JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Efficient and Accurate Free Energy Calculations on Trypsin Inhibitors

Anita de Ruiter and Chris Oostenbrink*

Institute for Molecular Modeling and Simulation, University of Natural Resources and Life Sciences (BOKU), Vienna, Austria

Supporting Information

ABSTRACT: Several combinations of free energy calculation methods have been applied to determine the relative free energies of binding between eight para-substituted benzamidines in complex with the serine protease trypsin. With the aim to improve efficiency and maintain accuracy, the linear response approximation (LRA), linear interaction energy (LIE) and third power fitting (TPF) are combined with the one-step perturbation (OSP) to determine the polar and apolar contributions to the free energy, respectively. It is shown that the combination TPF/OSP gives the most accurate results and is 4.5 times more efficient than the rigorous thermodynamic integration (TI). By projecting the electrostatic preorganization energy from the OSP simulations, an increase in efficiency of a factor 7.5 can even be achieved. Loss of accuracy with respect to the TI data is limited to 3.9 and 5.6 kJ/mol, respectively, making it an attractive approach for lead optimization programs in drug research.

INTRODUCTION

Molecular recognition is extremely important for biological function. It involves the specific, noncovalent formation of, e.g., receptor−ligand, antigen−antibody, and protein−DNA complexes, which are crucial for normal cell functioning. Therefore, accurate prediction of binding affinities is very helpful in, e.g., drug design.

A significant part of computational chemistry involves the determination of the free energies of ligands binding to proteins. In the last decades, a variety of methods have been developed for this purpose, showing large differences in their efficiency and accuracy. Naturally, the aim is to obtain the highest possible accuracy with the least amount of effort. One of the fastest approaches is the use of empirical scoring functions, which are most commonly applied in docking studies.[1,2] An estimation of the affinity is typically based on a single conformation, including several contributions like hydrogen bonds and penalties for the loss of rotational entropy. Empirical scoring functions are thus very fast and, therefore, suitable to assess a large number of ligands in a relatively short time. But at the same time, they lack the accuracy to discriminate ligands with similar binding affinities.[3,4] On the other hand, methods based on statistical mechanics, such as free energy perturbation (FEP)[5] and thermodynamic integration (TI),[6] can be applied to calculate free energies of binding. Both FEP and TI are based on molecular dynamics (MD) or Monte Carlo (MC) simulations and are used to calculate the relative free energy difference between two states. This involves lengthy simulations in order to sample all relevant conformations of the system in both end states and several (possibly nonphysical) intermediate states. Depending on the system, the number of intermediate states which are required can be quite large. Assuming that the simulations are long enough so that all relevant conformations have been sampled, both TI and FEP are rigorous methods, and high accuracy can be obtained. However, they quickly become too computationally demanding for larger numbers of ligands.

The linear response approximation (LRA)[7,8] and the linear interaction energy (LIE)[9] methods are more efficient approaches, requiring only simulations of the end states, without the need of sampling numerous (nonphysical) intermediate states. These semiempirical methods assume linear responses of both the polar and nonpolar free energy contributions. The binding free energy is then estimated based on the ensemble averages of the electrostatic and van der Waals contributions to the potential energy in simulations of the ligand free in solution and when bound to the protein. LRA includes the electrostatic contribution of the potential energy, calculated over a simulation in which all partial charges of the ligand are set to zero. This contribution, the so-called electrostatic preorganization, is assumed to be zero in LIE, abolishing the need for such a simulation of a neutralized ligand. Although this assumption is justified in the case of a simulation in water, it does not necessarily hold in case of protein simulations, as the protein is likely to have a nonzero net dipole in the active site.[10,11] Changing the theoretical scaling factor of 1/2 for the electrostatic interactions may in part compensate for this assumption.[12−14] While the linearity in the polar contribution in LIE has a reasonable theoretical background, the linearity of the nonpolar contribution comes from empirical observation.[12] The empirical parameter used to scale these contributions is, therefore, also used as a parameter to fit data to experimental values.[11,12] In order to reproduce experimental data, sometimes an additional empirical constant is used.[13]

Another efficient method to calculate binding free energies is the one-step perturbation.[15] This approach requires only the simulation of a single reference state in solution and bound to the protein. The reference state is defined such that it samples all the relevant conformations for all the ligands for which the free energy of binding is to be calculated. The reference state does not have to be a physical molecule and is often conveniently designed with soft-core interactions,[16] which allow for more extensive sampling. A disadvantage of this approach is that the ligands must be very similar to each other in other to be able to find an appropriate reference molecule which samples the relevant conformations. For groups of ligands with large variations in the partial charges, the choice of the

Special Issue: Wilfred F. van Gunsteren Festschrift

Received: October 25, 2011
Published: January 23, 2012

reference molecule becomes especially difficult, since a neutral reference state will most likely not sample the appropriate configurations for ligands with opposing dipoles.[17]

Because of the above observations for the different approaches to calculate free energy differences, we previously suggested to combine the strengths of LIE and the one-step perturbation.[18] Thus, the polar contributions from LIE are combined with the nonpolar contributions determined with the one-step perturbation. In addition, the electrostatic preorganization energy may also be determined from the simulation of the reference molecule. Combining LIE and one-step perturbation in this way eliminates the use of empirical parameters and does not lead to insufficient sampling due to inappropriate charges on the reference molecule. Details of this approach will be discussed in the Theory Section.

So far, we applied the combined approach to a model host–guest system only. Here we calculate the relative free energies of binding for eight benzamidine-like compounds binding to the serine protease trypsin. The ligands are shown in Figure 1. Trypsin is a well studied protein and has a reasonable size, and the benzamidines carry a full positive charge, making it a more challenging model system for this study.[19−26] Results from rigorous TI calculations will be used as reference values. This way errors due to the force field are eliminated, and the difference in efficiency and accuracy between the TI and the combined approaches can be evaluated.

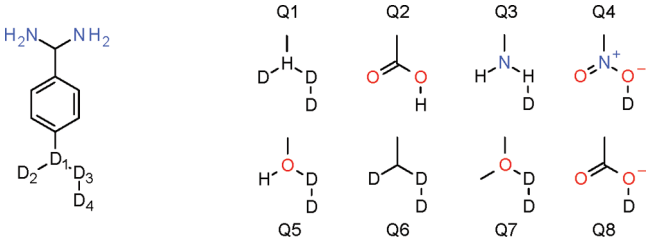

Figure 1. Molecular structure of the benzamidines, for which the relative binding free energies are calculated. The atoms D_1–D_4 in the left structure are being replaced by the atom types in the groups on the right, and D represents a noninteracting dummy atom.

THEORY

Thermodynamic Cycles. Calculating free energy differences is the easiest when the conformational spaces that are accessible to the two states are as similar as possible. Calculation of the absolute free energy of binding from simulations showing spontaneous complexation is thus difficult, although not impossible as recently shown by Buch et al. for the benzamidine–trypsin complex.[27] However, since free energy is a state function, one can use a thermodynamic cycle to calculate relative free energies of binding.[28] As the value around the cycle should be zero, the relative free energy of binding is equal to the difference in perturbing one ligand to the other, when bound to the protein and when free in solution.[28] When the two ligands are similar, the conformational spaces that need to be sampled are similar for the two ligands, making the calculation easier. Several thermodynamic cycles are used in this study to evaluate the results, the most complete one is shown in Figure 2.

