as SPC/E[18] and TIP*n*P[19]); however, some water models have additional Lennard-Jones sites centered on the hydrogens, and simulations of many other solvent types are frequently of interest. Alternate variants can be derived that remove the requirement for homogeneous particle types, as we discuss below.

For the next section, we will assume the use of a Lennard-Jones 6-12 potential with an abrupt cutoff to compute the dispersive part of the energy, unless noted otherwise. However, most of the equations are trivially generalizable to other models of long-range interactions, such as tapered cutoffs,[17,20] which provide greatly improved energy conservation. In most cases, the r^{-12} volume exclusion term will indeed be negligible outside of any commonly used cutoff distance, but the derivations presented will include this term as well, as at shorter distances it can become non-negligible. For clarity of discussion, here, however, we will refer to the neglected component of the energy as dispersive, as the large majority of the neglected energy is from the dispersion term.

With these assumptions, the corrections to the energy and the pressure can then be computed as:[13,17]

$$\begin{aligned} U_{\text{full}} &= U_c + U_{LRC} \\ &= U_c + \frac{N}{2}\rho \int_{r_c}^{\infty} U(r) g(r) 4\pi r^2 \, \mathrm{d}r \\ &= U_c + 2\pi N \rho \int_{r_c}^{\infty} U(r) r^2 \, \mathrm{d}r \\ &= U_c + 2\pi N \rho \int_{r_c}^{\infty} 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] r^2 \, \mathrm{d}r \\ &= U_c + 8N\pi\rho\epsilon\sigma^3 \left[\frac{1}{9}\left(\frac{\sigma}{r_c}\right)^9 - \frac{1}{3}\left(\frac{\sigma}{r_c}\right)^3 \right] \end{aligned} \quad (1)$$

$$\begin{aligned} (PV)_{\text{full}} &= (PV)_c + (PV)_{LRC} \\ &= (PV)_c + \frac{N}{2}\rho \int_{r_c}^{\infty} -\frac{1}{3} r \frac{\partial U(r)}{\partial r} g(r) 4\pi r^2 \, \mathrm{d}r \\ &= (PV)_c - \frac{2\pi N}{3}\rho \int_{r_c}^{\infty} r \frac{\partial U(r)}{\partial r} r^2 \, \mathrm{d}r \\ &= (PV)_c - \frac{2\pi N}{3}\rho \int_{r_c}^{\infty} 4\epsilon \left[-12\left(\frac{\sigma}{r}\right)^{12} + 6\left(\frac{\sigma}{r}\right)^{6} \right] r^2 \, \mathrm{d}r \\ &= PV_c - 16N\pi\rho\epsilon\sigma^3 \left[\frac{2}{9}\left(\frac{\sigma}{r_c}\right)^9 - \frac{1}{3}\left(\frac{\sigma}{r_c}\right)^3 \right] \end{aligned} \quad (2)$$

where U_{full} and $(PV)_{\text{full}}$ are the potential energy and instantaneous pressure-volume for a liquid with a non-truncated potential, U_c and $(PV)_c$ are the values determined using the cutoff for the potential, ρ is the average density of the pure solvent, N is the number of Lennard-Jones interaction sites in the entire system, and ϵ and σ are the Lennard-Jones well depth and characteristic radius and the subscript LRC indicates a long-range correction. The $N/2$ factor in eqs 1 and 2 is to correct for double counting of pairwise interactions, and the $r\,\mathrm{d}U/\mathrm{d}r$ term comes from the virial theorem and the definition of the virial. The part of the correction due to the r^{-6} dispersion term will dominate this correction term in most cases. For example, if $r_c = 3\sigma$, the repulsive r^{-12} term will contribute only 0.1% of the total correction energy. For Lennard-Jones van der Waals interactions, this energy will scale with ϵ. It will not necessarily scale with the number van der Waals sites per molecule—rather, it scales linearly with the number density of sites per unit volume in the system.

With this formula, we can estimate the amount of dispersion energy neglected with a given cutoff. For example, consider TIP3P water, for which $\epsilon = 0.1521$ kcal/mol and $\sigma = 0.315061$ nm, and assume a density of 1.0 g/cm^3.[19] Define the energetic contribution from the integral from the single particle pair potential minimum radius to infinity, divided by N, as the dispersion energy per molecule. This definition will somewhat underestimate the true dispersion energy, as $g(r) > 1$ near the minimum but provides the correct scale. The per molecule dispersion energy for TIP3P is then −1.9 kcal/mol. A cutoff of 0.9 nm neglects 0.11 kcal/mol or 6% of the dispersion energy. Although relatively small, this can result in a noticeable change in system observables. The heat of vaporization for most water models is about 10 kcal/mol, while the hydration free energy is about 6 kcal/mol. A value of 0.11 kcal/mol represents about 2% error in the free energy, which may be insufficient for high precision calculations.

2. Consequences for Pressures and Densities. If simulating a canonical ensemble, the volume (and thus the density) will not change during the simulation, and the dispersion correction can easily be applied afterwards. However, if the volume is allowed to change during the simulation, this dispersion correction to the virial must be included at each time step or the resulting density will be incorrect.

Take the case of a box of 900 TIP3P molecules, simulated for 5.0 ns (other simulation details described in section IV), with results shown in Table 1 for a NPT simulation at 298 K and 1 atm. All densities reported here are in grams per cubic centimeter, and uncertainties are less than 2×10^{-4} g/cm^3 in all cases. With a Lennard-Jones cutoff tapered from 1.2 to 1.3 nm, the density is 0.986 with a correction and 0.983 without the correction. Thus, even with these long cutoffs, the difference is statistically significant. With a Lennard-Jones cutoff tapered from 0.8 to 0.9 nm, however, the density is 0.986 with a correction, and 0.977 without. Extending down to a Lennard-

TABLE 1: Density of TIP3P Water with Different Dispersion Cutoffs

cutoff[a] nm	no correction[b] g/cm^3	with correction[b] g/cm^3	difference %
1.2−1.3	0.9829	0.9855	0.3
1.0−1.1	0.9815	0.9856	0.4
0.8−0.9	0.9775	0.9860	0.9
0.7−0.8	0.9731	0.9859	1.3

[a] Dispersion cutoffs are tapered over 0.1 nm. [b] Densities are computed with and without the isotropic analytical correction to the pressure from eq 2. All uncertainties in density are 0.0002 g/cm^3. Further simulation details in Methods.

Jones cutoff of 0.7 to 0.8 nm (short enough that the $g(r) = 1$ assumption may begin to fail), the density is 0.973 without a correction and 0.985 with the correction. This is a difference of more than 1% between cutoff values without the correction, while with the correction, the density values are statistically indistinguishable across different cutoffs. Assuming an approximate experimental value of $B = 4.5 \times 10^{-5}$ atm^{-1} for the isothermal compressibility of water and approximating $\delta P = -B^{-1}(\delta V)/V$, we find the difference in density between even the 1.2 nm cutoffs and the corrected results is equivalent to a pressure difference of about 50 atm. Clearly, some sort of dispersion correction is necessary to obtain values that are independent of cutoff in dense fluids.

3. Isotropic Fluids with Heterogeneous Lennard-Jones Sites. In systems with more than one type of Lennard-Jones site, we can rewrite eq 1 as a sum over multiple particle types and compute "average" Lennard-Jones terms (where the brackets indicate averages over all particle interaction pairs, not ensemble averages):

$$\begin{aligned} E_{LRC} &= \frac{N\rho}{2}\frac{1}{N(N-1)}\sum_{i<j}^{n}\int_{r_c}^{\infty} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12}-\left(\frac{\sigma_{ij}}{r}\right)^{6}\right]g(r)4\pi r^2\,\mathrm{d}r \\ &= 2\pi N\rho\int_{r_c}^{\infty}(\langle 4\epsilon_{ij}\sigma_{ij}^{12}\rangle^{-12}-\langle 4\epsilon_{ij}\sigma_{ij}^{6}\rangle\, r^{-6})r^2\,\mathrm{d}r \\ &= 8N\pi\rho\left[\frac{1}{9}\langle\,\epsilon_{ij}\sigma_{ij}^{12}\rangle r_c^{-9}-\frac{1}{3}\langle\epsilon_{ij}\sigma_{ij}^{6}\rangle r_c^{-3}\right] \end{aligned} \quad (3)$$

If the radial and angular distribution functions of all Lennard-Jones particles are uniform outside of the cutoff (for example, if the system is a isotropic collection of molecules, each molecule containing several Lennard-Jones sites), then this formulation will be as exact as the homogeneous case. This treatment can be trivially extended to the virial term as well, to obtain:

$$(PV)_{LRC} = 16N\pi\rho\left[\frac{2}{9}\langle\epsilon_{ij}\sigma_{ij}^{12}\rangle r_c^{-9}-\frac{1}{3}\langle\epsilon_{ij}\sigma_{ij}^{6}\rangle r_c^{-3}\right] \quad (4)$$

This heterogeneous isotropic analytical correction should be applied to get correct densities when simulating a system such as a protein or other heteropolymer in solvent or with a mixed solvent system. We have not been able to find this version of the analytical correction in the published literature, but it has nevertheless found its way into at least current versions of biomolecular simulation codes AMBER[21] and GROMACS[22] and is likely in others, although this is often not well-documented. CHARMM uses an alternate method that appears to solve the problem of cutoff dependent densities by calculating the pressure with very long cutoffs at infrequent intervals and keeping the long-range part of the pressure constant between updates.[23]

4. Corrections to the Chemical Potential. The same formalism used to correct the energy and pressure for the effect of Lennard-Jones cutoffs can be used to correct the excess chemical potential of one particle in solution as well, with ϵ and σ taken as the solvent−solute Lennard-Jones constants in place of solvent−solvent Lennard-Jones constants. Returning to the pure solvent case and integrating over all particles interacting with the solute particle, we can compute this as:[13]

$$\begin{aligned} \mu_{\mathrm{full}} &= \mu_c + \mu_{LRC} \\ &= \mu_c + \rho\int_{r_c}^{\infty} U(r)4\pi r^2\,\mathrm{d}r \\ &= \mu_c + 16\pi\rho\epsilon_{ij}{\sigma_{ij}}^3\left[\frac{1}{9}\left(\frac{\sigma_{ij}}{r_c}\right)^9-\frac{1}{3}\left(\frac{\sigma_{ij}}{r_c}\right)^3\right] \end{aligned} \quad (5)$$

This correction can easily be rederived for different van der Waals formulations and cutoff schemes. This long-range correction to the chemical potential can be calculated either at each step or after the simulation has completed, as long as the average density is correct (for example, by use of the correction from eq 2 during the simulation), and the other terms in eq 5 will not change significantly during a simulation.

