JCTC Journal of Chemical Theory and Computation

Evaluations of the Absolute and Relative Free Energies for Antidepressant Binding to the Amino Acid Membrane Transporter LeuT with Free Energy Simulations

Chunfeng Zhao,[†] David A. Caplan,[‡] and Sergei Yu. Noskov*,[†]

Institute for Biocomplexity and Informatics and Department of Biological Sciences, University of Calgary, 2500 University Drive, BI558, Calgary, AB, Canada T2N 1N4 and Molecular Structure and Function, Hospital for Sick Children and Department of Biochemistry, University of Toronto, Ontario, Canada

Received December 8, 2009

Abstract: The binding of ligands to protein receptors with high affinity and specificity is central to many cellular processes. The quest for the development of computational models capable of accurately evaluating binding affinity remains one of the main goals of modern computational biophysics. In this work, free energy perturbation/molecular dynamics simulations were used to evaluate absolute and relative binding affinity for three different antidepressants to a sodium-dependent membrane transporter, LeuT, a bacterial homologue of human serotonin and dopamine transporters. Dysfunction of these membrane transporters in mammals has been implicated in multiple diseases of the nervous system, including bipolar disorder and depression. Furthermore, these proteins are key targets for antidepressants including fluoxetine (aka Prozac) and tricyclic antidepressants known to block transport activity. In addition to being clinically relevant, this system, where multiple crystal structures are readily available, represents an ideal testing ground for methods used to study the molecular mechanisms of ligand binding to membrane proteins. We discuss possible pitfalls and different levels of approximation required to evaluate binding affinity, such as the dependence of the computed affinities on the strength of constraints and the sensitivity of the computed affinities to the particular partial charges derived from restrained electrostatic potential fitting of quantum mechanics electrostatic potential. Finally, we compare the effects of different constraint schemes on the absolute and relative binding affinities obtained from free energy simulations.

I. Introduction

In last five years we have seen rapid and amazing progress in structural studies of membrane transporters. Several crystal structures for sodium-coupled membrane transporters have been solved at high resolution with and without ligand bound.[1−4] One of the first complexes of a membrane transporter with a bound drug was obtained for the LeuT−andipressant complex. The structure, solved in a fully occluded state, contains bound ions, the transported solute, and one of three tricyclic antidepressants (TCAs). LeuT is a bacterial leucine transporter, belonging to the large family of neurotransmitter sodium symporters (NSS).[1] Transporters of this family are involved in the termination of synaptic transmission through the reuptake of neurotransmitters (including glycine, glutamate, serotonin, dopamine, and many others) from the synapse into the cytoplasm of neurons and glia. Dysfunction of these membrane transporters in mammals has been implicated in multiple diseases of the nervous system.[5] Depression, one of the most prevalent psychiatric disorders, is directly associated with perturbation of serotonergic neurotransmission.[6] Antidepressants including fluoxetine (Prozac) and TCAs are known to bind membrane

* Corresponding author. Telephone: (403) 210 7971. Fax (403) 220 8655. E-mail: snoskov@ucalgary.ca.

[†] University of Calgary.

[‡] University of Toronto.

10.1021/ct9006597
Published on Web 04/28/2010

transporters and block transport activity. Thus, understanding the mechanism of drug binding to these membrane transporters could possibly help the development of new therapeutics for depression. Recently, the crystal structures of LeuT bound to a variety of TCAs (clomipramine, CMI; imipramine, IMI; and desipramine, DSI) have been solved by two groups.[2,3] The TCAs bind in an extracellular-facing vestibule about 11 Å above the bound leucine substrate and the two sodium ions, as shown in the crystal structures.[2] It was demonstrated that they uncompetitively inhibit the transport of the leucine substrate, probably through the stabilization of the extracellular gate in a closed conformation.[2,3] With this in mind, the half-maximal inhibitory concentration (IC_{50}) of these TCAs should strongly correlate to their binding affinities. In dose−response experiments reported by Singh et al.,[2] CMI has an IC_{50} of inhibition of leucine transport of about eightfold lower than IMI. Thermodynamically, this roughly corresponds to a free energy decrease of ∼2 kT (about 1.5 kcal/mol at a temperature of 315 K). It was also shown that DSI is a less potent inhibitor compared to IMI (personal communication from S. Singh). Thus, Singh et al.'s data suggested the following affinity sequence: CMI > IMI > DSI. While ranking of the ligands for their binding affinity is available, it is difficult to measure absolute binding free energies for TCAs for their high nonspecific binding to chromatographic filters during separation, making evaluation of ratios between bound/unbound forms ambiguous. The closest measure to a K_d for TCA binding is the substrate uptake inhibition constant (K_i) measured by radiolabeled uptake experiments, which depends on drug binding efficacy. These values alone do not represent an absolute K_d but will provide an accurate measure for relative potency of these drugs known for their ability to inhibit substrate transport and may be used to rank them accordingly. A recent development from the Javitch group on the use of a scintillation proximity assay (SPA) for measuring the K_d of radiolabeled compounds to detergent-solubilized material may lead to experimental K_d values in the future.[7−9] This technique was recently used to measure [^{3}H]-citalopram binding to the presynaptic neuronal membrane serotonin transporter (SERT, homologous to LeuT), suggesting low μM to nM range of affinities.[3,4] The availability of experimental binding affinities,[2] together with the high-resolution crystal structures, make the TCA/LeuT system a rich platform for the testing and validation of various computational strategies for calculating binding free energies of drug binding.

The equilibrium thermodynamics of protein−ligand association is commonly described by binding affinity or Gibbs free energy (ΔG) and can be measured by a variety of standard biophysical or biochemical techniques, such as biospectroscopy, isothermal titration and differential scanning calorimetric techniques, electrophysiology, etc. However, to further our understanding of the process and its molecular determinants, it is important to obtain the quantitative contribution of different forces governing high affinity and specificity. Accurate prediction of binding affinity may, therefore, facilitate drug and protein design and optimization practices to attain better drugs with well-controlled binding specificity and/or affinity. Therefore, the calculations of binding free energy by means of molecular simulations has been a major area of research in theoretical and computational chemistry/biochemistry[10−20] over the past 40 years. Many different approaches to evaluations of ligand affinities have been developed and can loosely be categorized into three major classes.[21−24] The first class of methods encompasses empirically driven schemes based on training sets derived from complexes with known structures.[25,26] The features of the known protein binding pocket can be translated into set of potential parameters used for virtual computer screening of large compound libraries or for designing novel ligands de novo.[27,28] Although it is a very powerful approach, its usability is limited if the system under study is lacking an extensive training set data.[24]