Thermodynamic Integration (TI). TI is a method to calculate the free energy difference between two states, making use of multiple intermediate states as defined by a coupling parameter λ. The Hamiltonian $H(\lambda)$ connects the Hamiltonians H_A and H_B representing the end states A and B, respectively, such that the system is in state A when λ is equal to zero and in state B when λ is equal to 1. The free energy difference can then be calculated from the derivative of the free energy with respect to λ. It can be shown that this derivative equals the ensemble average of the derivative of the Hamiltonian with respect to λ from the simulation at a given λ value (eq 1).[6]

$$\Delta G_{BA} = G_B - G_A = \int_0^1 \frac{dG(\lambda)}{d\lambda} d\lambda = \int_0^1 \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_\lambda d\lambda \quad (1)$$

Here, ΔG_{BA} is the Gibbs free energy difference between state A and B, and $\langle\rangle_\lambda$ indicates an ensemble average obtained at λ. The numerical integration of these ensemble averages is performed by applying e.g. Simpson's rule.[29]

Since free energy is a state function, the relative free energy of binding between ligand 1 and 2 can be calculated from the difference of the free energy of mutating ligand 1 to 2 in water and when bound to the protein. Further thermodynamic cycles are

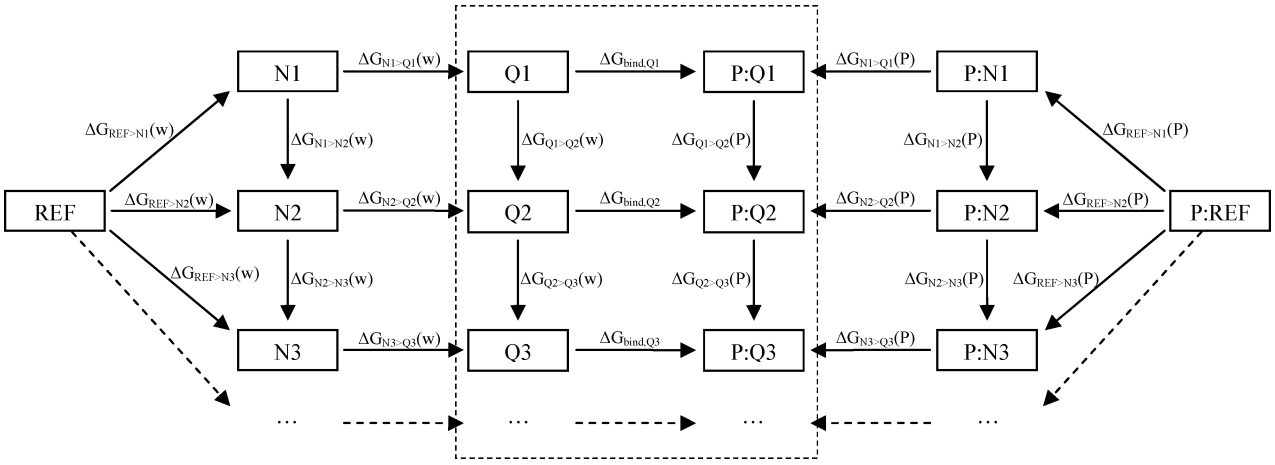

Figure 2. Thermodynamic cycle. The part in the dashed central box may be used to calculate the relative free energies of binding for compounds Q1, Q2, Q3, etc. to a protein P from the free energy differences between the compounds, bound to the protein and free in water (w). Alternatively, these free energy differences may be estimated from combined methods using the intermediate compounds N1, N2, etc. in which all partial charges were set to zero and a possibly nonphysical reference state REF.

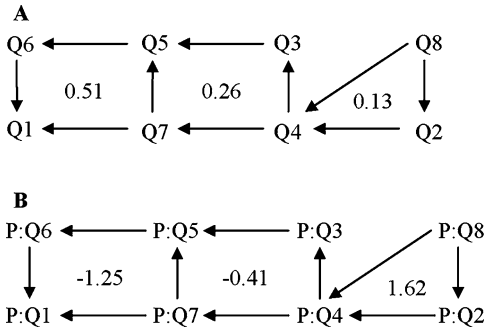

Figure 3. TI cycle closures in water (A) and bound to trypsin (B) for the normal, charged benzamidines (in kJ/mol).

designed, connecting the free energies of binding of 3−4 ligands, eventually connecting all ligands to each other (see Figure 3). The sum of the free energies of binding in such a cycle should yield zero, offering a convenient way to evaluate the performance of the TI calculations. One should however be careful with the design of the cycles. The cycles will only close if the end state of one TI is exactly the same as the first state of the next TI. Therefore, each atom that is removed during a perturbation should be kept as a dummy atom in the next perturbation. Also, the masses of the dummy atoms should be preserved or perturbed during the TI to match the next leg. This is necessary because we are using the SHAKE algorithm to constrain bond lengths, which makes the kinetic energy dependent on the positions of the atoms.[30] A second reason is that we are not only evaluating free energies of binding but also the thermodynamic cycles as indicated in Figure 3. The dummy atoms in these thermodynamic cycles do not have the same mass in the legs of the cycle, and thus these contributions do not cancel.[31]

Linear Response Approximation (LRA) and Linear Interaction Energy (LIE). The LRA and LIE methods only require the simulation of the ligands in water and the ligands bound to the protein. The electrostatic contribution is estimated as the difference between charging a neutral ligand in water and when bound to a protein. This is evaluated based on the difference in electrostatic energy of the ligand in the neutral and the charged states, both in water and bound to the protein. In LRA the neutral state of the ligand is explicitly simulated, and the electrostatic interaction energies for the charged state are calculated over the resulting ensemble, yielding the electrostatic preorganization energy.[7,8,32] The electrostatic contribution to the free energies is then calculated as

$$\Delta G^{\mathrm{LRA}}_{N>Q} = G_Q - G_N = \frac{1}{2}[\langle H_Q - H_N\rangle_N + \langle H_Q - H_N\rangle_Q] = \frac{1}{2}(\langle V^{\mathrm{el}}_{\mathrm{ls}}\rangle_N + \langle V^{\mathrm{el}}_{\mathrm{ls}}\rangle_Q) \quad (2)$$

Here, $\Delta G^{\mathrm{LRA}}_{N>Q}$ is the free energy difference of charging a ligand using LRA. H_Q and H_N indicate the Hamiltonians from the charged and neutral states, respectively, $\langle\rangle_N$ and $\langle\rangle_Q$ indicate ensemble averages over the neutral and charged states, respectively, and $V^{\mathrm{el}}_{\mathrm{ls}}$ is the electrostatic energy of the ligand with its surroundings. In LIE, however, it is assumed that the electrostatic interactions of the charged ligand obtained from the ensemble of neutral states will average to zero. The electrostatic contributions to the free energy of binding are then directly related to the ensemble average of the electrostatic interaction energy between the ligand and its surroundings obtained from the regular, charged simulations.

$$\Delta G^{\mathrm{LIE}}_{N>Q} = G_Q - Q_N = \frac{1}{2}[\langle H_Q - H_N\rangle_N + \langle H_Q - H_N\rangle_Q] \approx \beta\langle V^{\mathrm{el}}_{\mathrm{ls}}\rangle_Q \quad (3)$$

Here, $\Delta G^{\mathrm{LIE}}_{N>Q}$ is the free energy difference of charging a ligand using LIE. β has a theoretical value of 1/2 but can also be modified in order to better reproduce experimental data.