A similar expression for heterogeneous isotropic sites can be derived by analogy to eqs 3 and 4. In the case of the chemical potential, the averages of the *ij* parameters are over the pairs of solute−environment interactions, not all pairs of interactions in the system. However, this heterogeneous correction to the chemical potential can cause problems when applied to highly nonisotropic systems. If we are simulating the removal of a single ligand from complex with a solvated protein, then the value of the correction will depend on the size of the box, as the ratio of water to non-water van der Waals sites will change. Thus, the resulting chemical potential will also be box-size dependent, which is undesirable. Additionally, the bulk number density of Lennard-Jones sites in the solvent will be a poor approximation of the density of Lennard-Jones sites of a heterogeneous system such a solvated protein. One approach would be to use the actual density in the box, instead of the bulk solvent density, for ρ in eq 5. However, this number is also highly dependent on the ratio of solvent to protein on the system, which depends on the box size. Because of these difficulties, we do not recommend using a heterogeneous version of the analytical correction in eq 5 to compute the free energy of ligand binding.

However, using the homogeneous correction of eq 5 for hydration free energies is recommended. Comparison of free energies of hydration for small molecules computed over a wide range of cutoffs show that, with this correction, computed free energies are essentially independent of cutoff.[20] For large solutes and short cutoffs, the condition that $g(r) = 1$ might be violated slightly, but this contributes very little when the effective solute radius becomes noticeably larger than the cutoff distance.[14]

B. Failure of the Isotropic Assumption. The assumption that the system is isotropic outside the cutoff clearly is not satisfied in the case of simulations of heteropolymers, at least when the cutoff is smaller than the maximum dimension of the heteropolymer. Consider the case of computing the binding free energy of a ligand to a protein in solution. The ligand in complex is in a very heterogeneous environment. In a hydration free energy calculation, the end states are the solvated and vapor states; in binding, the end states are the bound state and the solvated state. If the contribution to the chemical potential from the dispersion energy neglected because of the cutoff were equal in both states, no correction would be necessary, as it would cancel. However, the correction terms are not equal. In

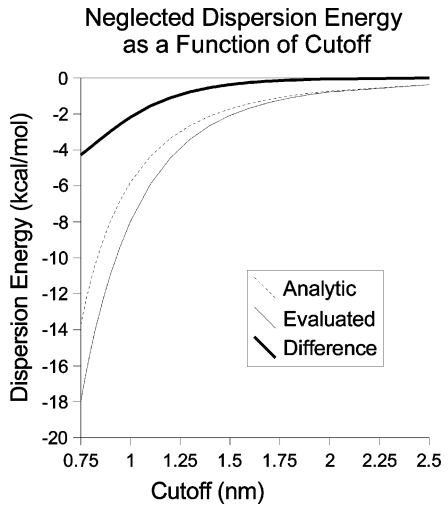

Figure 1. An approximate estimate of the neglected intermolecular dispersion potential energy and the analytic approximation to this energy, as a function of the cutoff, for FK506 bound to FKBP-12 simulated in TIP3P water. Also graphed is the difference between the two energies. The neglected dispersion energy is computed by averaging the evaluated energy over 400 configurations, with uncertainty less than 0.05 kcal/mol for all cutoffs. The difference is greater than 0.5 kcal/mol out to almost 1.5 nm, indicating that virtually all ligand binding energies calculated without a proper treatment of dispersion energies are not independent of simulation cutoff.

particular, in the complex, both the angular and the radial distribution functions are very far from uniform outside the cutoff. The ligand is fixed into position on or near the surface of the protein, meaning that the distributions now have substantial angular-dependence, whereas the distributions will be relatively flat around the solvated ligand. Additionally, the density of Lennard-Jones sites will be different around the solvated ligand relative to the bound ligand because of the difference between the densities of water and protein. Furthermore, the Lennard-Jones sites of water and the protein have substantially different dispersion parameters.

How large is this difference from the analytical correction? We take as an example the case of FK506 binding to FKBP-12 (simulation details discussed in Methods below). As a first order estimate of the problems with the analytical correction, we calculate the Lennard-Jones interaction energy between the solute and the environment in the region from 0.75 to 2.5 nm, averaged over a number of saved system configurations. Because the box is finite, we can only evaluate the Lennard-Jones energy out to the largest cutoff distance in the box, approximately 2.5 nm in this case. We assume that the distribution functions are uniform beyond 2.5 nm (a safe approximation, as the diameter of the protein is approximately 0.8 nm less than this) and add the analytical correction from 2.5 nm to infinity to the explicitly evaluated energy from cutoff r_c to 2.5 nm. We call this sum the evaluated dispersion energy neglected by cutoff at r_c. We then compare this with the analytical estimate from eq 5. As we will explain below, this approach is not adequate for providing quantitative corrections but provides a simple estimate of the rough magnitude of the correction.

In Figure 1, we compare the evaluated dispersion energy with the analytic dispersion energy, as well as the difference between the two, from 0.75 to 2.5 nm, as a function of the cutoff r_c. This is an estimate of the error inherent in the assumptions behind the analytical correction, which includes only an average dispersion energy from the part of the system outside the cutoff. The analytical correction is extremely poor for even moderate distance cutoffs in a protein−ligand complex. Even at a cutoff of 1.5 nm, the difference between the two correction methods is approximately 0.4 kcal/mol. If our target accuracy is 0.5 kcal/mol, this suggests that 1.5 nm would be the shortest cutoff that would give sufficiently precise results. These results are obtained with FK506, an unusually large ligand; a smaller molecule might have a dispersion correction only half of this value. In the case of such a molecule, the dispersion error would still be over 0.5 kcal/mol for cutoffs less than 1.2 nm.

For ligands in pure solvent, the analytical dispersion correction will be roughly correct, usually within the uncertainty of the hydration free energy. But for binding free energies, the free energy of decoupling the ligand in solvent must be subtracted from the free energy of decoupling in the complex. It is here that the analytical correction is not valid.

One potential approach to handle this correction is to use very long cutoffs. Frequently, long-range cutoffs of 1.2−1.4 nm are used, especially with twin-range potentials, where the longer range forces are not evaluated every step.[2,24] With an increase of dispersion cutoff to 1.6−1.8 nm, most of the neglected energy in ligand binding may be recovered, as can be seen in Figure 1. These simulations, however, can be computationally expensive: the computational effort goes roughly as r_c^3 for large r_c. Using twin-range multiple time step simulations can decrease this significantly but is still somewhat expensive and introduces some additional complications. Is it possible to measure observables that are independent of cutoff for heterogeneous systems without having to use such long cutoffs?

One possibility is to follow a procedure similar to what was done for Figure 1 above. We can evaluate the ligand−environment interaction energy from simulations of the bound and unbound ligand. The change in this interaction energy with cutoff, evaluated by reprocessing stored configurations, can be taken as the dispersion correction. This approach was applied in one study, removed much of the cutoff-dependence of the computed binding free energies,[14] and significantly improved fit to experiment. But this approach is fundamentally flawed. This analysis uses only the interaction energy between the ligand and the system, and therefore, it neglects any changes in the rest of the system as a function of cutoff. As the ligand is removed, water often occupies the binding site. Thus, changing the cutoff when the ligand is unbound should increase dispersion interactions between the water and the protein, as well. This effect will not be captured by considering only the interaction energy, and thus, this method is inadequate. In practice, for the FKBP system considered here, it seems to overestimate the magnitude of the correction by 10−20% because of the neglect of the changes in protein−environment interactions.

III. Correcting for Cutoffs in Nonisotropic Systems

In this work, we focus mainly on binding free energy calculations, as we can quantitatively demonstrate the importance of a proper treatment of long-range dispersion interactions for binding free energies. However, neglected dispersion interactions may also result in significant errors in other observables, as well, so we also briefly overview how to use similar reweighting strategies for more general observables.

A. Corrections to Binding Free Energies. Here, we present two methods appropriate for computing cutoff-independent binding free energies in heterogeneous, nonisotropic systems. In these methods, we simulate with the approximate, truncated function, using an analytical correction to the total pressure and energy to achieve (at least approximately) the right densities.