The second class of methods includes different extensions of popular molecular mechanics/Poisson−Boltzmann (generalized Born) surface area [MM/PB(GB)SA] algorithms,[29,30] where sampling of ligand/receptor coordinates achieved by molecular dynamics (MD) simulations and binding affinity is computed from collected trajectories.[16,31−34] The interaction energies (MM) are represented by respective force-field components for electrostatic and Lennard-Jones intermolecular terms, the nonelectrostatic component of the desolvation free energy is introduced via an empirical term (proportional to buried solvent accessible area), and the term accounting for the electrostatic penalty of water removal from the protein−ligand interface (desolvation penalty) is computed by means of continuum electrostatic models (such as Poisson−Boltzmann or generalized Born model). The collection of frames containing a protein−ligand complex can be extracted from MD/MC simulations, and each contribution can be averaged to obtain the binding free energies. This very popular and attractive method provides a straightforward and robust way for the computation of enthalphic contributions for a collection of frames extracted from MD but meets increasing difficulties when providing an accurate estimate for the loss in degrees of freedom incurred upon moving from bulk solution to the receptor-bound state and when accounting for dynamics/contribution due to explicit water present at the binding site.[35,36]

The third class includes approaches based on all-atom atomistic simulations.[15] The approaches vary from the application of thermodynamic integration[37] and free energy perturbation techniques to the computations of the potential of mean forces[38] with umbrella sampling methods,[39] adaptive biasing force MD,[40] or steered MD methods.[41] The difficulties associated with the requirement of sampling vast conformational space often lead to use of nonequilibrium simulations Jarzynski's equation,[42] modifications of Hamiltonian, such as metadynamics,[43] and a variety of enhanced sampling techniques.[44] Many of theses methods require certain knowledge of the pathway of the drug binding/releasing. A very promising theoretical approach to the problem is to compute absolute binding free energies using a molecular dynamics/free energy perturbation (MD/FEP) method with constraining forces.[35,45] In this method, the free energies are calculated from the thermodynamic reversible work along an unphysical transformation path with MD/FEP

using a potential function that depends on one or several coupling parameters, such that the appropriate potential energy is recovered at the end-points.[15,19,45−47] Arguably, this approach provides the best way to describe the thermodynamics of protein−ligand recognition providing achievement of an adequate conformational sampling. The most common use of the FEP techniques is for the evaluation of the relative binding free energies between two different ligands through the perturbation of one ligand into another with a dual-topology scheme.[17,48] However, calculation of the absolute (standard) free energy arguably provides more detailed information about the mechanisms of ligand binding. It can also lead to the direct connection between macroscopic experimental measurements of binding affinities and the microscopic structural data extracted from FEP simulations.

To overcome expensive conformational sampling in computations of the absolute binding free energies with FEP/MD, one can use restraining potentials.[18,35,45] Recent applications of this technique have consisted of FEP/MD simulations with a reduced generalized solvent boundary condition (GSBP)[49] model with enhanced sampling by using conformational, translational, and orientational restraining potentials.[35] The method was developed and tested on a simple system of FKBP12/ligand[35] as well as a model system of T4 lysozyme/ligand[45] with considerable success. It has also been successfully applied to explain the substrate specificity of the LeuT neurotransmitter transporter,[50] emphasizing its applicability to membrane proteins as well. Despite rigorous theoretical foundation for FEP/MD methods and apparent success in applications studying highly specific binding to proteins,[35] the evaluation of many methodological aspects of these computations and its applicability to studies of ligand binding to membrane proteins is still a future goal.

In this report, we extend FEP/MD simulations to studies of an antidepressant binding to the membrane transporter LeuT. We are exploring the dependence of computed absolute free energies on the choice of atomic models, the strength of restraining potentials, and the reference structure for conformational constraints. We also report on the role of ligand reorganization free energy in high-affinity binding to the protein. It was recently suggested that contribution due to ligand reorganization upon transfer from bulk solution to a protein binding pocket could be as high as 7−8 kcal/mol and could potentially be the major determinant of high- vs low-affinity binding.[28] To test the validity of different approximations (PBC vs reduced GSBP approximation), we compare the free energy results from absolute free energy calculations to those from more common relative free energy calculations.[17,48] This report is organized as follows: Theoretical formulation of the FEP/MD absolute free energy calculation method is briefly reviewed in Section II. Computational details and results are presented in Sections III and IV, respectively. The conclusion of this report is summarized in Section V.

II. Theoretical Formulation

A. Restraint Forces. To focus on relevant degrees of freedom, translational and orientational restraint potentials were implemented in all simulations, as described previously.[18,35,45,51−53] These two restraints are defined by three positions in the protein (Figure 1, center of mass Pc, and two randomly picked positions P1 and P2) and three positions in the ligand (center of mass Lc, and two randomly picked positions L1 and L2). The translational restraint is implemented to constrain the position of the center of mass of the ligand (Lc) relative to the protein. Its form is

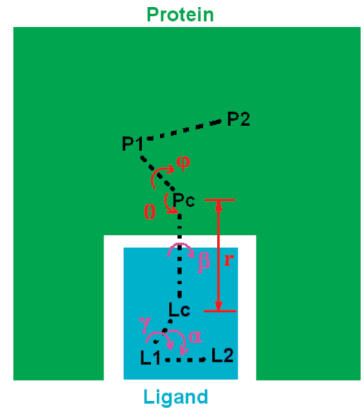


Figure 1. Translational and rotational restraints on the ligand. Three positions in the protein (center of mass Pc, and randomly picked P1 and P2) and three positions in the ligand (center of mass Lc, and randomly picked L1 and L2) were used to set up the translational and rotational restraints. The translational restraints, shown in red, are defined by distance r (Pc and Lc), angle θ (P1, Pc, and Lc), and dihedral φ (P2, P1, Pc, and Lc). The rotational restraints, shown in magenta, are defined by angle α (L1, Lc, and Pc) and dihedrals β (P1, Pc, Lc, and L1) and γ (Pc, Lc, L1, and L2).

$$u_t(r,\theta,\phi) = \frac{1}{2}k_{\text{dist}}(r - r_0)^2 + \frac{1}{2}k_{\text{ang}}(\theta - \theta_0)^2 + \frac{1}{2}k_{\text{ang}}(\phi - \phi_0)^2 \quad (1)$$

where r is the distance between Lc and Pc, θ is the angle P1−Pc−Lc, and φ is the dihedral angle P2−P1−Pc−Lc. The corresponding reference values derived from an average of the equilibration trajectory are r_0, θ_0, and φ_0. The force constants for restraints on distance and angles (including dihedral angles) are k_{dist} and k_{ang}. Similarly, the rotational restraint on the ligand has the form of

$$u_r(\alpha,\beta,\gamma) = \frac{1}{2}k_{\text{ang}}(\alpha - \alpha_0)^2 + \frac{1}{2}k_{\text{ang}}(\beta - \beta_0)^2 + \frac{1}{2}k_{\text{ang}}(\gamma - \gamma_0)^2 \quad (2)$$

where α is the angle Lc−L1−L2, β is the dihedral angle P1−Pc−Lc−L1, and γ is the dihedral angle Pc−Lc−L1−L2. The corresponding reference values derived from an average of the equilibration trajectory are α_0, β_0, and γ_0. The translational and rotational restraints ensure that the ligand is around its bound state. A configurational restraint (u_c), in the form of a harmonic potential with respect to the root-mean-square deviation (RMSD) of the ligand, relative to a reference configuration, is also applied to constrain the ligand configuration.