The nonpolar interactions are also assumed to have a linear relationship with the free energy difference, although the theoretical background is not as strong as for the electrostatic contributions. Equation 4 shows the general formula used in LIE.

$$\Delta G^{\mathrm{LIE}}_{\mathrm{bind}} = \alpha\cdot\Delta\langle V^{\mathrm{vdW}}_{\mathrm{ls}}\rangle_Q + \beta\cdot\Delta\langle V^{\mathrm{el}}_{\mathrm{ls}}\rangle_Q + \gamma \quad (4)$$

Here, β has a theoretical value of 1/2, and α and γ are empirical parameters usually used to fit results to experimental data from a training set. The delta sign, Δ, indicates the difference between the ensemble averages obtained from the simulation of the ligand in solution and when bound to the protein.

Cumulant Expansion and Third Power Fitting. As will be shown in the Results Section, both LIE and LRA did not give satisfactory results in this study. From TI simulations of charging the benzamidine-like compounds (both in solvent and bound to trypsin), it became clear that the linear response approximation applied in LIE and LRA does not necessarily hold. Therefore, we searched for an alternative method to calculate the electrostatic contribution to the free energy of binding, using only simulations from the neutral and charged states, without assuming a linear response. It is straightforward to show that for a linear combination of H_N and H_Q, the derivative of the free energy with respect to λ is defined as in eq 5.

$$\frac{\mathrm{d}G}{\mathrm{d}\lambda} = \left\langle \frac{\partial H}{\partial \lambda}\right\rangle_\lambda = \langle V^{\mathrm{el}}_{\mathrm{ls}}\rangle_\lambda \quad (5)$$

We observed that in many cases this quantity does not correlate linearly with λ, but that it typically shows a slightly curved behavior. A third-order polynomial can be used to describe this curvature, where four constraints are needed to find the best fit. Apart from the $\mathrm{d}G/\mathrm{d}\lambda$ values at the end states, the slope ($d^2G/\mathrm{d}\lambda^2$) in these points is also included. This second derivative of the free energy with respect to λ can be obtained from the simulations at the end points using a cumulant expansion (see, e.g., Chipot and Pohorille).[33] The free energy of charging is then given by eq 6.

$$\Delta G^{\mathrm{TPF}}_{N>Q} = \int_0^1 f(\lambda)\mathrm{d}\lambda \quad (6)$$

Here, $f(\lambda)$ is a third-order polynomial in λ with parameters a, b, c, and d. The parameters are defined by fitting to the four constraints:

$$\left.\frac{\mathrm{d}G}{\mathrm{d}\lambda}\right|_0, \left.\frac{\mathrm{d}G}{\mathrm{d}\lambda}\right|_1, \left.\frac{\mathrm{d}^2G}{\mathrm{d}\lambda^2}\right|_0, \left.\frac{\mathrm{d}^2G}{\mathrm{d}\lambda^2}\right|_1$$

Which are defined by

$$\left.\frac{\mathrm{d}G}{\mathrm{d}\lambda}\right|_\lambda = \left\langle \frac{\partial H}{\partial \lambda}\right\rangle_\lambda = \langle V^{\mathrm{el}}_{\mathrm{ls}}\rangle_\lambda \quad (7)$$

$$\left.\frac{d^2G}{d\lambda^2}\right|_\lambda = \frac{1}{k_BT}\left(\left\langle\frac{\partial H}{\partial\lambda}\right\rangle_\lambda^2 - \left\langle\left(\frac{\partial H}{\partial\lambda}\right)^2\right\rangle_\lambda\right) = \frac{1}{k_BT}(\langle V_{ls}^{el}\rangle_\lambda^2 - \langle(V_{ls}^{el})^2\rangle_\lambda) \quad (8)$$

Here, k_B is Bolzmann's constant, and T is the temperature in K. Note that the second derivative of the free energy with respect to λ is equal to the negative of the fluctuations of the derivative of the Hamiltonian with respect to λ. These four constraint values are thus easily obtained from simulations at the neutral and charged states without additional costs when compared to LRA. We will refer to this method as third power fitting (TPF). In the past, polynomial fits to simulations at multiple λ-values[34] or free energy extrapolations using higher order expansions[15] have been suggested. Note that the method proposed here still relies on end-state simulations only and that higher order expansions are usually very difficult to converge.[15]

One-Step Perturbation (OSP). In the combined approaches, the nonpolar contributions to the free energy of binding will be determined with the OSP method. The general idea behind the OSP method is again based on the fact that free energy is a state function. Because of this, the free energy difference between two states will be independent of the chosen path. The free energy difference between states N1 and N2 can thus also be obtained by going from state N1 to some intermediate state REF and then from REF to N2:

$$\Delta G_{N1>N2} = \Delta G_{N1>REF} + \Delta G_{REF>N2} \quad (9)$$

In the OSP method, REF is an unphysical reference state which is closely related to state N1 and N2 (and usually several other states Nx as well). After simulation of the reference state, the free energy difference with respect to the other states can be calculated with the perturbation formula of Zwanzig:[5]

$$\Delta G_{N1>REF} = -k_BT\cdot\ln\langle e^{-(H_{N1}-H_{REF})/k_BT}\rangle_{REF} \quad (10)$$

This a posteriori calculation of the free energy difference is much faster than the actual simulation, as only the differences in the Hamiltonians of the two states have to be calculated. However, this equation only applies if the simulation of the reference state encompasses the relevant conformations of the real states (N1, N2, etc.), making the choice of the reference state very important. More overlap between the ensemble sampled with the reference state and that of the real states can be achieved by using soft atoms.[16] Soft-core potentials can be used at sites in REF where not all real states have an atom or where the atom types differ between the real states. During the reference simulation these atoms behave normally at larger distances, but the singularity for overlapping atoms is removed. This leads to more extensive sampling, which might resemble the real states better. Here, both the reference state and the real states are neutral so that only the nonpolar contributions to the free energy are evaluated.

Projection of Neutral States From OSP. The simulations that are required to get the relative free energy of binding of the para-substituted benzamidine to trypsin are those of the reference state, the neutral and charged forms of each para-substituted benzamidine, and each of these both when solvated and when bound to trypsin. A further increase in efficiency is possible if the neutral states required for TPF are not explicitly simulated but projected from the reference state. This requires energetic properties of the real states, which cannot be calculated directly from the reference simulation but need to be reweighted to the appropriate ensemble.[35] This was already performed in the combined LIE/OSP method[18] and is equivalent to umbrella sampling with a biasing potential equal to $V^{bias} = H_{REF} - H_{N1}$.[36,37] The properties in eqs 7 and 8 for $\lambda = 0$ can then be calculated from

$$\langle X\rangle_{N1} = \frac{\langle X e^{-(H_{N1}-H_{REF})/k_BT}\rangle_{REF}}{\langle e^{-(H_{N1}-H_{REF})/k_BT}\rangle_{REF}} \quad (11)$$

The quantity X is calculated over the snapshots of a simulation of the reference state and reweighted to obtain the proper ensemble average.