We then reprocess a relatively small number of (ideally uncorrelated) individual configurations with the long-range energy function. Finally, we use reweighting to estimate the free energy difference appropriate for the simulations with the long cutoffs. These methods do not make any assumptions about the shape of the van der Waals potential function.

To introduce and illustrate these methods, we apply them in the context of a particular approach to computing absolute binding free energies. We use an alchemical thermodynamic cycle in which we separately compute the free energy of decoupling a ligand from solvent and the free energy of decoupling a ligand from the binding site of a protein.[25,26] These corrections are applicable more broadly, as well, to other approaches, to alchemical calculations, or to potential of mean force calculations. "Decoupling" the ligand involves computing the free energy of eliminating the interactions between the ligand and the solvent or the ligand and the protein. Considering only the protein−ligand part of the cycle (for simplicity, we assume the electrostatic interactions between the ligand and the protein have already been removed), we seek to perform the transformation:

$$PL_N \rightarrow P + L_{N,LJ} \tag{6}$$

where P denotes the protein, L_N denotes the ligand with its electrostatic interactions with the environment removed, and $L_{N,LJ}$ denotes the ligand with its electrostatic and Lennard-Jones interactions removed. We will denote the free energy of this decoupling transformation as ΔG_{LJ}.

Now, we compute the free energy of decoupling by running separate simulations at a series of alchemical intermediate states (with associated intermediate λ values) spanning these two end states PL_N and $P + L_{N,LJ}$, using a relatively short Lennard-Jones cutoff r_{SR} at all states. We can compute an estimate of this free energy $\Delta G_{LJ}^{\mathrm{SR}}$ directly from the these short-range simulations using standard methods like thermodynamic itegration TI,[27] the Bennett acceptance ratio (BAR) approach,[28,29] or the weighted histogram analysis method (WHAM).[30,32] Reanalyzing configurations stored from these simulations, we can then compute an estimate of $\Delta G_{LJ}^{\mathrm{LR}}$, the free energy difference using a Hamiltonian computed with a cutoff large enough for the analytic long-range van der Waals correction to be appropriate. In general, the subscript SR represents quantities obtained at the simulated cutoff, and LR represents quantities corresponding to a longer dispersion cutoff, at which no simulations were actually performed.

1. EXP-LR Approach. The simplest way to do this is to reevalute the energy of each saved configuration using this longer cutoff, r_{LR}, from the simulations where the ligand is fully interacting (with $\lambda = 0$) and the simulations with the ligand fully decoupled (with $\lambda = 1$). For convenience, define a reduced potential $u(\lambda) = \beta(U(\lambda) + PV)$, where U is the potential energy, P is the external pressure, and V is the volume, the PV term being included if we are simulating an isobaric ensemble. F represents the free energy, Helmholtz or Gibbs, depending on the ensemble. Other terms such as chemical potential of the solvent or electrical work can also be included in this reduced potential. Then, at each of these two endpoint λ values, we use exponential averaging[33] (EXP; sometimes known as the Zwanzig relation) to compute the free energy of extending the cutoff from r_{SR} to r_{LR}:

$$\Delta F^{\mathrm{LR}}(\lambda) = -\beta^{-1} \ln\langle \exp[-\Delta u(\mathbf{q}^{\mathrm{SR}}, \lambda)]\rangle_{\mathrm{SR}} \tag{7}$$

where $\Delta u(\mathbf{q}^{\mathrm{SR}}, \lambda) = u^{\mathrm{LR}}(\mathbf{q}^{\mathrm{SR}}, \lambda) - u^{\mathrm{SR}}(\mathbf{q}^{\mathrm{SR}}, \lambda)$, and the angular brackets denote an average over all configurations, $\mathbf{q}^{\mathrm{SR}}$, obtained from a simulation using the short-range cutoff at that particular λ. u^{SR} and u^{LR} are the reduced potentials at short and long range, respectively. To obtain the total correction, we compute the difference between the corrections at $\lambda = 0$ and $\lambda = 1$. The total correction is then $\Delta F^{LRC} = \Delta F^{\mathrm{LR}}(\lambda = 1) - \Delta F^{\mathrm{LR}}(\lambda = 0)$.

The free energy $\Delta F_{LJ}^{\mathrm{LR}}$ will asymptotically (in the limit of a sufficiently large number of uncorrelated configurations sampled) satisfy $\Delta F_{LJ}^{\mathrm{SR}} + \Delta F_{LJ}^{\mathrm{LRC}} = \Delta F_{LJ}^{\mathrm{LR}}$. We will refer to this method as EXP-LR in this paper.

A known drawback of exponential averaging is that the most important energy differences are those at the extremes of the distribution of energy differences. The variance scales exponentially with the variance of the distribution of energy differences, meaning that it requires a large amount of data to obtain adequate statistics when the variance of the distributions becomes large with respect to kT.[34−36] Thus, we can expect EXP-LR to approach to work well only as long as the distribution of potential energy differences is relatively narrow with respect to kT, as we will typically have only a relatively small number of configurations at which we compute the long-range energies. Note that the magnitude of the difference between the short and the long-range potentials can be much larger than kT, as long as the variance is low.

2. WHAM-LR Approach. An alternative method can partially overcome this limitation of EXP-LR by using additional information from the intermediate states that were also simulated. Essentially, in EXP-LR, we compute a correction to the binding free energy by reweighting the end states. Another approach, instead of computing a correction at the end states, is to use all of the intermediate states and recompute a binding free energy using the samples first reweighted to long cutoffs. This can be done by the weighted histogram analysis method (WHAM),[30] which provides a way to estimate the free energy difference between adjacent pairs of alchemical states at an arbitrary potential, given the sampled configurations.[31] We can therefore use samples collected from simulations run with the short-ranged potential $u^{\mathrm{SR}}(\mathbf{q}; \lambda)$ to estimate the free energy difference $\Delta F_n^{\mathrm{LR}} \equiv F_{n+1}^{\mathrm{LR}} - F_n^{\mathrm{LR}}$ between two adjacent alchemical states, λ_n and λ_{n+1}, had we sampled from the long-ranged potential $u^{\mathrm{LR}}(\mathbf{q}; \lambda)$ instead:

$$F_n^{\mathrm{LR}} = -\beta^{-1} \ln \sum_{k=n}^{n+1} \sum_{t=1}^{T_k} \frac{\exp[-u^{\mathrm{LR}}(\mathbf{q}_{kt}^{\mathrm{SR}}; \lambda_n)]}{\sum_{k'=n}^{n+1} T_{k'} \exp[\beta F_{k'}^{\mathrm{SR}} - u^{\mathrm{SR}}(\mathbf{q}_{kt}^{\mathrm{SR}}; \lambda_{k'})]}$$

$$F_{n+1}^{\mathrm{LR}} = -\beta^{-1} \ln \sum_{k=n}^{n+1} \sum_{t=1}^{T_k} \frac{\exp[-u^{\mathrm{LR}}(\mathbf{q}_{kt}^{\mathrm{SR}}; \lambda_{n+1})]}{\sum_{k'=n}^{n+1} T_{k'} \exp[\beta F_{k'}^{\mathrm{SR}} - u^{\mathrm{SR}}(\mathbf{q}_{kt}^{\mathrm{SR}}; \lambda_{k'})]} \tag{8}$$

Here, T_n denotes the number of configurations stored from the simulation conducted at potential $U^{\mathrm{SR}}(\mathbf{q}; \lambda_n)$, and the $\mathbf{q}_{nt}^{\mathrm{SR}}$, $t = 1, \ldots, T_n$ denote the configurations themselves. As both equations contain the free energies F_n and F_{n+1} (uniquely determined only up to an additive constant), they must be iterated to convergence. As can be seen, this scheme requires that, for each neighboring pair of alchemical intermediates, we compute both u^{SR} and u^{LR} at both λ_n and λ_{n+1}. $\Delta F_{LJ}^{\mathrm{LR}}$ will be equal to $\sum_{i=n}^{N-1} \Delta F_n^{\mathrm{LR}}$. In this study, we only include weighting over nearest states n and $n + 1$, but it could easily be generalized to more states. We will term this method WHAM-LR.

This expression does provide the total decoupling free energy using a long cutoff, which is what we seek. However, it uses only the (perhaps infrequently) stored configurations which we have reprocessed with a long cutoff. When the short-range component of the free energy calculations is done internally in a simulation package, the short-range free energy may have greater accuracy because of averaging over additional configurations. In such cases, it may be desirable to only estimate a correction using reweighting and add this to the short-range component computed with the full set of configurations. The correction can be computed by first using the stored configurations $\mathbf{q}_{nt}^{\mathrm{SR}}$ to compute:

$$F_n^{\mathrm{SR}} = -\beta^{-1} \ln \sum_{k=n}^{n+1} \sum_{t=1}^{T_k} \frac{\exp[-u^{\mathrm{SR}}(\mathbf{q}_{kt}^{\mathrm{SR}}; \lambda_n)]}{\sum_{k'=n}^{n+1} T_{k'} \exp[\beta F_{k'}^{\mathrm{SR}} - u^{\mathrm{SR}}(\mathbf{q}_{kt}^{\mathrm{SR}}; \lambda_{k'})]}$$

$$F_{n+1}^{\mathrm{SR}} = -\beta^{-1} \ln \sum_{k=n}^{n+1} \sum_{t=1}^{T_k} \frac{\exp[-u^{\mathrm{SR}}(\mathbf{q}_{kt}^{\mathrm{SR}}; \lambda_{n+1})]}{\sum_{k'=n}^{n+1} T_{k'} \exp[\beta F_{k'}^{\mathrm{SR}} - u^{\mathrm{SR}}(\mathbf{q}_{kt}^{\mathrm{SR}}; \lambda_{k'})]} \quad (9)$$

Again, these equations must be iterated to convergence. The necessary correction is then given by $\Delta F_n^{\mathrm{LRC}} = \Delta F_n^{\mathrm{LR}} - \Delta F_n^{\mathrm{SR}}$, and $\Delta F^{\mathrm{LRC}} = \sum_{n=1}^{N=1} \Delta F_n^{\mathrm{LRC}}$. The advantage of computing both the short and the long-range energy using these configurations is that because we use the same potential energies to compute the short-range free energies as the long-range free energies, variance from the short-range part of the total potential energy is eliminated. This leaves a free energy whose uncertainty depends only on the variances in the differences between the long-range and the short-range potential.