B. Standard Free Energy. The free energy (ΔG_b) of a ligand (L) binding to a receptor protein (R) correlates to the equilibrium constant K_b of the binding reaction L + R⇔L · R by

$$K_b = \frac{[\mathrm{L}\cdot\mathrm{R}]}{[\mathrm{L}][\mathrm{R}]} = \exp[-\beta\Delta G_b] \tag{3}$$

where $\beta \equiv 1/(k_B T)$, with k_B being the Boltzmann constant and T being the absolute temperature. Assuming the binding of a ligand is defined as moving one ligand molecule from the bulk solution to the binding site, K_b can be expressed as the following equation at low ligand concentration:

$$K_b = \frac{\int_{\mathrm{site}} d\mathbf{L} \int d\mathbf{X} e^{-\beta U}}{\int_{\mathrm{bulk}} d\mathbf{L}\delta(r_{\mathrm{L}} - r^*) \int d\mathbf{X} e^{-\beta U}} \tag{4}$$

where **L** and **X** are the coordinates of the ligand molecule and the remaining atoms (including solvent, receptor protein, counterions, and others), respectively. U is the total potential energy of the system, r_{L} is the position of the center of mass of ligand L, and r^* is some arbitrary position in the bulk. The δ function is a result of the translational invariance of the ligand in the bulk. To evaluate the binding free energy with molecular simulations, Deng et al.[53] wrote eq 4 in the form of the multiple of a series of intermediate states connecting the initial (the ligand in the binding site) and final (the ligand in the bulk) states:

$$\begin{aligned} K_b = {} & \frac{\int_{\mathrm{site}} d\mathbf{L} \int d\mathbf{X} e^{-\beta U_1}}{\int_{\mathrm{site}} d\mathbf{L} \int d\mathbf{X} e^{-\beta(U_1+u_{\mathrm{c}})}} \times \\ & \frac{\int_{\mathrm{site}} d\mathbf{L} \int d\mathbf{X} e^{-\beta(U_1+u_{\mathrm{c}})}}{\int_{\mathrm{site}} d\mathbf{L} \int d\mathbf{X} e^{-\beta(U_1+u_{\mathrm{c}}+u_{\mathrm{t}})}} \times \\ & \frac{\int_{\mathrm{site}} d\mathbf{L} \int d\mathbf{X} e^{-\beta(U_1+u_{\mathrm{c}}+u_{\mathrm{t}})}}{\int_{\mathrm{site}} d\mathbf{L} \int d\mathbf{X} e^{-\beta(U_1+u_{\mathrm{c}}+u_{\mathrm{t}}+u_{\mathrm{r}})}} \times \\ & \frac{\int_{\mathrm{site}} d\mathbf{L} \int d\mathbf{X} e^{-\beta(U_1+u_{\mathrm{c}}+u_{\mathrm{t}}+u_{\mathrm{r}})}}{\int_{\mathrm{site}} d\mathbf{L} \int d\mathbf{X} e^{-\beta(U_0+u_{\mathrm{c}}+u_{\mathrm{t}}+u_{\mathrm{r}})}} \times \\ & \frac{\int_{\mathrm{site}} d\mathbf{L} \int d\mathbf{X} e^{-\beta(U_0+u_{\mathrm{c}}+u_{\mathrm{t}}+u_{\mathrm{r}})}}{\int_{\mathrm{bulk}} d\mathbf{L} \int d\mathbf{X} e^{-\beta(U_0+u_{\mathrm{c}}+u_{\mathrm{t}})}} \times \\ & \frac{\int_{\mathrm{bulk}} d\mathbf{L} \int d\mathbf{X} e^{-\beta(U_0+u_{\mathrm{c}}+u_{\mathrm{t}})}}{\int_{\mathrm{bulk}} d\mathbf{L}\delta(r_{\mathrm{L}} - r^*) \int d\mathbf{X} e^{-\beta(U_0+u_{\mathrm{c}})}} \times \\ & \frac{\int_{\mathrm{bulk}} d\mathbf{L}\delta(r_L - r^*) \int d\mathbf{X} e^{-\beta(U_0+u_{\mathrm{c}})}}{\int_{\mathrm{bulk}} d\mathbf{L}\delta(r_L - r^*) \int d\mathbf{X} e^{-\beta(U_1+u_{\mathrm{c}})}} \times \\ & \frac{\int_{\mathrm{bulk}} d\mathbf{L}\delta(r_L - r^*) \int d\mathbf{X} e^{-\beta(U_1+u_{\mathrm{c}})}}{\int_{\mathrm{bulk}} d\mathbf{L}\delta(r_L - r^*) \int d\mathbf{X} e^{-\beta(U_1)}} \end{aligned} \tag{5}$$

where the subscript 1 and 0 of U indicate fully interacting and fully decoupled ligand.

In terms of free energy contributions, the binding constant can be written as

$$\begin{aligned} K_b = {} & \exp(+\beta\Delta G_{\mathrm{c}}^{\mathrm{site}}) \times \exp(+\beta\Delta G_{\mathrm{t}}^{\mathrm{site}}) \times \exp(+\beta\Delta G_{\mathrm{r}}^{\mathrm{site}}) \times \\ & \exp(-\beta\Delta G_{\mathrm{int}}^{\mathrm{site}}) \times F_{\mathrm{r}} \times F_{\mathrm{t}} \times \exp(+\beta\Delta G_{\mathrm{int}}^{\mathrm{bulk}}) \times \\ & \exp(-\beta\Delta G_{\mathrm{c}}^{\mathrm{bulk}}) \end{aligned} \tag{6}$$

where the terms sequentially correspond to the terms in eq 5. All the terms involving ΔG can be calculated by the standard free energy perturbation method,[15,54,55] while the free energy components associated with the configurational constraint ($\Delta G_{\mathrm{c}}^{\mathrm{site}}$ and $\Delta G_{\mathrm{c}}^{\mathrm{bulk}}$) can be better obtained by an umbrella sampling scheme and the translational (F_{t}), and rotational (F_{r}) factors can be evaluated directly with numerical integration schemes, since the interaction between the ligand molecule and the environment is turned off.