METHODS

The energy minimizations and molecular dynamics simulations were performed using the GROMOS11 simulation package[38] with the 53a6 force field parameter set.[39] Topologies of the different p-substituted benzamidines were created using the structures as shown in Figure 1. Parameters for the individual atoms can be found in the Figures S1 and S2 and Tables S1−S6 in the Supporting Information.

Initial coordinates of trypsin with bound benzamidine were obtained from the crystal structure with PDB entry code 3PTB.[40] In order to create initial bound conformations for the other p-substituted benzamidines, they were simply structurally aligned to benzamidine, since it was obvious that the binding site was large enough to accommodate even the *p*-carboxy-benzamidine. The ligands were solvated in rectangular periodic boxes containing 815 SPC water molecules[41] and one chloride ion to obtain a neutral system. Simulations of the ligands bound to trypsin were performed in rectangular period boxes containing 9492 SPC water molecules and 9 chloride ions.

For the simulations of the benzamidines bound to trypsin, thermalisations and initial equilibrations were performed with the complex of *p*-carboxy-benzamidine (Q2) and trypsin. Initial velocities were generated from a random Maxwell−Boltzmann distribution at 50 K. The complex was then slowly warmed up, by increasing the temperature with steps of 50 K and simulating 20 ps at each temperature. Initially the position restraints on the backbone atoms had a force constant of 2.5×10^4 kJ mol^{-1} nm^{-2}, which was decreased 10-fold with each increase in temperature. The last simulation was performed at 298 K, without position constraints for 40 ps. The ligand simulations in solvent were immediately started at 298 K, again by generating initial random velocities from a Maxwell−Boltzmann distribution.

The solute and solvent were controlled by separate temperature baths, using the weak coupling method with a relaxation time of 0.1 ps.[42] The SHAKE algorithm with a geometric accuracy of 1×10^{-4} was used to constrain bond lengths,[43] which allowed the use of a 2 fs time step. Simulations were performed at constant pressure, realized by isotropic weak coupling with a compressibility of 4.575×10^{-4} (kJ mol^{-1} nm^{-3})$^{-1}$ and a relaxation time of 0.5 ps.[42] Center of mass movement was removed every 1000 steps. A triple range cutoff scheme was used for the calculation of the long-range interactions, and a pairlist was generated every fifth time step. Within a distance of 0.8 nm, every interaction is calculated from the pairlist. The interactions between 0.8 and 1.4 nm were evaluated only with pairlist updates and kept constant during intermediate steps. Beyond 1.4 nm, a reaction field contribution[44] was applied to electrostatic energies and forces using a dielectric constant of 61.[45]

Initial simulations showed that the natively bound calcium ion sometimes drifts away from its original binding position, about 2.2 nm away from the benzamidine binding site. In view of its importance for the structure and function of trypsin,[46] distance restraints with a force constant of 2×10^3 kJ mol^{-1} nm^{-2} were applied on the calcium ion with respect to four neighboring residues in all simulations in order to prevent it from drifting away.

TI calculations for the perturbation of charged ligands in complex with trypsin were performed at 21 evenly spaced λ values, with up to 3 extra λ values in order to create smoother integration curves. At each λ value, an equilibration phase of 100 ps and a production phase of 400 ps were performed. In some cases the production phases were prolonged up to 800 ps in order to reduce error estimates. In the perturbations between benzamidines 2, 4, and 8, some convergence problems occurred. These issues were solved by performing the simulations in the reverse direction and redoing a simulation starting from one of the obtained structures. In the end the same phase space was sampled in all simulations. A softness parameter of 0.5 for the van der Waals interactions was applied to atoms for which the atom type was perturbed. Atoms undergoing changes in their charges had a softness parameter of 0.5 nm^2 for the electrostatic interactions.

In order to evaluate the performance of OSP, LIE, LRA, and TPF separately, TI calculations of perturbing neutral ligands and charging of ligands were also performed. Thus, for comparison with the methods which calculate the electrostatic contributions to the free energy difference, TI calculations were performed where the ligands are perturbed from their neutral state to the charged state. The neutral state is defined here as having zero partial charges on the atoms which are changing between different ligands (atoms 17−21, see Supporting Information), the other atoms of the benzamidines still have their normal charges. This setup was chosen to prevent the charging free energy from being predominated by the full charge change of the benzamidine moiety, which is the same in all compounds. The LIE method has successfully been applied to parts of interacting molecules before.[24,25,47] Simulations were performed at five evenly spaced λ values, without the use of soft-core potentials to ensure that eq 7 holds. The systems are equilibrated for 200 ps at $\lambda = 0$ and for 100 ps at other λ-values. The production runs are 1 ns for the end states and 400 ps for the intermediate λ-values. TI calculations for the perturbation of the neutral ligands were performed with 21 evenly spaced λ values, with the same simulation times as for the charging TI calculations. The softness parameter of 0.5 for the van der Waals interactions was again applied for the atoms undergoing perturbation.

OSP. The parameters of the reference state for the OSP calculations can be found in the Supporting Information. It was designed based on the largest atoms, longest bond lengths, and largest angles of the eight benzamidines. The benzamidine part which is identical for all ligands was kept as normal (with partial charges), whereas the para-substituted atoms were neutral and were described by a soft-core potential with a softness parameter for the van der Waals interactions of 1.8. This high softness parameter allows for relevant sampling for even the smallest ligand, as the soft atoms can be penetrated by surrounding atoms. Coordinates and velocities were written out every 0.2 ps for later evaluation. The free energy difference between the reference state and a real, neutral ligand could then be calculated by applying eq 10. Prior to the energy evaluation of the Nx states, bond lengths were reset to the appropriate values (Table S4, Supporting Information). As bond lengths were constrained, this does not lead to an additional energy term in the bonds,[48] but the value of the improper dihedral angle may change slightly. The differences between the Hamiltonians consist of van der Waals, improper dihedral, angle, and kinetic energy contributions because the masses of the atoms are changing. The reference state in water was simulated for 10 ns in water and for 5 ns when bound to trypsin.

Calculation of the Electrostatic Contributions. For LIE, the normal, charged states of the ligands needed to be simulated both in solution and when bound to trypsin. Simulation conditions were the same as for the TI calculations for the free energy difference between the normal, charged ligands. Each ligand was simulated for 2 ns in water and 1.6 ns when bound to trypsin. The β value was determined by least-squares fitting to the TI reference data.

For the LRA and TPF methods, all ligands were additionally simulated for 1 ns in solution and when bound to trypsin in their neutral states. For the calculation of the ensemble averages of the neutral states from the OSP simulations, eq 11 was applied.