Since the variation of these long-range interactions between different configurations is generally much lower than the variance of the other energy terms, the number of configurations that must be stored and reprocessed in this manner to obtain results with a given uncertainty is relatively small compared with the amount of data needed to generate the same amount of uncertainty at short cutoffs. The computational expense involved in reprocessing will be much less than that in running the original simulation.

This method does not directly depend on the variance in the potential energy differences between r_{SR} and r_{LR} but rather on the overlap of the distributions of potential energy differences between different λ values, which means that it has a different limitation than EXP-LR. If the phase space overlap between λ values is low and few configurations are saved, then the uncertainty in $\Delta G_{LJ}^{\mathrm{LR}}$ will be dominated by the uncertainty in the free energies between intermediate states. In other words, we trade the need for tight distributions in the potential energy difference between short-range and long-range cutoffs with a need for overlapping distributions in the potential energy difference between the intermediate states.

3. Comparison of Approaches. In limit that the distribution of potential energies between r_{SR} and r_{LR} is sufficiently narrow, and the distribution of energies between intermediate states is narrow, both methods should give similar estimated free energy differences. In this case, EXP-LR will be preferable, as it involves fewer energy evaluations in the reprocessing step. With large differences in the phase space between intermediates, EXP-LR will also be preferable, if few configurations are stored for reprocessing. WHAM-LR can handle much larger variances in potential energy difference distributions. However, for efficient calculations, one typically uses λ states that are already widely spaced. If snapshots are stored less frequently than the energy differences are evaluated for the short-range calculations, the overlap between λ states in the reprocessed energies may be insufficient. It is not always clear a priori which method will give the best performance for a given choice of biomolecular system, short-range cutoff, and frequency of sampling configurations, although, as we show here, either correction gives far more consistent values than leaving results uncorrected.

In either case, error analysis should be done, as for any simulation method. Methods for estimating uncertainties for WHAM are given in ref 37 and Appendix A of ref 32 and for EXP are given in ref 34. Where specific uncertainty formulas are not available or unsuitable, block bootstrap analysis can be performed.[26]

B. Corrections to Other Observables. In general, other observables computed from molecular simulations may be sensitive to the Lennard-Jones cutoff used in the simulation. In this work, we focus on corrections to binding free energies. However, reweighting techniques can also be used to estimate corrections to other observables. One can run simulations with short cutoffs, reprocess configurations with longer cutoffs, and use reweighting to estimate corrected observables appropriate for the long cutoffs.

The corrected equilibrium expectation of an observable $A(\mathbf{q})$ for cutoff r_{LR} is given in terms of expectations with cutoff r_{SR} by the well-known umbrella reweighting formula

$$\begin{aligned}\langle A \rangle_{\mathrm{LR}} &= \frac{\int d\mathbf{q} \exp(-u^{\mathrm{LR}}(\mathbf{q}))A(\mathbf{q})}{\int d\mathbf{q} \exp(-u^{\mathrm{LR}}(\mathbf{q}))} \\ &= \frac{\int d\mathbf{q} \exp(-u^{\mathrm{SR}}(\mathbf{q})) \exp(-[u^{\mathrm{LR}}(\mathbf{q}) - u^{\mathrm{SR}}(\mathbf{q})])A(\mathbf{q})}{\int d\mathbf{q} \exp(-u^{\mathrm{SR}}(\mathbf{q})) \exp(-[u^{\mathrm{LR}}(\mathbf{q}) - u^{\mathrm{SR}}(\mathbf{q})])} \\ &= \frac{\langle \exp(-\Delta u(\mathbf{q}))A(\mathbf{q}) \rangle_{\mathrm{SR}}}{\langle \exp(-\Delta u(\mathbf{q})) \rangle_{\mathrm{SR}}} \end{aligned} \quad (10)$$

where $\Delta u(\mathbf{q}) \equiv u^{\mathrm{LR}}(\mathbf{q}) - u^{\mathrm{SR}}(\mathbf{q})$, and the average is over all configurations are sampled from the short-range ensemble.

IV. Simulation Methods

In section V, we present two different examples of the application of these methods in ligand binding. The first, the computation of the binding free energies of two molecules to FKBP-12 using the Folding@Home distributed computing system,[38] demonstrates the magnitude of the needed correction and the robustness of these methods in computing this correction in the limit of a large amount of sampling. With the second, the binding of toluene to the engineered binding pocket of T4 lysozyme, we give an example of what the performance of these correction methods may be like in a more typical, small computer cluster setting. In this system, we demonstrate some problems in signal-to-noise analysis when applying these methods and explore the small cutoff limits of the method.

The FKBP-12 simulations were run with the modified version of GROMACS 3.1.4 described in previous papers,[14,39−41] running in double precision and using the Folding@Home distributed computing network.[38] The two ligands considered were L20 (FK506) and L2 from Holt et al.[42] (Figure 2). Integration of the equations of motion was performed using the

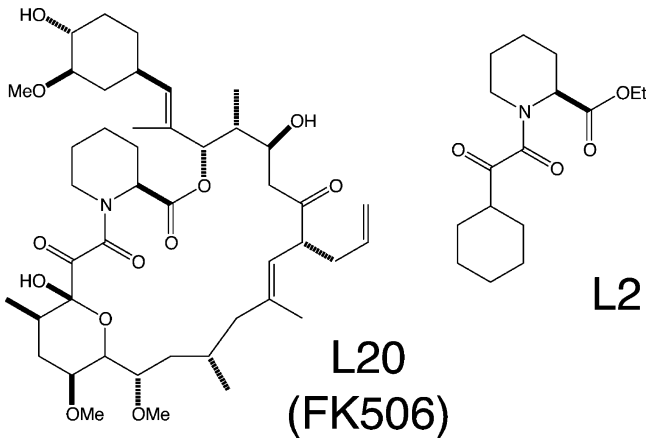

Figure 2. Free energies of binding of the ligands L20 (FK506) and L2 from Holt et al.[42] were calculated with different dispersion cutoffs and corrections in this study.

velocity Verlet algorithm[43] with a time step of 2.0 fs. Bond lengths were constrained with SETTLE[44,45] for waters and Rattle,[46] with a relative accuracy of 10^{-6}, for all other bonds. Twenty copies of each simulation were run from the same initial structure but different velocities. The inital structures were taken from the 10 ns point of 20 trajectories of a simulation of the protein−ligand complexes in a previous study.[14]

Andersen pressure and temperature control[47] were used to ensure sampling from an isobaric−isothermal (NPT) ensemble. The pressure was calculated using an atom-center based virial. Isotropic scaling of the truncated octahedron at each step was performed by scaling the atomic positions and then correcting the intramolecular distances with SETTLE[44] (for waters) or Rattle.[46] The Andersen piston mass was chosen to correspond with a compressibility with 4.5×10^{-5} bar^{-1} and a time constant of 1.67 ps. Andersen temperature control was implemented by complete reassignment of all velocities every 2.0 ps (every 1000 steps). The thermostat temperature was 283 K, the experimental temperature for the binding experiments.[42]

All FKBP simulations used the same electrostatic parameters. The neighborlist was set to 1.2 nm, updated every 10 steps. Particle mesh Ewald (PME)[1] was used, with the short-range Ewald cutoff set at 1.1 nm, spline interpolation of the charges with order 6, Fourier spacing of at most 0.15 nm, and a Gaussian width ($1/\beta$) of 0.323334 nm. Three different Lennard-Jones cutoffs, with the switching function implemented in GROMACS 3.1.4, were used in separate sets of simulations: switched off from 0.75 to 0.9 nm (referred to in this paper as LJ75), from 0.9 to 1.0 (LJ9), and from 1.0 to 1.1 nm (LJ10). A Lennard-Jones correction term to the total energy and virial, a switched-cutoff version of eqs 3 and 4, for both the r^{-12} and the r^{-6} terms was included. AMBER-99ϕ[48] parameters were used for the protein. The ligand parameters were identical to a previous FKBP binding paper,[49] using generalized Amber force field (GAFF) parameters[50] and partial charges generated by AM1-BCC[51] using MOPAC, with TIP3P water as the solvent. All topology files and initial coordinate files are included in Supporting Information.

To compute binding free energies, we used the thermodynamic cycle discussed elsewhere.[14,26,52] For this cycle, it is necessary to remove the ligand both from a solvent environment and from the binding site. We broke each of these calculations into Coulombic and Lennard-Jones parts. The Coulombic parts are done by scaling to zero the ligand electrostatic interactions entirely, while the Lennard-Jones components are done by decoupling the ligand Lennard-Jones interactions (leaving in intramolecular Lennard-Jones interactions).