The standard binding free energy, defined relative to the standard concentration of 1 mol/L, is

$$\begin{aligned} \Delta G_b^0 \equiv -k_{\mathrm{B}}T\ln[K_b C^0] = {} & \Delta\Delta G_{\mathrm{int}} + \Delta\Delta G_{\mathrm{c}} + \\ & \Delta\Delta G_{\mathrm{t}}^0 + \Delta\Delta G_{\mathrm{r}} \end{aligned} \tag{7}$$

where $C^0 = 1$ mol/L and $K_b C^0$ gives the standard binding constant. The free energy contributions are grouped as: $\Delta\Delta G_{\mathrm{int}} = \Delta G_{\mathrm{int}}^{\mathrm{site}} - \Delta G_{\mathrm{int}}^{\mathrm{bulk}}$, $\Delta\Delta G_{\mathrm{c}} = \Delta G_{\mathrm{c}}^{\mathrm{bulk}} - \Delta G_{\mathrm{c}}^{\mathrm{site}}$, $\Delta\Delta G_t^0 = -\Delta G_{\mathrm{t}}^{\mathrm{site}} - k_{\mathrm{B}}T \ln (F_{\mathrm{t}}C^0)$, and $\Delta\Delta G_{\mathrm{r}} = -\Delta G_{\mathrm{r}}^{\mathrm{site}} - k_{\mathrm{B}}T \ln (F_r)$, where $\Delta\Delta G_{\mathrm{int}}$corresponds to the free energy difference associated with removing the ligand, restrained by the potential u_c from the bulk and inserting it to the binding site, restrained by u_{c}, u_{t}, and u_r; $\Delta\Delta G_{\mathrm{t}}^0$and $\Delta\Delta G_{\mathrm{r}}$ correspond to the free energy changes associated with turning on and off the translational and rotational restraints on the ligand; $\Delta\Delta G_{\mathrm{c}}$ corresponds to free energy associated with the application of RMSD restraints. Recently, Deng and Roux further incorporated a grand canonical Monte Carlo step into the absolute binding free energy calculation to account for the fluctuation of the number of water molecules in highly occluded binding sites during alchemical perturbation.[19,56] For the TCA/LeuT systems, however, the TCA binding sites are open to access by water molecules from the extracellular side of the membrane. Thus, standard molecular dynamics trajectories offer adequate sampling for fluctuations in the number of water molecules around the binding site (Figure S2 in the Supporting Information).

C. Decomposition of the Interaction Free Energy. The interaction free energy, $\Delta G_{\mathrm{int}}^{\mathrm{site}}$ and $\Delta G_{\mathrm{int}}^{\mathrm{bulk}}$, are decomposed into the contributions from electrostatic and Lennard-Jones (LJ) components. Further more, with the application of the Weeks−Chandler−Andersen (WCA) scheme,[57] the LJ potential, written as eq 8 in the CHARMM 27 all-atom force field, is uniquely separated into the repulsive (eq 9) and the dispersive (eq 10) potentials.

$$U^{\mathrm{LJ}}(r) = \varepsilon\left[\left(\frac{R_{\min}}{r}\right)^{12} - 2\left(\frac{R_{\min}}{r}\right)^{6}\right] \tag{8}$$

$$U^{\mathrm{repu}}(r) = U^{\mathrm{LJ}}(r) + \varepsilon \quad \text{when} \quad r < R_{\min}; \qquad U^{\mathrm{repu}}(r) = 0 \quad \text{when} \quad r \geq R_{\min} \tag{9}$$

$$U^{\mathrm{disp}}(r) = -\varepsilon \quad \text{when} \quad r < R_{\min}; \quad U^{\mathrm{disp}}(r) = U^{\mathrm{LJ}}(r) \quad \text{when} \quad r \geq R_{\min} \tag{10}$$

where ε has the dimension of energy, and $R_{\min}$ has the dimension of length. When the separation r of two atoms is at $R_{\min}$, the LJ potential reachs its well depth $-\varepsilon$.

The interaction free energies ($\Delta G^{\mathrm{a}}_{\mathrm{int}}$ where a represents *site* or *bulk*) are thus further separated into three components due to the contributions from electrostatic ($\Delta G^{\mathrm{a}}_{\mathrm{elec}}$), dispersive LJ ($\Delta G^{\mathrm{a}}_{\mathrm{disp}}$), and repulsive LJ ($\Delta G^{\mathrm{a}}_{\mathrm{repu}}$) potentials.[53] This decomposition has shown to increase the statistical accuracy of the calculation of hydration free energies of molecules.[53] More importantly, while the decomposition of the interaction free energies is arbitrary to some degree, the values of these contributions helps to understand the nature of ligand binding.[35,45,53,56]

D. Free Energy Perturbation. The interaction free energies are evaluated by alchemical transformations using the standard free energy perturbation approach.[15,54,55] Briefly, the free energy contribution is calculated by gradually turning the potential on or off using a coupling parameter λ valued from 0 to 1. For example, to calculate the dispersive free energy for the ligand binding to the receptor binding site ($\Delta G^{\mathrm{site}}_{\mathrm{disp}}$), a coupling parameter λ_{disp} is introduced. When $\lambda_{\mathrm{disp}} = 0$, the dispersive interaction between the ligand and the environment is completely turned off, and when $\lambda_{\mathrm{disp}} = 1$, the dispersive interaction between the ligand and the environment is completely turned on. The resulting auxiliary potential energy with the coupling constant is as follows

$$U(\lambda_{\mathrm{disp}}) = U^0 + U^{\mathrm{disp}}(\lambda_{\mathrm{disp}}) \tag{11}$$

where U^0 is the total potential when the dispersive interactions between the ligand and the environment are completely turned off, and $U^{\mathrm{disp}}(\lambda_{\mathrm{disp}})$ is the total dispersive potential between the ligand and the environment scaled by the coupling constant. Several windows are applied to gradually increase the coupling constant from 0 to 1. For each window (from $\lambda_{\mathrm{disp},i}$ to $\lambda_{\mathrm{disp},j}$), the ensemble average $\langle \exp\{-\beta[U(\lambda_{\mathrm{disp},j}) - U(\lambda_{\mathrm{disp},i})]\}\rangle_{U(\lambda_{\mathrm{disp},i})}$ is obtained, and the sum of these windows corresponds to $\exp(-\beta \Delta G^{\mathrm{site}}_{\mathrm{disp}})$. Similar FEP/MD methods can be applied to calculate the other ΔG components. For a detailed description of these FEP/MD procedures, readers are referred to the work by Deng et al.[45]

E. Calculation of the Different Free Energy Components. Combining the sequential process in eq 5 and the decomposition of the interaction energy described in section C, the steps corresponding to the dissociation of the fully interacting ligand in the protein binding site [system U_1(site)] as the initial state are listed in Table S1 in the Supporting Information. For each step, the second column gives the initial system, and the third column gives the system change. The corresponding free energy component and the method used to compute it are listed in the third and fourth columns, respectively. The free energies associated with the configurational constraints of the ligand to the reference configuration, $\Delta G^{\mathrm{site}}_{\mathrm{c}}$ and $\Delta G^{\mathrm{bulk}}_{\mathrm{c}}$, are calculated by integrating the Boltzmann factor of the RMSD potential of mean force (PMF) obtained from umbrella sampling simulations. The translational factor (F_{t}) and the rotational factor (F_{r}) are numerically integrated from the expressions of