RESULTS

The TI reference data for the normal, charged perturbations in solution and in complex with trypsin were evaluated based on the closure of the smaller thermodynamic cycles, as shown in Figure 3. The largest deviations from cycle closure are obtained for the cycle including Q8 when bound to trypsin. This is not unexpected, since this ligand contains an extra full charge, which causes larger absolute errors in the electrostatic energy contributions during the TI calculations. Note that because all simulations are performed using the same electrostatic methodology and boundary conditions, no additional corrections to the free energy[49] were required in order to assess the accuracy of the various methods. With a maximum cycle closure error of −1.6 kJ/mol reliable reference data is obtained. The cycle closures for the TI of the neutral perturbations in solution are comparable and are given in Table S7 in the Supporting Information. The integration of the TI data was performed using Simpsons rule. The most commonly applied numerical integration method in TI calculations would be the trapezoidal rule; however, it was pointed out by Bruckner and Boresch that Simpsons rule performs better.[29] Both methods were applied, and it was shown that by applying the Simpsons rule, cycle closures were indeed better. The cycle closures using the trapezoidal rule can be found in Table S8 in the Supporting Information.

The overall aim of this work is to calculate the relative free energies of binding, but we first separate this problem into smaller parts in order to evaluate the performance of each of the methods. As discussed in the Theory and Methods Sections, several methods were used to calculate the electrostatic contributions to the free energy of binding. First of all, reference data was obtained from TI calculations. These results, plus the results from LIE, LRA, and TPF, are shown in Table 1 for the simulations in water and Table 2 for the simulations of the ligand−protein complexes. The differences of the methods with respect to the TI reference data are also shown as well as the overall root-mean-square deviations (rmsd) for each method. It immediately becomes clear that LIE is not performing well with a rmsd of >27 kJ/mol. The LRA method is doing better with rmsd's of around 8 and 10 kJ/mol for the water and complex simulations, respectively. In the current case, this improvement seems to be due to the fact that we only

Table 1. Charging Free Energies of Benzamidines in Water[a]

water	TI	LIE	diff	LRA	diff	TPF	diff
N1→Q1	22.8	10.9	11.9	22.8	0.0	22.8	0.0
N2→Q2	−111.1	−84.0	−27.0	−119.5	8.5	−114.7	3.6
N3→Q3	−111.6	−71.5	−40.1	−118.2	6.6	−112.0	0.4
N4→Q4	89.8	39.5	50.3	89.1	0.7	90.0	−0.2
N5→Q5	−105.4	−76.2	−29.1	−115.5	10.2	−108.9	3.5
N6→Q6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
N7→Q7	−54.2	−31.4	−22.8	−55.8	1.6	−54.5	0.3
N8→Q8	−336.8	−317.2	−19.7	−353.4	16.6	−339.3	2.5
rmsd			29.1		7.9		2.0

[a]Charging free energies in kJ/mol, obtained with TI, LIE, LRA and TPF. For LIE and LRA, $\beta = 1/2$. Differences with respect to TI are indicated in the columns 'diff', and rmsd's with respect to TI are shown in the last row.

Table 2. Charging Free Energies of Benzamidines Bound to Trypsin[a]

trypsin	TI	LIE	diff	LRA	diff	TPF	diff
N1→Q1	22.9	11.2	11.7	23.3	−0.4	23.3	−0.4
N2→Q2	−115.2	−82.7	−32.5	−117.1	1.9	−116.8	1.6
N3→Q3	−109.2	−77.4	−31.9	−121.2	12.0	−112.6	3.3
N4→Q4	87.6	39.3	48.3	88.3	−0.7	88.8	−1.2
N5→Q5	−106.2	−77.4	−28.9	−117.3	11.1	−110.7	4.4
N6→Q6	0.0	0.0	0.0	0	0.0	0.0	0.0
N7→Q7	−53.4	−30.6	−22.8	−54.8	1.4	−53.4	0.0
N8→Q8	−291.8	−297.5	5.7	−314.5	22.8	−296.8	5.0
rmsd			27.2		9.9		2.7

[a]Charging free energies in kJ/mol, obtained with TI, LIE, LRA and TPF. For LIE and LRA, $\beta = 1/2$. Differences with respect to TI are indicated in the columns 'diff', and rmsd's with respect to TI are shown in the last row.

include the atoms that are being modified in the estimates of $\langle V_{ls}^{el} \rangle$, while the benzamidine part of the molecules is considered part of the surroundings. The contribution of the neutral states (ignored in LIE) is non-negligible in this case. However, the TPF method is clearly the best method, with rmsd's of around 2 and 2.7 kJ/mol. In the water simulations, the TPF method always matches the TI data better than LIE or LRA. When the ligands are bound to trypsin, LRA outperforms TPF only for the N4→Q4 transition, while both methods agree very well with the TI data. The various results are illustrated for the charging of *p*-amino-benzamidine (N3→Q3) in water in Figure 4. From this figure it becomes clear that we cannot assume that the electrostatic preorganization energy does not play a role; including this with LRA makes a significant difference. It is also obvious that the linear response approximation used in both LIE and LRA does not exactly hold in this case. In order to determine if the curvature observed in this graph may be due to the fact that only half the molecule is considered for the charging calculations, we have performed the same calculation for a hypothetical molecule in which all remaining partial charges (for atoms 1−16, in Figure S1, Supporting Information) were set to zero (Figure 4b). It is clear from this figure that the preorganization energy largely disappears but the curvature remains. As TPF shows the best match to the reference data, this method was chosen to calculate the electrostatic contributions to the free energy of binding.

For the nonpolar van der Waals contributions, we compare the results from OSP to the reference data by considering the difference between states N1 and N2 as obtained from eq 9, with direct estimates, obtained with TI. For the simulations in water, these results are shown in Table 3. With a rmsd of 1.9, the OSP simulations nicely match the reference data. For the bound states, only three of the lengthy TI simulations were performed, with a rmsd of 1.8 kJ/mol.

Separately, the OSP and TPF methods seem to work properly. Now, these methods are combined in order to determine the free energy of perturbing a normal, charged ligand to another, e.g., Q1→Q2 in Figure 2. As this is done separately for the simulations in water and for the ones with trypsin, the performance of the combined OSP and TPF can be evaluated separately for the ligands in water and when bound to the protein as indicated in Table 4. In the water cycle, there is a maximum deviation of around 5 kJ/mol, with a rmsd of around 3 kJ/mol. The deviations in the trypsin cycle are a bit larger. The maximum deviation in this cycle is around 10 kJ/mol, with an overall rmsd of 4.5 kJ/mol. The remaining columns in Table 4 indicate the results for the relative binding free energies, which correspond to the differences between the trypsin and water values and are equal to the actual relative free energies of binding. Here, the largest deviation is around 6 kJ/mol, and the overall rmsd is around 3.9 kJ/mol, indicating that some errors made in the water and trypsin cycle are compensating. Eight of the 10 relative free energy differences that were calculated show the correct sign, leading to a Kendall τ value of 0.6. These results are very reasonable, especially considering the increase in efficiency with respect to the TI simulations. Including equilibrations, a total simulation time of around 245 ns is required for the TI simulations, whereas this is only 52 ns for the TPF/OSP method.