For calculation of the free energies, potential energy differences between intermediates were output every 50 steps or 0.1 ps. The λ values of 0, 0.1, 0.25, 0.40, 0.55, 0.70, 0.85, and 1.0 were used for Coulombic decoupling, and 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.65, 0.70, 0.75, 0.80, 0.85, 0.9, 0.95, and 1.0 were used for decoupling of Lennard-Jones terms, using a soft core potential. Short-range free energies of decoupling were computed using BAR[28,29] (the equivalent of WHAM for two states) between neighboring intermediate λ values, again as described in a previous study.[40] Simulations were equilibrated for 2.0 ns (in the case of L20) or 1.0 ns (in the case of L2), after which the free energy data was collected. Data were collected from 20 independent runs to compute the short cutoff free energy. These runs were approximately 10.0 ns (depending on the intermediate state and run) for L20 and the longest cutoff for L2 and approximately 6.0 ns for the other two cutoffs for L2. Reprocessed energy differences were generated at each snapshot between all pairs of cutoffs, every 200 ps, for a total of approximately 700 snapshots per intermediate state for L20, 800 snapshots for the longest cutoff for L2, and 400 for the other two cutoffs. Uncertainties in the short-range free energies were taken as the standard deviation in the average of the free energies computed from the 20 separate copies.

For Figure 1, with L20, the 200 ps interval snapshot structures were reprocessed starting from 2.0 ns and extending to 6.0 ns, for a total of 400 snapshots. The uncertainty of the average excluded energy was estimated by taking the standard deviation in this energy over all snapshots and dividing by the square root of the total number of snapshots. Uncertainty in the average energies was less than 0.05 kcal/mol for all points, too small to display on the graph.

Pure TIP3P water simulations were also run with the same modified GROMACS version 3.1.4. These boxes were equilibrated for 100 ps at their new cutoffs, starting from a previously equilibrated water box, and then run for 1.0 ns. The PME parameters used order 6 interpolation, with a grid size of 0.1 nm, and Gaussian width ($1/\beta$) of 0.215556 nm. The temperature was set to 298 K, and velocities were reassigned every 5.0 ps. A neighborlist cutoff of 1.4 nm was used for all pure water simulations. In all other details, the simulations were identical to the FKBP simulations.

Calculations for the T4 lysozyme L99A mutant were run using GROMACS 3.3[3,53] with several crucial bugfixes described previously.[26] The system was prepared as discussed previously,[26] taking the initial ligand orientation from docking (to the apo structure) and using the apo structure of the protein. GAFF parameters[50] and AM1-CM2 charges[54] were used for the ligand, and AMBER 96 parameters[55] for the protein. For every system simulated, the system was first minimized with up to 5000 steps of L-BFGS[56] minimization, then 500 steps of steepest descent minimization;[57] then dynamics were begun after assigning velocities randomly from the Maxwell−Boltzmann distribution at 300 K. The Langevin integrator was used both to integrate the equations of motion and to provide temperature control at 300 K. The friction coefficient was 1 ps^{-1}. Dynamics used a time step of 2.0 fs; bond lengths for water were constrainted with SETTLE,[44] and protein and ligand bonds were constrained with LINCS.[58] Equilibration was done first for 10 ps using constant volume, then 100 ps using constant pressure and the Berendsen[59] barostat for temperature control. The compressibility was set to 4.5×10^{-5} bar^{-1}, and the time constant was set to 0.5 ps. Production simulations followed; they were 1 ns

in length at every alchemical intermediate state for the ligand in the protein and 5 ns each for the ligand alone in solvent.

All simulations for lysozyme used the same electrostatic parameters. The neighborlist was set to 1.2 nm and updated every 10 steps. Particle-mesh Ewald (PME)[1] was used, with the short-range Ewald cutoff set at 0.9 nm, spline interpolation order 6, Fourier spacing at most 0.1 nm, and a Gaussian width ($1/\beta$) of 0.260197 nm. We tried a variety of different Lennard-Jones cutoff schemes using switched cutoffs. Switching distances were 0.55 to 0.65 nm (scheme LJ55), 0.6 to 0.7 nm (scheme LJ6), 0.65 to 0.75 nm (LJ65), 0.7 to 0.8 nm (LJ7), 0.75 to 0.9 nm (LJ75), 0.8 to 0.9 nm (LJ8), 0.9 to 1.0 nm (LJ9), and 1.0 to 1.1 nm (LJ10). The GROMACS analytical Lennard-Jones correction to the energy and pressure (with average Lennard-Jones parameters) was included for the r^{-6} term. A key difference is that, in GROMACS 3.3, the dispersion potential depends on λ, and this adds an effective heterogeneous analytical dispersion correction to the short-range Lennard-Jones energy. In the lysozyme simulations, for the free energy of Coulombic charging, we used λ values of 0, 0.25, 0.5, 0.75, and 1.0. For Lennard-Jones decoupling, we used 0.0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.65, 0.7, 0.8, 0.85, 0.9, 0.95, and 1.0. As in the FKBP simulations, we used BAR to compute the free energies. However, because of differences in the codes used, to apply BAR, we saved full (single) precision simulation snapshots every 1.0 ps from every simulation and then reprocessed these snapshots at adjoining λ values in order to obtain the necessary potential energy differences to compute the short-range free energy.

Single precision structures were reprocessed using double-precision energy evaluation to avoid introducing additional numerical noise in the long-range corrections; preliminary analysis indicated that single precision energy evaluations could result in errors of up to 0.05 kcal/mol in the correction term. Reprocessing was done using the full saved trajectory in every case, so snapshots were spaced every 1.0 ps. We also used the confine-and-release approach, combined with umbrella sampling, to account for kinetic trapping of a valine side chain in the binding site, as described previously.[60] For this particular ligand, toluene, both valine rotameric states contribute to the overall binding free energy.

Because our focus here is on the Lennard-Jones decoupling portion of the calculation when the ligand is in complex with the protein, we directed our computational effort there. In particular, all other components of the binding free energy calculation were run with only one cutoff scheme (L8), and we performed only one trial at each λ value for the other components of the calculation. For the Lennard-Jones decoupling calculations in the complex, however, we performed three sets of simulations beginning from the same starting coordinates. The final results were obtained by averaging over these three sets.

A potential limitation of the approaches, as applied here, is that they are applied only to the Lennard-Jones decoupling portion of the cycle, when the ligand's electrostatic interactions have already been decoupled. Thus, implicitly, they assume that the density of Lennard-Jones sites around the bound ligand, outside the original short cutoff, does not depend on whether the ligand is interacting electrostatically with the protein. This is probably a good assumption for both systems considered here, as both are quite rigid and undergo minimal conformational change on ligand binding,[42,62] essentially all of which is localized to the region immediately around the ligand (thus falling within the short cutoff region). However, this assumption may break down in the case of protein conformational changes that are coupled to ligand binding and electrostatics and that involve protein regions outside the original cutoff. In such cases, the correction approaches described here should be applied not just to the Lennard-Jones decoupling portion of the binding calculation, but to the full binding calculation, including the electrostatics component. Such an extension was not done here, but is straightforward.

All uncertainties for lysozyme calculations and uncertainties in the long-range dispersion corrections for FKBP were computed using the block bootstrap approach described previously,[26] taking the standard deviation over 40 bootstrap trials, with block lengths taken to be equal to the autocorrelation time.

V. Results

As we discuss in Methods above, computed ligand binding free energies can depend on the Lennard-Jones cutoff used in the simulations, even when using an analytical correction which assumes the system is isotropic outside the cutoff. We presented approaches for computing corrections (EXP-LR and WHAM-LR) to obtain consistent free energies regardless of the cutoff with which the simulation was originally run. We present here the results of applying these corrections to the systems described above.

In the first of these, FKBP, we studied two ligands. FK506 (a relatively large molecule, over 700 Da) binds with high affinity, with an experimental binding free energy of -12.1 ± 0.2 kcal/mol at 283 K (from a K_d of 0.45 ± 0.15 nM); L2 binds with significantly lower affinity (-7.38 ± 0.08 kcal/mol, from a K_d of 2000 ± 300 nM). With simulations run using three different cutoff schemes, as described in Methods, we can directly test the consistency of the long-range correction methods over a range of cutoff distances. Not only can we compare the sum of the short-range free energy plus the long-range correction term, but also we can measure the free energy to change the cutoffs using the correction. If the formalism is correct, we will obtain the same free energy difference between cutoffs r_A and r_B whether we sample using cutoff r_A or cutoff r_B.

Table 2 shows the free energy difference computed between simulations with different Lennard-Jones cutoffs, evaluated by EXP-LR and WHAM-LR, for the two FKBP ligands studied. Both methods clearly produce corrections that agree well with one another, forward and backward, within statistical uncertainty, though EXP-LR appears to have slightly lower variance, for reasons we discuss later.