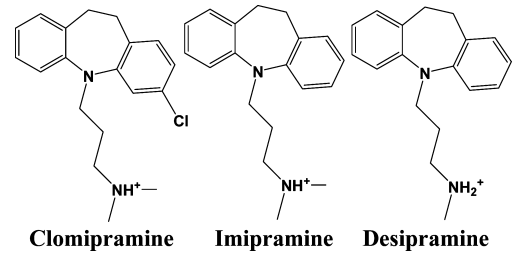


Figure 2. Structural formulas of the three tricyclic antidepressants (TCAs) used for studies. The side-chain nitrogens are set in protonated form in accordance to the experimental conditions.

$$F_{\mathrm{t}} = \int_0^{\infty} dr r^2 \int_0^{\pi} d\theta \sin(\theta) \int_{-\pi}^{\pi} d\phi \exp[-\beta u_{\mathrm{t}}(r, \theta, \phi)] \tag{12}$$

$$F_{\mathrm{r}} = \frac{1}{8\pi^2} \int_0^{\pi} d\alpha \sin(\alpha) \int_0^{\pi} d\beta \int_{-\pi}^{\pi} d\gamma \exp[-\beta u_{\mathrm{r}}(\alpha, \beta, \gamma)] \tag{13}$$

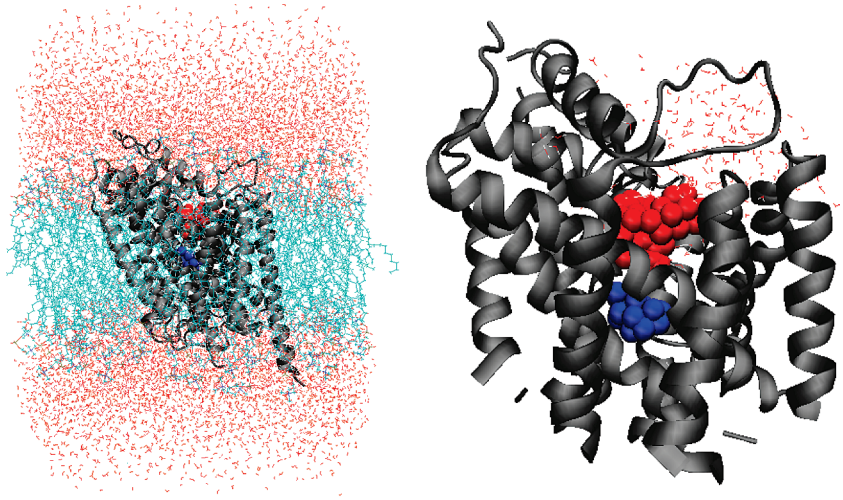
where $r, \theta, \varphi, \alpha, \beta, \gamma$ are the constrained internal coordinates illustrated in Figure 1, and u_t and u_r are the translational and rotational restraining potentials applied to the bound ligand presented in eqs 1 and 2. FEP/MD simulations are applied to get all of the other components of the absolute binding free energy.[45]

From Table S1 in the Supporting Information, it should be noted that when $\Delta G^{\mathrm{site}}_{\mathrm{elec}}$, $\Delta G^{\mathrm{site}}_{\mathrm{disp}}$, and $\Delta G^{\mathrm{site}}_{\mathrm{repu}}$ are calculated, restraints are applied during the alchemical transformations, thus making the values conditional on the restraining forces. Nonetheless, a comparison to their corresponding values in the bulk is informative about the binding.

III. Methods

A. Molecular Models of TCAs. The chemical structures of the TCAs are shown in Figure 2. A neutral model for clomipramine (CMI) was originally developed in ref 58. New models with a positive charge for CMI, imipramine (IMI), and desipramine (DSI) were developed with the same procedure as described for CMI. Briefly, the geometric parameters (bond lengths, angles) were extracted from the crystallographic data. The positions of hydrogen atoms were unavailable from the crystal structures and were obtained by the HBUILD utility implemented with the CHARMM program.[59] The CHARMM27 force field[60] was used for the intramolecular potentials, and the nonbonded Lennard-Jones potential was used for the three drugs. Some of the torsional potentials are not available in the CHARMM27 force field and were obtained by fitting the B3LYP/6-31G* torsion profiles.

Electrostatic potentials are crucial for reproducing the binding affinities of drugs to protein receptors. We assigned specific partial charges to each atom of the drugs using the



(a) (b)

Figure 3. Simulation system for the TCA and substrate-bound leucine transporter LeuT. (a) The full simulation system for the MD equilibration. The protein is shown in cartoon mode (gray). A leucine substrate (blue) and a clomipramine drug (red) bound to the protein are shown in space-fill mode. The DPPC membrane (cyan) and water (red) molecules are shown in line mode, while the bound sodium ions and counterions (100 mM NaCl) are not shown. The whole system contains about 60 000 atoms. (b) A sphere containing LeuT (gray cartoon) and substrate (blue space-fill), the bound TCA (red space-fill), and water molecules (red lines) in the GSBP simulation system used for FEP/MD. Only atoms in a sphere (20 Å radius) centered on the ligand are represented explicitly. All atoms outside the sphere are represented implicitly using a continuum electrostatic approach.

RESP fitting approach described by Anisimov et al.[61] Briefly, for each drug, an electrostatic potential (ESP) grid was created on several Connolly surfaces[62] of the molecule by the CGRID program. ESP calculation at the B3LYP/6-31G* level was applied to obtain the electrostatic map at the grid points, which was then used for partial charge fitting with the FITCHARGE module of the CHARMM program. The initial charge set ($CS_{initial}$) was obtained for each atom type from the CHARMM nonpolarizable force field. A parabolic penalty function was used to restrain the values of the fitted charge set (CS_{fitted}) with restraining forces of 10^{-4} Å^2 to the initial charge set ($CS_{initial}$). Thus, the partial charges of CS_{fitted} have a better match to the electrostatic map. The initial ($CS_{initial}$) and fitted (CS_{fitted}) charges are presented in the Supporting Information. The total charge of each TCA molecule is +1 with the side-chain nitrogen protonated based on the reported pK_a values of the TCAs.[63,64]