So far, the results of the projection of the energy contributions of the neutral state from the OSP simulation have not been included yet. Table 5 shows the comparison of the cycle closures using a separate neutral simulation or the projection from the OSP simulation, for the water, trypsin and binding free energy cycle. Although there are some differences in the cycle closures in the water cycle, there is no significant change in the overall rmsd of the water cycle. In the trypsin cycle, the increase in the overall rmsd by using the projection is

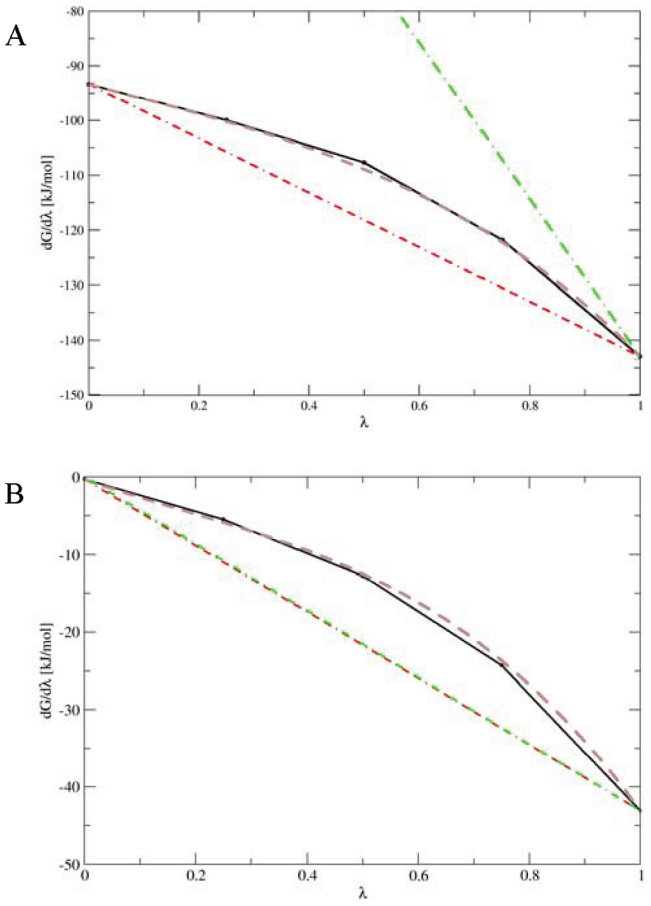

Figure 4. Comparison of TI, LIE, LRA, and TPF for charging BZA3 in water. The solid black line represents the TI data, the green (stripe–dot) line represents the LIE results using β = 1/2, the red (stripe–stripe–dot) line represents the LRA results, and the brown (striped) line represents the TPF results. Panels A and B show the results for the calculations with a charged and a neutral benzamidine moiety, respectively.

Table 3. Comparison of OSP with TI Calculations Connecting the Neutral States of the Ligands in Water and When Bound to the Protein[a]

	water			trypsin		
	TI	OSP	diff.	TI	OSP	diff.
N2→N4	3.7	3.0	0.7	2.7	−0.3	3.0
N2→N8	4.6	4.3	0.3			
N3→N5	−5.1	−3.5	−1.7			
N4→N3	14.9	16.6	−1.7			
N4→N7	11.3	15.9	−4.6			
N4→N8	1.0	1.3	−0.3			
N5→N6	0.8	2.1	−1.3			
N6→N1	6.7	8.6	−2.0	6.2	6.6	−0.4
N7→N1	5.9	8.0	−2.1	5.1	5.9	−0.8
N7→N5	−1.8	−2.8	0.9			
rmsd			1.9			1.8

[a]Differences between the methods are indicated in the columns labeled with 'diff', and values are in kJ mol^{-1}.

also very small, around 0.8 kJ/mol. The rmsd of the binding free energy cycle shows an increase of around 1.7 kJ/mol up to 5.6 kJ/mol. The total required simulation time has now decreased to around 33 ns, because the simulations of the neutral ligands are no longer necessary.

DISCUSSION

The combined TPF/OSP method as described above is significantly more efficient than the frequently used TI approach. Using TPF/OSP including a neutral simulation or by projecting the energy contributions from the OSP simulation, the efficiency increases with a factor of more than 4.5 and almost 7.5, respectively. Obviously, these factors depend on the number of ligands which are investigated, since no extra simulations are needed for additional ligands in the OSP part of the method. It might be necessary to simulate the reference state longer however, in order to still sample the relevant states for all ligands. This also strongly depends on how versatile the set of ligands is. The price that one pays for the increase in efficiency is a slight loss of accuracy, which is below 4 kJ/mol (<1 kcal/mol) for TPF/OSP when the neutral state is explicitly simulated. When its contributions are projected from the OSP simulation, the rmsd increases to 5.6 kJ/mol, which is still very small considering the factor 7.5 efficiency increase. Note that a difference of 5.6 kJ/mol in binding affinity corresponds to a factor 10 difference in the dissociation constant K_D and that a difference below 4 kJ/mol is within chemical accuracy. For a discussion about the definition of chemical accuracy, the reader is referred to Shirts et al.[50] For comparison, using linear response to calculate the free energy of (de)protonation of aspartic acid in aqueous solution resulted in a rmsd of 27 kJ/mol with respect to TI data and 16 kJ/mol when only considering single protonation site changes.[51]

We emphasize again that the goal was not to reproduce experimental data, but to eliminate force field inaccuracies by comparing the results to the TI data. Having said this, the comparison of the TI and TPF/OSP results with respect to the experimental data will now be briefly discussed (see Table 6). Unfortunately no experimental data is available for Q8, for all other ligands absolute binding free energies are given in ref 52. The difference in relative free energy of binding between TI and experiment (excluding the ones with Q8) has a rmsd value of 6.9 kJ/mol. Strikingly, comparing the TPF/OSP results with the experimental values leads to a rmsd of only 4.1 kJ/mol. By using the absolute free energy of binding of the reference state as an empirical parameter, this method also allows for the calculation of absolute free energies for the real compounds.[53] This procedure leads to a rmsd of only 2.8 kJ/mol (Table 7). This agreement with experiment is slightly better than, e.g., described in the study of Leiros et al. in which LIE was applied to this system previously.[26] The large deviations of the (relative) free energies of binding including Q4 with respect to experiment are likely to be caused by wrong partial charges on the nitro group, as these parameters are not so well validated in this force field. Furthermore, a polarizable force field might be required in order to further improve the agreement to experiment.[19]

Apart from just the efficiency of the method, the advantage of TPF/OSP is that more time is spent on simulating physical states, rather than on unphysical states. Simulations of physical states are of course more useful as they can give insight into the dynamical behavior of the ligands and several properties can be analyzed a posteriori. When the contributions from the N-states are calculated from the projection of the OSP simulation, the unphysical simulations are restricted to a single reference state in water and in complex with trypsin. The other simulations are the physical, charged states of the ligands in water and in complex with trypsin. In the TI simulations only the end states