We can then examine the effect of these new dispersion corrections on the total free energy of binding. In our thermodynamic cycle,[14,26] the total binding energy (ΔG°) is the free energy of decoupling the ligand from the complex with the solvated protein (ΔG_P) minus the free energy of decoupling the ligand from the solvent (ΔG_S). Both of these decoupling free energies can be written as the sum of the free energy to turn off the Coulombic interactions ($\Delta G_{S,C}$) and ($\Delta G_{P,C}$), plus the free energy to turn off the Lennard-Jones interactions ($\Delta G_{S,LJ}$) and ($\Delta G_{P,LJ}$), plus a free energy for restraining the ligand in complex and releasing it to the standard state when unbound, ΔG_r°. The Lennard-Jones decoupling free energy, using a longer cutoff (ΔG_{LJ}^{LR}), is the sum of the short-range Lennard-Jones energy (ΔG_{LJ}^{SR}), and any dispersion correction (ΔG_{LJ}^{LRC}). The dispersion correction will not affect either of the Coulombic decoupling energies. We will assume that the Lennard-Jones decoupling energy in pure solvent includes the analytical correction, which is presumed to be valid for this smaller system,

TABLE 2: Free Energy Differences between Simulations with Different Dispersion Cutoffs Evaluated by EXP-LR and WHAM-LR for L20(FK506) and L2 Bound to FKBP-12

	L20 (FK506)		L2	
simulations[a]	WHAM-LR (kcal/mol)	EXP-LR (kcal/mol)	WHAM-LR (kcal/mol)	EXP-LR (kcal/mol)
	From LJ75			
LJ9	−4.23 ± 0.05	−4.22 ± 0.02	−1.71 ± 0.05	−1.78 ± 0.02
LJ10	−6.39 ± 0.05	−6.42 ± 0.02	−2.50 ± 0.05	−2.66 ± 0.03
LONG[b]	−12.20 ± 0.07	−12.28 ± 0.03	−4.72 ± 0.10	−4.89 ± 0.04
	From LJ9			
LJ75	4.27 ± 0.05	4.19 ± 0.01	1.71 ± 0.05	1.76 ± 0.02
LJ10	−2.24 ± 0.02	−2.20 ± 0.01	−0.88 ± 0.03	−0.85 ± 0.01
LONG	−8.12 ± 0.03	−8.02 ± 0.01	−3.09 ± 0.04	−3.07 ± 0.02
	From LJ10			
LJ75	6.50 ± 0.04	6.37 ± 0.02	2.59 ± 0.05	2.58 ± 0.02
LJ9	2.20 ± 0.03	2.16 ± 0.01	0.86 ± 0.02	0.84 ± 0.01
LONG	−5.86 ± 0.02	−5.82 ± 0.01	−2.21 ± 0.02	−2.21 ± 0.01
	Sums[b]			
(LJ75 to LJ9)+(LJ9 to LONG)	−12.35 ± 0.06	−12.24 ± 0.03	−4.80 ± 0.06	−4.85 ± 0.03
(LJ75 to LJ10)+(LJ10 to LONG)	−12.25 ± 0.05	−12.24 ± 0.03	−4.71 ± 0.05	−4.87 ± 0.03
(LJ75 to LJ9)+(LJ9 to LJ10)+(LJ10 to LONG)	−12.33 ± 0.06	−12.24 ± 0.02	−4.80 ± 0.06	−4.84 ± 0.02

[a] LJ75 indicates a cutoff tapered between 0.75 and 0.9 nm, LJ9 indicates a cutoff between 0.9 and 1.0 nm, and LJ10 indicates a cutoff between 1.0 and 1.1 nm. LONG indicates the cutoff independent binding energy described in the paper. Free energies computed with WHAM-LR and EXP-LR agree very well in both forward and reverse directions. The total correction to go from LJ75 to LONG is also consistently estimated. [b] Summing the free energies between intermediate cutoff ranges demonstrates even shorter cutoffs correctly predict the total correction to very long-range cutoffs.

TABLE 3: Free Energies of Binding for L20 and L2 to FKBP-12, Computed with Three Different Dispersion Cutoffs[a]

simulation[b]	ΔG_{LJ}^{SR} (kcal/mol)[c]	ΔG_{LJ}^{LRC} analytic (kcal/mol)[c]	ΔG_{LJ}^{LRC} EXP-LR (kcal/mol)[c]	$\Delta G°$ analytic (kcal/mol)[d]	$\Delta G°$ EXP-LR (kcal/mol)[d]	$\Delta G°_{exp}$ (kcal/mol)[e]
			L20 (FK506)			
LJ75	−1.13 ± 0.17	−10.40	−12.67 ± 0.03	−13.02 ± 0.17	−15.29 ± 0.17	−12.1 ± 0.20
LJ9	−5.60 ± 0.22	−6.77	−8.41 ± 0.01	−13.86 ± 0.22	−15.50 ± 0.22	−12.1 ± 0.20
LJ10	−7.35 ± 0.23	−5.01	−6.21 ± 0.01	−13.85 ± 0.23	−15.05 ± 0.23	−12.1 ± 0.20
			L2			
LJ75	−3.65 ± 0.18	−3.88	−5.04 ± 0.04	−5.48 ± 0.18	−6.64 ± 0.18	−7.38 ± 0.08
LJ9	−5.37 ± 0.13	−2.53	−3.22 ± 0.02	−5.85 ± 0.13	−6.54 ± 0.13	−7.38 ± 0.08
LJ10	−6.38 ± 0.11	−1.87	−2.36 ± 0.01	−6.20 ± 0.11	−6.69 ± 0.11	−7.38 ± 0.08

[a] When properly corrected, computed binding free energies should be cutoff-independent; they are not, except in the second to last column. The cutoff dependence is, in fact, drastic except with the EXP-LR or WHAM-LR corrections. Using either EXP-LR or WHAM-LR as described in this study yields free energies that are consistent within the mututal statistical uncertainty. [b] Different cutoff schemes are tapered from 0.75 to 0.9 nm (LJ75), tapered from 0.9 to 1.0 nm (LJ9), and tapered from 1.0 to 1.0 nm (L10). [c] First three columns are the short-range Lennard-Jones component of the binding free energy, the analytical estimate of the long-range correction, and the long-range correction from EXP-LR. [d] Last two columns are the computed binding free energies using different correction schemes (the analytical correction and the EXP-LR scheme). The binding energy includes additional terms, independent of the dispersion cutoff, as detailed in the text, but exclude the uncertainty of these extra terms, as they are independent of the Lennard-Jones cutoffs in these tables. [e] Final column is the experimental value.

and thus the total $\Delta G_{S,LJ}^{LR}$ does not depend on cutoff. We can then write the binding energy as:

$$\begin{aligned}\Delta G° &= \Delta G_P - \Delta G_S + \Delta G_r° \\ &= (\Delta G_{P,C} + \Delta G_{P,LJ}) - (\Delta G_{S,C} + \Delta G_{S,LJ}) + \Delta G_r° \\ &= (\Delta G_{LJ}^{SR} + \Delta G_{LJ}^{LRC}) + \Delta G_{REM} \\ &= \Delta G_{LJ}^{LR} + \Delta G_{REM} \end{aligned} \quad (11)$$

where ΔG_{REM} includes all Coulombic decoupling terms and the solvent Lennard-Jones decoupling term, as well as the restraining/standard state component, and we have dropped the subscript P from the protein-bound Lennard-Jones terms for simplicity in the rest of the section. For FKBP, this remainder term, computed with BAR as described in ref 14 is −1.49 ± 0.24 kcal/mol for L20 and 2.05 ± 0.08 kcal/mol for L2.

For L20, ΔG_{LJ}^{SR} computed with different cutoffs has a spread of more than 6 kcal/mol (second column of Table 3). Adding the analytic correction ΔG_{LJ}^{LRC} (third column) results in binding free energies with a spread of 0.8 kcal/mol (fifth column). This is larger than the mutual uncertainty between different cutoff schemes and still more than 2.0 kcal/mol less than the total binding free energy when EXP-LR or WHAM-LR is applied. With ΔG_{LR}^{LRC} computed with EXP-LR (third column), the spread in computed binding free energies (sixth column) drops to only 0.45 kcal/mol, which is almost within the mutual uncertainty of the calculations. In the case of L2, ΔG_{LJ}^{SR} for different cutoffs are spread over 3 kcal/mol, while the spread in analytically corrected binding energies is approximately 0.7 kcal/mol. Again, this is not nearly as good as the results with the EXP-LR or WHAM-LR correction, which have a spread of only 0.15 kcal/mol in the total binding energy, well within the mutual uncertainties of the calculations with the three different cutoffs. $\Delta G_{P,LJ}^{SR}$ is much more sensitive to the dispersion cutoff than is the binding energy, because, for the binding free energy, the cutoff dependence of $\Delta G_{P,LJ}^{SR}$ partially cancels with that of $\Delta G_{S,LJ}^{SR}$.[68]

These simulations demonstrate the importance of including such a correction in ligand binding, as a small change in cutoff

TABLE 4: Lennard-Jones Decoupling Free Energies for Toluene in the Apolar Engineered Site in T4 Lysozyme Using Different Dispersion Cutoff Schemes[a]

simulation[b]	ΔG_{LJ}^{SR} (kcal/mol)	ΔG_{LJ}^{LRC} WHAM-LR (kcal/mol)	ΔG_{LJ}^{LRC} EXP-LR (kcal/mol)	ΔG_{LJ}^{LR} WHAM-LR (kcal/mol)	ΔG_{LJ}^{LR} EXP-LR (kcal/mol)
LJ55	-6.20 ± 0.03	-1.48 ± 0.34	-1.68 ± 0.20	-7.68 ± 0.35	-7.88 ± 0.20
LJ6	-6.47 ± 0.03	-1.22 ± 0.14	-1.11 ± 0.08	-7.68 ± 0.15	-7.58 ± 0.09
LJ65	-6.73 ± 0.03	-0.90 ± 0.08	-1.11 ± 0.07	-7.63 ± 0.09	-7.85 ± 0.07
LJ7	-7.02 ± 0.03	-0.69 ± 0.06	-1.03 ± 0.03	-7.71 ± 0.06	-8.05 ± 0.04
LJ75	-7.54 ± 0.03	-0.43 ± 0.03	-0.67 ± 0.02	-7.97 ± 0.05	-8.21 ± 0.04
LJ8	-7.29 ± 0.03	-0.41 ± 0.03	-0.50 ± 0.02	-7.70 ± 0.04	-7.79 ± 0.04
LJ9	-7.57 ± 0.03	-0.29 ± 0.01	-0.29 ± 0.01	-7.87 ± 0.03	-7.86 ± 0.03
LJ10	-7.89 ± 0.04	-0.15 ± 0.01	-0.13 ± 0.01	-8.04 ± 0.04	-8.02 ± 0.04