B. Equilibrium MD Simulation. The starting configuration of drug and substrate-bound LeuT were taken from the X-ray coordinates revealed by Singh et al.[2] (Protein Data Bank entries 2Q6H, 2Q72, and 2QB4). The complexes were embedded in a lipid membrane using a multistep membrane-building procedure used in previous studies.[50] The simulation box contains the LeuT transporter, bound sodium and/or chloride ions, one leucine substrate, one antidepressant (clomipramine, imipramine, or desipramine) bound at the extracellular gate, and 148 dipalmitoylphosphatidylcholine (DPPC) lipid molecules solvated in an explicit 100 mM NaCl aqueous solution. A snapshot of the full simulation box is shown in Figure 3a. All computations were carried out by CHARMM, version c34b2, with the CHARMM27 force fields for proteins and lipids. MD simulation methods used here are similar to those used in previous studies of membrane systems.[50] Briefly, constant temperature/pressure algorithms were applied (with pressure at 1 atm and temperature at 315 K). Periodic boundary conditions were used for the hexagonal system. Electrostatic interactions were treated with the particle mesh Ewald (PME) algorithm with a 96 × 96 × 96 Å grid for fast Fourier transform, $\kappa = 0.34$ Å^{-1}, and a sixth-order spline interpolation. The nonbonded interactions were smoothly switched off at 12−14 Å. All simulation systems were equilibrated for 5 ns each without any configurational constraints.

C. Absolute Binding Free Energy Calculation. Following the equilibrium MD simulation, drug binding free energies were calculated using the protocol described in section II. To decrease computational cost, only the atoms in and surrounding the binding site (within 20 Å of the bound drug) were treated explicitly. All other atoms in the system were considered implicit, using a GSBP[49] generated for each system. It has been shown that the use of GSBP significantly decreases the size of the system (in our case from ∼60 000 to ∼7000 atoms, Figure 3b) while keeping the statistical error relatively low (∼1−2 kcal/mol).[35,45] After the GSBP maps were generated, the reduced systems were minimized and equilibrated for 0.5 ns. Using the free energy decomposition protocol, the free energy components resulting from electrostatic, dispersive, repulsive, and constraining forces were measured independently. The CHARMM PERT function with the additional CHEMPERT option[59] was used for the FEP/MD simulations. All FEP/MD runs were equilibrated for 0.1 ns before collecting data during a 0.4 ns run. For the electrostatic component, both forward and reverse windows were calculated with the values for the coupling parameter λ set to [0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0]. The dispersive component was measured through four forward windows with λ set to [0.0, 0.25, 0.75, 1.0]. The repulsive component was calculated by annihilating the

Table 1. Absolute Free Energy of Hydration for Clomipramine (C), Imipramine (I), and Desipramine (D)[a]

		G_{rep}	G_{disp}	G_{elec}	G_{conf}	G_{tot}	G_{conf_vac}	G_{solv}
CS_{fitted}	C	39.4 ± 0.5	−35.3 ± 0.2	−50.7 ± 0.1	−4.9 ± 0.6	**−51.4 ± 0.4**	−3.1	−48.3
	I	38.3 ± 0.5	−33.6 ± 0.2	−51.4 ± 0.2	−5.5 ± 0.8	**−52.2 ± 0.8**	−7.2	−45.1
	D	37.3 ± 0.9	−32.4 ± 0.2	−56.2 ± 0.1	−3.6 ± 1.2	**−54.9 ± 1.9**	−5.0	−49.9
$CS_{initial}$	C	39.9 ± 0.8	−35.0 ± 0.2	−54.9 ± 0.1	−7.2 ± 1.4	**−57.2 ± 2.4**	−10.0	−47.1
	I	38.5 ± 0.7	−33.6 ± 0.2	−55.5 ± 0.3	−4.6 ± 0.8	**−55.2 ± 0.6**	−9.3	−45.8
	D	37.3 ± 0.8	−32.2 ± 0.1	−62.3 ± 0.1	−3.9 ± 0.1	**−61.1 ± 0.8**	−8.5	−52.6

[a] The data are obtained with the application of SSBP. The first group (rows 2−4) of results is obtained from simulations with the fitted charge set (CS_{fitted}), and the second group (rows 5−7) of results is obtained with the initially guessed charge set ($CS_{initial}$). G_{tot} is the sum of the free energy components of repulsive (G_{rep}), dispersive (G_{disp}), electrostatic (G_{elec}), and conformational (G_{conf}).

molecule from binding site or bulk solvent with the application of a soft-core potential.[53] The annihilation is done from both forward and reverse windows with λ set to [0.0, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0]. The translational and rotational constraint components were measured through forward windows with λ set to [0.0, 0.0025, 0.0050, 0.0075, 0.01, 0.02, 0.04, 0.06, 0.08, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0]. Similar coupling constants were applied in the work of Deng et al.[45] and Wang et al.[35] The results were processed with the weighted histogram analysis method[65] (the WHAM module of CHARMM) to remove any bias due to the restraining forces. Hydration free energies of the TCAs were calculated by FEP/MD, using a model system of TCA solvated by 400 water molecules. The spherical solvent boundary potential (SSBP)[53,66] was applied to account for the influence outside of the solvation sphere. The FEP protocol used to determine hydration free energy is the same as that described above for the computation of absolute binding free energy. Equilibration without constraints was performed for 100 ps, and window lengths for evaluation of free energy were 200 ps. Statistical uncertainty is reported with the standard deviations of each free energy component and the total solvation, site, and binding free energies obtained through separating the FEP/MD simulations to three blocks and obtaining the free energies for each block using WHAM analysis.[65]

D. Relative Binding Free Energy Calculation. The binding free energy difference between CMI and IMI as well as IMI and DSI were calculated with the CHARMM PERT function. For example, in the case of CMI/IMI, the equilibrated membrane system of LeuT/CMI (equilibrated from PDB entry 2Q6H) is used as the starting configuration ($\lambda = 0$), and in the final configuration, CMI is perturbed to IMI ($\lambda = 1$). For the perturbation, 11 windows were used varying between 0.0 and 1.0 by increments of 0.1. For each perturbation window, a 10 ps equilibration run and a 190 ps production run were applied. We also use the equilibrated membrane system of LeuT/IMI (equilibrated from PDB entry 2Q72), as the $\lambda = 0$ state and similar procedures were carried out to calculate the relative binding free energy by perturbing IMI to CMI. The relative binding free energies from the forward and backward simulations are within 1 kcal/mol of difference, and the average relative free energy is reported.

IV. Results and Discussion

A. Hydration Free Energy of TCAs. The hydration free energies of the three TCAs are listed in Table 1. G_{tot}, the sum of the free energy components of repulsive (G_{rep}), dispersive (G_{disp}), electrostatic (G_{elec}) and conformational (G_{conf}), will be subtracted from G_{tot}(site) (Table 2) to obtain the absolute binding free energy. To obtain the hydration free energy, the component caused by RMSD constraining force for the solute in vacuum (G_{conf_vac}) needs to be subtracted from G_{tot}. The resulting hydration free energies for all the three TCAs from the calculations with SSBP potential are in the order of −50 kcal/mol. The calculations indicate that the electrostatic component (G_{elec}), compared to the LJ component ($G_{repu} + G_{disp}$) of the free energy, mostly accounts for the favorable hydration free energies for the three TCAs.