Table 4. Free Energies of Changing Q*a*→Q*b*, in Water and Bound to Trypsin, Calculated With TI and the Combined TPF/OSP Approach[a]

	water			trypsin			binding free energy		
	TI	TPF/OSP	diff.	TI	TPF/OSP	diff.	TI	TPF/OSP	diff.
Q2→Q4	204.4	207.7	3.3	208.5	205.3	−3.1	4.1	−2.3	−6.4
Q2→Q8	−221.5	−220.4	1.1	−176.1	−177.1	−0.9	45.4	43.3	−2.1
Q3→Q5	1.4	−0.3	−1.7	−0.3	−1.7	−1.3	−1.7	−1.3	0.4
Q4→Q3	−186.1	−185.4	0.7	−179.4	−185.0	−5.5	6.7	0.4	−6.3
Q4→Q7	−132.7	−128.6	4.2	−124.4	−125.0	−0.6	8.3	3.5	−4.8
Q4→Q8	−425.8	−428.0	−2.3	−383.0	−382.4	0.6	42.8	45.7	2.9
Q5→Q6	106.1	110.9	4.8	104.0	114.3	10.2	−2.1	3.3	5.4
Q6→Q1	29.5	31.4	1.9	29.4	29.9	0.5	−0.1	−1.5	−1.4
Q7→Q1	82.9	85.3	2.4	79.8	82.6	2.8	−3.1	−2.7	0.4
Q7→Q5	−52.2	−57.1	−4.9	−54.9	−61.6	−6.7	−2.7	−4.5	−1.8
rmsd			3.1			4.5			3.9

[a]In the final columns (binding free energy), the free energy difference between the bound and free states equal to the relative binding free energies of the compounds is shown. The differences between TI and the TPF/OSP approach are given in the columns labeled with 'diff'. Values are in kJ mol^{-1}.

Table 5. Comparison of the Closure of the Water, Trypsin and Binding Free Energy Cycles Obtained with the Combined TPF/OSP Approach[a]

	water		trypsin		binding free energy	
	neutral	OSP	neutral	OSP	neutral	OSP
Q2→Q4	−3.3	−3.2	3.1	5.8	−6.4	−9.0
Q2→Q8	−1.1	−0.3	0.9	1.5	−2.1	−1.8
Q3→Q5	1.7	2.3	1.3	−3.4	0.4	5.7
Q4→Q3	−0.7	−2.0	5.5	8.5	−6.3	−10.5
Q4→Q7	−4.2	−4.0	0.6	−1.0	−4.8	−3.0
Q4→Q8	2.3	3.0	−0.6	−2.7	2.9	5.7
Q5→Q6	−4.8	−4.0	−10.2	−9.4	5.4	5.7
Q6→Q1	−1.9	−2.3	−0.5	−1.5	−1.4	−0.8
Q7→Q1	−2.4	−2.9	−2.8	−3.5	0.4	0.7
Q7→Q5	4.9	3.9	6.7	6.5	−1.8	−2.6
rmsd	3.1	3.0	4.5	5.3	3.9	5.6

[a]The data reported in Table 4, obtained using explicit simulations of the neutral state (neutral) is compared to the corresponding cycle closure obtained from estimating the quantities for the neutral state from OSP using eq 11. Values are in kJ mol^{-1}.

Table 6. Comparison of the Relative Free Energies of Binding Obtained with TI and TPF/OSP with Respect to Experimental Results (in kJ/mol)[a]

	exptl.	TI	diff.	TPF/OSP	diff.
Q2→Q4	0.3	4.1	−3.8	−2.3	2.6
Q2→Q8	n.a.	45.5	n.a.	43.3	n.a.
Q3→Q5	1.3	−1.7	3.0	−1.3	2.6
Q4→Q3	−8.0	6.7	−14.6	0.4	−8.4
Q4→Q7	−3.2	8.3	−11.5	3.5	−6.7
Q4→Q8	n.a.	42.8	n.a.	45.7	n.a.
Q5→Q6	1.6	−2.1	3.6	3.3	−1.8
Q6→Q1	−1.1	−0.1	−1.0	−1.5	0.4
Q7→Q1	−3.2	−3.1	−0.1	−2.7	−0.5
Q7→Q5	−3.6	−2.7	−0.9	−4.5	0.9
rmsd			6.9		4.1

[a]In TPF/OSP, the preorganization energy contribution is calculated from the neutral simulations. Experimental results obtained from Guia et al. (1975), and n.a. indicates not available.

are physical, whereas all the intermediate states are unphysical. Here, at least 19 intermediate, unphysical states were required per perturbation.

Table 7. Comparison of the Absolute Free Energies of Binding Obtained with TPF/OSP with Respect to Experimental Results (in kJ/mol)[a]

	exptl.	TPF/OSP	diff.
Q1	−26.5	−24.7	−1.7
Q2	−20.5	−23.3	2.8
Q3	−28.1	−25.2	−2.9
Q4	−20.2	−25.6	5.5
Q5	−26.9	−26.5	−0.3
Q6	−25.3	−23.2	−2.1
Q7	−23.3	−22.1	−1.2
rmsd			2.8

[a]Absolute free energies are calculated by introducing an estimate of the binding energy for the reference structure.

Especially simulations of unphysical states, such as the reference state for the OSP simulation, need to be monitored in order to make sure that the configurations that are sampled are relevant. Including irrelevant configurations (e.g., partial unfolding of the protein or ligand moving out of the active site) in the post-analysis of the OSP can lead to free energies that no longer correspond to the bound state. For this reason the simulation of the reference state bound to trypsin was not prolonged for more than 5 ns. Occasionally also a TI simulation had to be repeated as the ligand was turning out of the active site. Analysis showed that SER195, which is part of the catalytic triad, is very important for hydrogen-bond interactions with the benzamidine-like ligands. The presence or absence of this hydrogen bond can determine if the simulations are converging or not.

One of the largest observed errors in Table 4 is associated with Q5 (*p*-hydroxy-benzamidine) bound to trypsin. It seems to be caused by the polar and nonpolar interactions (Tables 2 and 4). Q5 is a difficult ligand because of its hydroxyl group, which makes it a perfect candidate for hydrogen-bond interactions. It forms stable hydrogen bonds with Gln174 and Ser177 in the individual TI simulations, but transitions between the two partners have not been observed.

The nonpolar interactions may not be optimally predicted since there are relatively few configurations of the reference state which contribute to N5. In general, the smaller the ligand, the fewer relevant configurations are sampled because higher energy is required for surrounding atoms to penetrate the soft

dx.doi.org/10.1021/ct200750p | J. Chem. TheoryComput. 2012, 8, 3686−3695

atoms. A preliminary simulation of the reference state might be required to determine the optimal softness parameter in order to sample relevant configurations for all ligands.

CONCLUSION

Relative free energies of binding have been calculated for para-substituted benzamidines in complex with trypsin using several combinations of methods. Various methods and simulations are used for the calculation of the polar and apolar contributions to the free energy of binding. For the polar contributions, the accuracy and efficiency of LIE, LRA, and TPF have been evaluated. Clear advantages of TPF, with respect to the other two, are the much higher accuracy and the fact that it does not require empirical parameters. For the apolar contributions to the free energy of binding, the OSP approach has been applied. Contributions from the neutral state of the ligand can be determined from separate simulations of (partially) neutralized ligands or can be calculated from a projection from the simulation of the reference state used in the OSP. The latter leads to a slight loss in accuracy but a large gain in efficiency. With respect to TI, the TPF/OSP approach is up to 7.5 times more efficient, and moreover, a much larger part of the simulation time is spent on simulation of physical, rather than unphysical states. It was shown that the TPF/OSP combination is an efficient, yet still reasonably accurate approach for the calculation of relative free energies of binding of a protein ligand system. As it is most efficient for a set of related ligands, it has the potential of being very valuable for lead optimization studies.