[a] Shown are the short-range component, the long-range corrections computed with WHAM-LR and EXP-LR, and then the overall decoupling free energies with the corrections. The total here is a favorable contribution for binding, so without additional corrections, the shorter cutoffs neglect more than 1 kcal/mol of the total binding free energy. [b] Different simulations use different dispersion cutoff schemes, from short to long, and are as detailed in Methods.

of a dispersion term can significantly affect computed binding free energies; in this example, by almost 1 kcal/mol for typical cutoffs. Even with a relatively long cutoff of 1.0 to 1.1 nm, the result differs from the full dispersion binding energy by more than 0.5−1.5 kcal/mol. Table 3 demonstrates that these methods can very effectively compensate for extremely large differences caused by differing dispersion cutoffs.

We also note that the uncertainty in the long-range correction for L2 is not any smaller than that in the correction for L20, even though the correction itself is smaller.[61] The fact that the uncertainty does not decrease with ligand size highlights a potential problem with the methods presented here. The variance in the potential energy differences from configuration to configuration, and thus the uncertainty in WHAM-LR and EXP-LR, is primarily dependent on the environment. Thus, as the size of the system grows, the uncertainty grows larger. For large systems, uncertainties may grow unacceptably large.

We also studied the binding of toluene in the apolar T4 lysozyme L99A binding cavity. This cavity is interesting as an extremely simple cavity binding site that binds relatively small nonpolar molecules like benzene and toluene, among others.[62−65] As in FKBP, we examine the binding of toluene in this cavity by computing its absolute binding free energy. The simulation lengths we use (1 ns at each λ value) are relatively short and easily accessible for a variety of systems on small computer clusters. We show uncorrected and corrected Lennard-Jones decoupling free energies for the ligand-in-complex portion of the calculation in Table 4.

The total binding free energy for toluene, in this case, is given by $\Delta G^\circ = \Delta G_{LJ} + 3.55 \pm 0.05$ kcal/mol, where ΔG_{LJ} is the Lennard-Jones decoupling energy in the complex. The other term includes the remainder of the binding free energy including ligand desolvation and other factors, computed as described previously.[26,66] With ΔG_{LJ} taken as the uncorrected ΔG_{LT}^{SR} from the L55 cutoff in Table 4, this gives a binding free energy of -2.65 ± 0.05 kcal/mol (using the analytical correction implemented in GROMACS), while the value from the LJ10 cutoff scheme gives a binding free energy of -4.33 ± 0.06 kcal/mol (using the same correction). This is very large difference and is solely due to the choice of Lennard-Jones cutoff. The experimental value is -5.52 ± 0.06 kcal/mol.[63] While neither of the computed values agree perfectly with experiment, obviously the value from the LJ10 scheme agrees much better.

The goal of the methods described here is to provide corrections to yield computed free energies that are independent of cutoff. When these corrections are applied, the ΔG_{LJ}^{LR} values (Table 4) are all within a much tighter range (-7.63 to -8.04 kcal/mol with WHAM-LR and -7.58 to -8.21 kcal/mol with EXP-LR). This leads to a much smaller range of possible resulting binding free energies with different cutoff schemes (-4.08 to -4.48 kcal/mol with WHAM-LR and -4.03 to -4.66 kcal/mol with EXP-LR), rather than the range of nearly 1.7 kcal/mol when no correction was applied. Thus, the corrections greatly improve the consistency of computed binding free energies and improve agreement with experiment, even down to cutoffs as small as 0.6 nm.

As discussed in Methods, a potential pitfall of the methods as presented here is that the correction is applied only for the part of the thermodynamic cycle with the decoupling of Lennard-Jones particles and not for the decoupling of electrostatic interactions. As the results presented for FKBP and lysozyme are essentially independent of cutoff, any additional correction for the change in dispersion correction when electrostatic interactions are decoupled appears to be negligible but might need to be included in other cases.

The corrections computed with lysozyme are less precise than the corrections with FKBP. For example, EXP-LR and WHAM-LR diverge to a greater extent than for FKBP, even with the LJ75 cutoff scheme. This is due in large part to fewer decorrelated snapshots being used in the computations for lysozyme but also because the lysozyme system is larger, with a total potential energy twice as large. This results in larger errors in the correction term, primarily because of noise in the water−water interactions. For example, the variances in $U^{LR}(\mathbf{q}^{SR}, \lambda = 1) - U^{SR}(\mathbf{q}^{SR}, \lambda = 1)$ and $U^{LR}(\mathbf{q}^{SR}, \lambda = 0) - U^{SR}(\mathbf{q}^{SR}, \lambda = 0)$ are 0.41 and 0.43 kcal/mol, respectively, for the LJ75 cutoff for both FKBP ligands, but are 0.52 and 0.60 kcal/mol for LJ75 in lysozyme, and rising to over 1.7 for LJ55. Since EXP-LR grows exponentially with energy scale *kT*, as the width of the distribution becomes comparable to *kT*, its variance begins to grow.[34]

VI. Conclusions

Our main goal in this study is to obtain computed binding free energies that are independent of the Lennard-Jones cutoff with which the original simulation was run; any improved agreement with experiment is a bonus. In this study, we show that, without corrections or by using only analytical corrections, computed binding free energies can depend strongly on the cutoff used in the simulation. We find that this cutoff dependence can be up to 0.8 kcal/mol in the FKBP case and up to 1.6 kcal/mol for lysozyme.[69] By using the two new methods we introduce (EXP-LR and WHAM-LR), the computed ligand binding free energies are essentially cutoff-independent. This strongly suggests that previous free energy calculations done

without long-range corrections may have produced results with significant cutoff dependence.

In most cases, EXP-LR will likely be the preferred method. It requires fewer energy re-evaluations, only at the states where $\lambda = 0$ and $\lambda = 1$, and thus will be significantly more efficient if there are a large number of intermediate states. If the width of the distribution of energy differences is significantly larger than *kT*, or the system is particularly large, however, WHAM-LR should also be tested. Error analysis is essential to identify signs of potential problems.

We have primarily described our method in terms of correcting to a very long-range cutoff in order to remove an arbitrary unphysical simulation parameter. This is not unprecedented and can improve agreement with experiment. Several studies have shown that using corrections to include long-range dispersion can improve fit to experiment for many of the current biomolecular force fields.[20,40] Additionally, here, applying corrections usually improves the agreement with experiment relative to the uncorrected free energies. But this is not universally true. For FKBP, the binding free energy of L20 actually becomes somewhat further away from the experimental binding affinity with the addition of the correction.

The method presented here can alternatively be used to correct observables to any alternate cutoff scheme for dispersion. For example, if a particular parameter set performs best with 0.9 nm abrupt dispersion cutoffs, the formalism here can be used to obtain correct free energies for this cutoff independent of the cutoff used to obtain the sampling, while tapered cutoffs with better energy conservation properties are used for molecular dynamics.

In the future, an alternate cutoff independent method to include Lennard-Jones interactions is to implement a fast Ewald summation over Lennard-Jones sites, which was already proposed and derived by the originators of the smooth PME method.[1] This summation would presumably be very fast, because of the fast decay of both the r^{-12} and the r^{-6} functional forms would likely add little overhead to most simulations if implemented efficiently. However, it becomes complicated to implement with additive *ij* combination rules such as the current CHARMM or AMBER potential functions[1] and will not work with non-polynomial dispersion functions. If Ewald summation of long-range Lennard-Jones treatments were implemented, the corrections presented here would thus be unnecessary, and for ease of use, this approach may eventually turn out to be preferable to methods presented here. Introducing periodicity into the dispersion interaction does add artifacts, but since the Lennard-Jones interactions fall off much faster than electrostatic interactions, if simulation boxes are large enough that electrostatic errors are small, any Lennard-Jones artifacts will be negligible.

Other non-isotropic systems may also require improved treatment of long-range dispersion interactions. For example, protein aggregation is important in a number of diseases,[67] and small changes to the strength of dispersion interactions (i.e., due to the handling of long-range dispersion interactions) could potentially make the difference between aggregation and solubility in simulations of protein aggregation. However, the methods here might not be sufficient to handle such cases, where the phase space difference between long and short-range dispersion cutoffs is large.

The formalisms presented in this paper may also be generalized to correct for the exclusion of other relatively low variance potential energy terms in a simulation; for example, one could potentially simulate with a fast and coarse set of parameters for PME treatment of long-range electrostatics, to gain computational speed, and then use reprocessing with better parameters and reweighting to obtain corrections.