The influence of applied partial charge sets on the solvation free energy is apparent. While there are only slight differences (within 1 kcal/mol) on solvation free energies for the two charge sets (CS_{fitted} and $CS_{initial}$) of CMI and IMI, there is a decrease of about 3 kcal/mol in hydration free energy of DSI from CS_{fitted} to $CS_{initial}$. The most significant change is the electrostatic free energy component which has been decreased from −56.2 for CS_{fitted} to −62.3 kcal/mol for $CS_{initial}$. This indicates that, given relatively small changes in the partial charges (Table S2, columns 7 to 6 in the Supporting Information), the electrostatic contribution of hydration free energy could be dramatic. It should be noted, however, that the hydration free energies obtained here are for drug transfer from vacuum to bulk solution. The solubility data reported for many pharmaceutical agents is not readily comparable to theoretical results, as one has to first correct for the drug sublimation free energy.

B. Absolute Binding Free Energies of TCAs Binding to LeuT. The computed absolute binding free energies of TCAs to LeuT are listed in Table 2. The absolute free energy of binding is the difference between the free energies of site binding and hydration (G_{tot} in Table 1). As described, the absolute energy is decomposed into the contributions from electrostatic, dispersive, and repulsive parts as well as from constraint potentials on RMSD, orientation, and translation. The final absolute binding free energies are −13.6, −13.4, and −12.0 kcal/mol for CMI, IMI, and DSI, respectively. All of them are negative, indicating favorable binding. As discussed in Section II, while the decomposition of the free energies is path dependent, the comparison between the same contributions for different ligands can still provide valuable information about the nature of binding (Figure 4). CMI has a lower binding free energy than IMI mainly due to its gain in the dispersive free energy (−4.1 kcal/mol), which may be due to favorable interactions achieved by its additional chlorine with the LeuT side chain.[2] Despite the overall

Table 2. Absolute Free Energy of Binding for Clomipramine (C), Imipramine (I), and Desipramine (D) to LeuT in Reduced GSBP System[a]

						G_{pos}				
			G_{rep}	G_{disp}	G_{elec}	G_{const}	$-k_BT\ \ln(F_tC^0)$	$-k_BT\ \ln(F_r)$	G_{conf}	G_{tot}
CS_{fitted}	site	C	30.0 ± 0.7	−54.6 ± 0.3	−48.7 ± 0.6	−3.0 ± 0.1	5.4	6.5	−0.5 ± 0.1	−64.9 ± 1.2
		I	30.5 ± 1.1	−48.8 ± 0.4	−55.3 ± 0.8	−2.6 ± 0.1	5.4	6.4	−1.2 ± 0.1	−65.6 ± 0.4
		D	28.0 ± 0.4	−50.8 ± 0.1	−53.6 ± 0.1	−2.1 ± 0.1	5.6	6.4	−0.5 ± 0.1	−67.0 ± 0.6
	binding	C	−9.4 ± 0.9	−19.4 ± 0.6	2.0 ± 0.6		8.8 ± 0.1		4.4 ± 0.5	**−13.6 ± 1.4**
		I	−7.8 ± 1.4	−15.3 ± 0.6	−3.9 ± 0.9		9.3 ± 0.1		4.3 ± 0.8	**−13.4 ± 1.0**
		D	−9.3 ± 0.4	−18.3 ± 0.1	2.6 ± 0.1		9.9 ± 0.1		3.0 ± 1.2	**−12.1 ± 1.4**
$CS_{initial}$	site	C	30.6 ± 0.9	−54.4 ± 0.3	−54.3 ± 0.4	−1.5 ± 0.1	5.4	6.5	−0.5 ± 0.1	−68.2 ± 0.6
		I	33.1 ± 0.2	−48.3 ± 0.1	−58.3 ± 0.4	−1.9 ± 0.1	5.4	6.4	−1.6 ± 0.1	−65.2 ± 0.4
		D	26.4 ± 1.1	−49.0 ± 0.5	−60.2 ± 0.6	−2.6 ± 0.1	5.4	6.4	−1.3 ± 0.1	−74.9 ± 1.7
	binding	C	−9.3 ± 1.5	−19.4 ± 0.2	0.6 ± 0.3		10.3 ± 0.1		6.7 ± 1.3	−11.1 ± 2.5
		I	−5.5 ± 0.7	−14.7 ± 0.3	−2.7 ± 0.3		9.9 ± 0.1		3.0 ± 0.7	−10.0 ± 0.7
		D	−10.9 ± 1.6	−16.8 ± 0.4	2.1 ± 0.6		9.2 ± 0.2		2.6 ± 0.2	−13.8 ± 2.1

[a] The first group (rows 2−7) of results is obtained from simulations with the fitted charge set (CS_{fitted}), and the second group (rows 8−13) of results is obtained with the initially guessed charge set ($CS_{initial}$)*. For the site free energy, G_{tot} is the sum of the free energy components of repulsive (G_{rep}), dispersive (G_{disp}), electrostatic (G_{elec}), positional (G_{pos}, including translational and rotational constraints), and conformational (G_{conf}). For the binding free energy, the values are obtained by subtracting the corresponding components of solvation free energy (Table 1) from the site free energy. Note that $CS_{initial}$ is the initial charge set based on CHARMM force field, and the atomic charges were derived from similar atom types in the CHARMM27 force field and were not parametrized. CS_{fitted} is the RESP charge fitted for the electrostatic density map from QM calculation, and it reflects the chemical environment of the atoms of the drugs better.

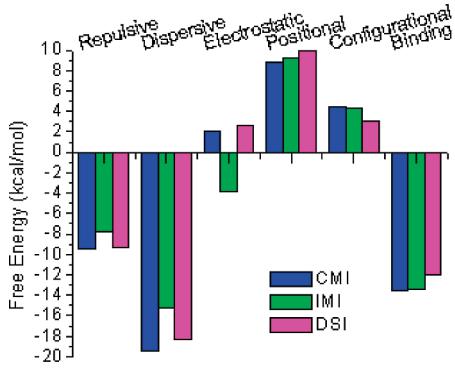


Figure 4. Column illustration of the binding free energy components of three TCA's (CMI in blue, IMI in green, and DSI in magenta) binding to LeuT. From left to right: repulsive, dispersive, electrostatic, positional (including translational and rotational constraints), configurational, and total binding free energy.

unfavorable binding free energy for DSI compared to IMI, the DSI binding free energy exhibits a more favorable repulsive component. This is explained by the fact that DSI has one less methyl group on its "tail" and is accommodated better by the binding pocket.