ASSOCIATED CONTENT

Supporting Information

Force field parameters for the ligands, TI cycle closures for the neutral simulations, and comparison of different integration methods. This information is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Email: chris.oostenbrink@boku.ac.at.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to thank Wilfred van Gunsteren for his continuing support and interest in free energy calculations. We congratulate him with his 65th birthday and look forward to many fruitful future collaborations. This work was funded by the European Research Council (ERC grant 260408) and the Vienna Science and Technology Fund (WWTF), through project LS08-QM03. The authors thank Andrianopsyah Mas Jaya Putra for preliminary simulations.

REFERENCES

(1) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. *Nat. Rev. Drug Discovery* **2004**, *3*, 935−949.
(2) Wang, R.; Lu, Y.; Wang, S. *J. Med. Chem.* **2003**, *46*, 2287−2303.
(3) Perola, E.; Walters, W. P.; Charifson, P. S. *Proteins* **2004**, *56*, 235−249.
(4) Warren, G. L.; Andrews, C. W.; Capelli, A.-M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. *J. Med. Chem.* **2006**, *49*, 5912−5931.
(5) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.
(6) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300.
(7) Lee, F. S.; Chu, Z.-T.; Bolger, M. B.; Warshel, A. *Protein Eng.* **1992**, *5*, 215−228.
(8) Sham, Y. Y.; Chu, Z. T.; Tao, H.; Warshel, A. *Proteins* **2000**, *39*, 393−407.
(9) Åqvist, J.; Medina, C.; Samuelsson, J.-E. *Protein Eng.* **1994**, *7*, 385−391.
(10) Åqvist, J.; Hansson, T. *J. Phys. Chem.* **1996**, *100*, 9512−9521.
(11) Hansson, T.; Marelius, J.; Åqvist, J. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 27−35.
(12) Marelius, J.; Hansson, T.; Åqvist, J. *Int. J. Quantum Chem.* **1998**, *69*, 77−88.
(13) Åqvist, J.; Luzhkov, V. B.; Brandsdal, B. O. *Acc. Chem. Res.* **2002**, *35*, 358−365.
(14) Almlöf, M.; Carlsson, J.; Åqvist, J. *J. Chem. Theory Comput.* **2007**, *3*, 2162−2175.
(15) Mark, A. E.; Xu, Y.; Liu, H.; van Gunsteren, W. F. *Acta Biochim. Pol.* **1995**, *42*, 525−535.
(16) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529−539.
(17) Oostenbrink, C.; Van Gunsteren, W. F. *J. Comput. Chem.* **2003**, *24*, 1730−1739.
(18) Oostenbrink, C. *J. Comput. Chem.* **2009**, *30*, 212−221.
(19) Jiao, D.; Golubkov, P. A.; Darden, T. A.; Ren, P. *P. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6290−6295.
(20) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395−2417.
(21) Essex, J. W.; Severance, D. L.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **1997**, *101*, 9663−9669.
(22) Åqvist, J. *J. Comput. Chem.* **1996**, *17*, 1587−1597.
(23) Wang, W.; Wang, J.; Kollman, P. A. *Proteins* **1999**, *34*, 395−402.
(24) Brandsdal, B. O.; Smalås, A. O.; Åqvist, J. *FEBS Lett.* **2001**, *499*, 171−175.
(25) Brandsdal, B. O.; Åqvist, J.; Smalås, A. O. *Protein Sci.* **2001**, *10*, 1584−1595.
(26) Leiros, H. S.; Brandsdal, B. O.; Andersen, O. A.; Os, V.; Leiros, I.; Helland, R.; Otlewski, J.; Willassen, N. P.; Smalås, A. O. *Protein Sci.* **2004**, *13*, 1056−1070.
(27) Buch, I.; Giorgino, T.; De Fabritiis, G. *P. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 10184−10189.
(28) Tembe, B. L.; Mc Cammon, J. A. *Comput. Chem.* **1984**, *8*, 281−283.
(29) Bruckner, S.; Boresch, S. *J. Comput. Chem.* **2011**, *32*, 1320−1333.
(30) Van Gunsteren, W. F. Methods for calculation of free energies and binding constants: Successes and problems. In *Computer Simulation of Biomolecular Systems, Theoretical and Experimental Applications*; Van Gunsteren, W. F., Weiner, P. K., Eds.; Escom Science Publishers: Leiden, The Netherlands, 1989; pp 27−59.
(31) Boresch, S. *Mol. Simul.* **2002**, *28*, 13−37.
(32) Warshel, A.; Russell, S. T. *Q. Rev. Biophys.* **1984**, *17*, 283−422.
(33) *Free Energy Calculations. Theory and Applications in Chemistry and Biology*; Chipot, C., Pohorille, A., Eds.; Springer Verlag: Berlin, New York, 2007.
(34) Hummer, G.; Pratt, L. R.; García, A. E. *J. Phys. Chem.* **1996**, *100*, 1206−1215.
(35) Oostenbrink, C.; van Gunsteren, W. F. *Chem.—Eur. J.* **2005**, *11*, 4340−4348.
(36) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.
(37) König, G.; Bruckner, S.; Boresch, S. *J. Comput. Chem.* **2009**, *30*, 1712−1718.
(38) Schmid, N.; Christ, C. D.; Christen, M.; Eichenberger, A. P.; van Gunsteren, W. F. *Comput. Phys. Commun.* **2012**, *183*, 890−903.
(39) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656−1676.
(40) Marquart, M.; Walter, J.; Deisenhofer, J.; Bode, W.; Huber, R. *Acta Crystallogr., Sect. B: Struct. Sci.* **1983**, *39*, 480−490.
(41) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Reidel: Dordrecht, The Netherlands, 1981; Vol. *11*, pp 331−342.

(42) Berendsen, H. J. C.; Postma, J. P. M; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(43) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(44) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451.
(45) Heinz, T. N.; van Gunsteren, W. F.; Hünenberger, P. H. *J. Chem. Phys.* **2001**, *115*, 1125.
(46) Papaleo, E.; Fantucci, P.; De Gioia, L. *J. Chem. Theory Comput.* **2005**, *1*, 1286−1297.
(47) Almlöf, M.; Andér, M.; Åqvist, J. *Biochemistry* **2011**, *46*, 200−209.
(48) Severance, D. L.; Essex, J. W.; Jorgensen, W. L. *J. Comput. Chem.* **1995**, *16*, 311−327.
(49) Kastenholz, M. A.; Hünenberger, P. H. *J. Chem. Phys.* **2006**, *124*, 224501.
(50) Shirts, M. R.; Mobley, D. L.; Brown, S. P. Free-energy calculations in structure-based drug design. In *Drug Design: Structure- and Ligand-Based Approaches*; Merz, K. M., Ringe, D., Reynolds, C. H., Eds.; Cambridge University Press: New York, 2010; pp 61−86.
(51) Bürgi, R.; Läng, F.; Van Gunsteren, W. *Mol. Simulat.* **2001**, *27*, 215−236.
(52) Mares-Guia, M.; Nelson, D. L.; Rogana, E. *J. Am. Chem. Soc.* **1977**, *99*, 2331−2336.
(53) Oostenbrink, C.; van Gunsteren, W. F. *Proteins* **2004**, *54*, 237−246.