We have presented two general approaches for correcting for the effect of truncating long-range Lennard-Jones interactions, in order to obtain conistency in computed free energies regardless of the cutoff. This study demonstrates that computational estimates of binding free energies from studies using different cutoffs might be significantly less comparable to each other than previously assumed, as these studies neglected significant long-range dispersion interactions beyond the cutoff. Discrepancies in binding free energies between different studies on the same system might partially be explained by different cutoffs. We believe that future binding free energy studies should include this or another accurate treatment of the long-range correction, to allow easier comparisons of binding free energies computed with different cutoff schemes. Additionally, future force fields should be parametrized with this cutoff dependence in mind.

Acknowledgment. M.R.S. and D.L.M. contributed equally to this work. D.L.M. thanks Ken A. Dill (UC-San Francisco) for postdoctoral support. M.R.S. thanks the NIH for support via a NRSA Ruth L. Kirschstein postdoctoral fellowship, as well as Richard A. Friesner (Columbia University). J.D.C. thanks HHMI and IBM predoctoral fellowships and Ken A. Dill for support. Calculations for lysozyme were performed on the UCSF QB3 shared computing facility. The authors thank the users of Folding@Home for donating their CPU cycles for the FKBP simulations.

Supporting Information Available: Input topology and coordinate files for the simulations discussed in this work. This information is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(2) Tuckerman, M.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990−2001.
(3) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306−317.
(4) Steinbach, P. J.; Brooks, B. R. *J. Comp. Chem.* **1994**, *15*, 667−683.
(5) Sagui, C.; Darden, T. A. *Ann. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 155−179.
(6) Deserno, M.; Holm, C. *J. Chem. Phys.* **1998**, *109*, 7678−7693.
(7) Kitchen, D. B.; Hirata, F.; Westbrook, J. D.; Levy, R. M.; Kofke, D.; Yarmush, M. *J. Comp. Chem.* **1990**, *11*, 1169.
(8) Guenot, J.; Kollman, P. A. *J. Comp. Chem.* **1993**, *14*, 295−311.
(9) Hünenberger, P. H.; Börjesson, U.; Lins, R. D. *Chimia* **2001**, *55*, 861−866.
(10) Lisal, M.; Kolafa, J.; Nezbeda, I. *J. Chem. Phys.* **2002**, *117*, 8892−8897.
(11) Mark, P.; Nilsson, L. *J. Comp. Chem.* **2002**, *23*, 1211−1219.
(12) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451−5459.
(13) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids;* Oxford University Press: New York, 1987.
(14) Shirts, M. R. Ph.D. Dissertation, Stanford University, 2005.
(15) Head-Gordon, T.; Hura, G. *Chem. Rev.* **2002**, *102*, 2651−2670.
(16) Hura, G.; Russo, D.; Glaeser, R. M.; Head-Gordon, T.; Krack, M.; Parrinello, M. *Phys. Chem. Chem. Phys.* **2003**, *5*, 1981−1991.
(17) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004**, *120*, 9665−9678.
(18) Kollman, P. A.; Dixon, R.; Cornell, W.; Fox, T.; Chipot, C.; Pohorille, A. The development/application of a 'minimalist' organic/biochemical molecular mechanic force field using a combination of *ab initio* calculations and experimental data. In *Computer Simulation of Biomolecular Systems*, Vol. 3; Wilkinson, A., Weiner, P., van Gunsteren, W. F., Eds.; Elsevier: Amsterdam, The Netherlands, 1997.

(19) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(20) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. *J. Chem. Phys.* **2003**, *119*, 5740−5761.
(21) Personal communication.
(22) GROMACS manuals 3.1.1, 3.2, or 3.3, at http://www.gromacs.org/gromacs/documentation/paper-manuals.html.
(23) Lagüe, P.; Pastor, R. W. *J. Phys. Chem. B* **2004**, *108*, 363−368.
(24) van Gunsteren, W. F.; Berendsen, H. J. C. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 992−1023.
(25) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047−1069.
(26) Mobley, D. L.; Chodera, J. D.; Dill, K. A. *J. Chem. Phys.* **2006**, *125*, 084902.
(27) Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1991**, *95*, 1175−1188.
(28) Bennett, C. H. *J. Comp. Phys.* **1976**, *22*, 245−268.
(29) Shirts, M. R.; Bair, E.; Hooker, G.; Pande, V. S. *Phys. Rev. Lett.* **2003**, *91*, 140601.
(30) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(31) The derivation of the WHAM estimator used here, eq 21 of ref 30, is too involved to reproduce here, so we refer the reader there for a detailed exposition.
(32) Chodera, J. D.; Swope, W. C.; Pitera, J. W.; Seok, C.; Dill, K. A. *J. Chem. Theor. Comput.* **2007**, *3*, 26−41.
(33) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420−1426.
(34) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 144107.
(35) Lu, N. D.; Kofke, D. A. *J. Chem. Phys.* **2001**, *114*, 7303−7311.
(36) Lu, N. D.; Kofke, D. A. *J. Chem. Phys.* **2001**, *115*, 6866−6875.
(37) Janke, W. Statistical analysis of simulations: Data correlations and error estimation. In *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*, Vol. 10; Grostendorst, J., Marx, D., Murmatsu, A., Eds.; Rolduc Conference Centre: Kerkrade, The Netherlands, 2002.
(38) Shirts, M.; Pande, V. S. *Science* **2000**, *290*, 1903−1904.
(39) Jayachandran, G.; Shirts, M. R.; Park, S.; Pande, V. S. *J. Chem. Phys.* **2006**, *125*, 084901.
(40) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 134508.
(41) A bug fix was added to correctly compute the reciprocal space forces of mutating atoms; the effect of this on the results presented for smaller molecules in the previous study[40] was orders of magnitude smaller than the uncertainty but became sigificant with the large proteins in this study.
(42) Holt, D. A.; Luengo, J. I.; Yamashita, D. S.; Oh, H. J.; Konialian, A. L.; Yen, H. K.; Rozamus, L. W.; Brandt, M.; Bossard, M. J.; Levy, M. A.; Eggleston, D. S.; Liang, J.; Schultz, L. W.; Stout, T. J.; Clardy, J. *J. Am. Chem. Soc.* **1993**, *115*, 9925−9938.
(43) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. *J. Chem. Phys.* **1982**, *76*, 637−649.
(44) Miyamoto, S.; Kollman, P. A. *J. Comp. Chem.* **1992**, *13*, 952−62.
(45) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004**, *120*, 9665−78.
(46) Andersen, H. C. *J. Comp. Phys.* **1983**, *52*, 24−34.
(47) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384−2393.
(48) Sorin, E. J.; Pande, V. S. *Biophys. J.* **2005**, *88*, 2472−2493.
(49) Fujitani, H.; Tanida, Y.; Ito, M.; Shirts, M. R.; Jayachandran, G.; Snow, C. D.; Pande, E. J. S. V. S. *J. Chem. Phys.* **2005**, *123*, 084108.
(50) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comp. Chem.* **2004**, *25*, 1157−1174.
(51) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Comp. Chem.* **2000**, *21*, 132−46.
(52) Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. *J. Phys. Chem. A* **2003**, *107*, 9535−9551.
(53) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comp. Chem.* **2005**, *26*, 1701−1718.
(54) Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **1996**, *100*, 16385−16398.
(55) Kollman, P. A. *Acc. Chem. Res.* **1996**, *29*, 461−469.
(56) Liu, D. C.; Nocedal, J. *Math. Prog. B* **1989**, *45*, 503−528.
(57) We found that the L-BFGS minimizer in GROMACS occasionally terminates extremely early, after only a few steps, so we used additional steepest descents minimization to ensure that minimization was adequate.
(58) Hess, B.; Bekker, H.; Berendsen, H. J. C. *J. Comp. Chem.* **1997**, *18*, 1463−1472.
(59) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3584−3690.
(60) Mobley, D. L.; Chodera, J. D.; Dill, K. A. *J. Chem. Theor. Comput.* **2007**, *3*, 1231−1235.
(61) The listed uncertainties are actually slightly higher for L2 than for L20 using the 0.75−0.9 and 0.9−1.0 cutoffs, but this is principally because of the smaller number of samples collected for L2 at these cutoffs.
(62) Morton, A.; Matthews, B. W. *Biochemistry* **1995**, *34*, 8576−8588.
(63) Morton, A.; Baase, W. A.; Matthews, B. W. *Biochemistry* **1995**, *34*, 8564−8575.
(64) Wei, B. Q.; Baase, W. A.; Weaver, L. H.; Matthews, B. W.; Shoichet, B. K. *J. Mol. Biol.* **2002**, *322*, 339−355.
(65) Graves, A. P.; Brenk, R.; Shoichet, B. K. *J. Med. Chem.* **2005**, *48*, 3714−3728.
(66) Mobley, D. L.; Graves, A. P.; McReynolds, A. C.; Shoichet, B. K.; Dill, K. A. *J. Mol. Biol.* **2007**, *371*, 1118−1134.
(67) Koo, E. H.; Peter T. Lansbury, J.; Kelly, J. W. *Proc. Natl. Acad. Sci., U.S.A.* **1999**, *96*, 9989−9990.
(68) If different cutoffs are used to compute the free energies of the complexed ligand and solvated ligand, this cancellation of error will not fully occur, and omitting a proper treatment of the long-range correction will lead to even larger errors in the total binding free energy than exist between columns five and six of Table 3.
(69) In GROMACS 3.3, the heterogeneous dispersion correction to the chemical potential was implemented and used for these calculations. As discussed in the text, this can lead to inaccuracies, and likely explains part of the larger cutoff-dependence for toluene in lysozyme than the much larger FK506. The fully buried binding site of toluene also may explain part of the larger discrepancy.