The complete expression for the binding constants (eq 4−5) contains a term that describes conformational dynamics of the receptor. To perform their function, membrane transporters may undergo large conformational changes, binding and unbinding ions, and opening and closing extracellular and intracellular gates. Such events take place over large time intervals in the μs to s range. These changes in the protein's structure are not present in the ns MD/free energy simulations. Only one state of the transporter (the "occluded" state) is considered, leading, therefore, to a large overestimation of the absolute drug binding affinities. The contribution of the binding site's conformational changes to the complete partition function is expected to be unfavorable. This problem is well-known, and the interested reader may refer to an excellent review by Mobley and Dill.[67] Clearly, this component plays a dominant role in differences between computed and experimental binding affinities. However, assuming that these large conformational changes are ion induced and independent of the particular drug, one may conclude that relative free energies or drug ranking based on the absolute free energy computations will be robust, since the term describing receptor dynamics will cancel out. Regardless, the absolute binding free energies, though lacking in contribution from receptor allosteric changes, provide molecular insights on key factors governing formation of the high-affinity/-specificity complex between the protein and the drug. Below we will provide detailed discussion on the importance of different factors in computations of absolute binding affinities with FEP/MD Simulations.

C. Restraints of Different Strengths. The application of biasing restraints on configuration, translation, and orientation of the ligand greatly reduces the configuration space and enhances the sampling efficiency. However, the choice of the restraining force constants has been shown to affect the outcome of individual free energy contributions (i.e., dispersive, repulsive, electrostatic, etc.), despite the resulting binding free energy being largely unaffected.[35] In this section, several sets of different force constants are used in calculating the absolute free energy of CMI binding to LeuT. The effect of these contributions on the total absolute free energy is evaluated. From Table 3, the distance constant (k_t, in kcal/mol/Å^2) has little effect on the value of G_{int}(site), the sum of electrostatic, dispersive, and repulsive free energies as well as the total absolute binding free energy, which is evident from the comparison of the results from ($k_c = 10$, $k_t = 10$, and $k_a = 200$) to ($k_c = 10$, $k_t = 1$, and $k_a = 200$). Reducing the strength of the angular and dihedral force constant (k_a, in kcal/mol/rad^2) makes G_{int}(site) more unfavorable but only slightly affects the final binding free energy. Lowering the RMSD force constant (k_c, in kcal/mol/Å^2) to 1 kcal/mol/Å^2 only slightly changes both G^{int} and the final binding free energy, as evident in the set ($k_c = 1$, $k_t = 10$, and $k_a = 200$). Thus, the choice of the constraining forces is robust to some

Table 3. Computed Binding Free Energy for the Clomipramine/LeuT Complex at Different Force Constants for the RMSD Potentials, the Translational Restraint, and the Rotational Restraint[a]

k_c:k_t:k_a	G_{repu} (site)	G_{disp} (site)	G_{elec} (site)	$-G_{const}$ (site)	$-k_BT\ \ln(F_tC^0)$	$-k_BT\ \ln(F_r)$	$-G_{conf}$ (site)	G^0_{tot} (site)	$\Delta G^0_{binding}$
10:10:200	30.0 ± 0.7	−54.6 ± 0.3	−48.7 ± 0.6	−3.0 ± 0.1	5.4	6.5	−0.5 ± 0.1	−64.9 ± 1.2	**−13.6 ± 1.4**
1:1:20	32.5 ± 0.8	−52.2 ± 0.5	−48.4 ± 0.6	−0.3 ± 0.1	3.2	4.3	−0.1	−61.0 ± 0.7	**−13.3 ± 1.0**
10:1:200	30.9 ± 1.2	−54.2 ± 0.5	−48.3 ± 0.3	−1.4 ± 0.1	4.4	6.2	−0.4 ± 0.1	−62.8 ± 1.0	**−12.0 ± 0.5**
10:10:20	32.2 ± 3.4	−53.8 ± 0.3	−48.9 ± 0.2	−0.4 ± 0.1	3.9	4.3	−0.5	−63.1 ± 3.2	**−10.3 ± 2.1**
1:10:200	32.2 ± 2.6	−53.8 ± 0.4	−49.1 ± 0.7	−1.6 ± 0.1	5.1	6.2	−0.1	−61.0 ± 3.2	**−13.8 ± 3.1**
100:100:2000	28.6 ± 1.0	−55.0 ± 0.4	−49.1	−4.3 ± 0.1	7.5	8.6	−47.4 ± 0.9	−111.1 ± 1.0	**−18.3 ± 1.0**
0:0:0	38.9 ± 1.7	−53.0 ± 0.3	−47.6 ± 0.4	N/A	N/A	N/A	N/A	−61.7 ± 2.2	**−15.4 ± 1.8**

[a] The values k_c (in kcal/mol/Å^2), k_t (in kcal/mol/Å^2), and k_a (in kcal/mol/rad^2) are the force constants for the RMSD potentials, the distance force constant for the translational restraint, and the angle/dihedral force constant for the translational and rotational restraints, respectively.

Table 4. Absolute Free Energy of Binding for Clomipramine (C), Imipramine (I), and Desipramine (D) to LeuT in Reduced GSBP System[a]

					G_{pos}				
		G_{rep}	G_{disp}	G_{elec}	G_{const}	$-k_BT\ \ln(F_tC^0)$	$-k_BT\ \ln(F_r)$	G_{conf}	G_{tot}
site	C	36.5 ± 0.6	−54.2 ± 0.2	−48.9 ± 0.8	−3.0 ± 0.2	5.4	6.5	−10.7 ± 0.3	−68.4 ± 0.8
	I	32.3 ± 1.3	−49.2 ± 0.1	−51.3 ± 0.5	−2.6 ± 0.1	5.4	6.4	−5.3 ± 0.3	−64.3 ± 1.6
	D	30.3 ± 0.7	−51.7 ± 0.2	−52.2 ± 0.7	−2.6 ± 0.1	5.6	6.4	−5.1 ± 0.4	−69.3 ± 0.9
solv	C	40.4 ± 0.3	−35.9 ± 0.3	−51.7 ± 0.1		N/A		−3.5 ± 1.0	−50.7 ± 1.6
	I	39.1 ± 0.6	−34.0 ± 0.2	−51.5 ± 0.1		N/A		−4.4 ± 0.5	50.9 ± 0.3
	D	36.7 ± 0.7	−32.6 ± 0.3	−56.0 ± 0.1		N/A		−2.7 ± 0.2	−54.6 ± 0.6
bind-ing	C	−3.9 ± 0.9	−18.3 ± 0.4	2.7 ± 0.8		8.8 ± 0.2		−7.2 ± 0.8	−17.9 ± 1.5
	I	−6.7 ± 0.8	−15.2 ± 0.2	0.3 ± 0.4		9.3 ± 0.1		−0.9 ± 0.8	−13.2 ± 1.3
	D	−6.4 ± 0.4	−19.0 ± 0.3	3.9 ± 0.7		9.4 ± 0.1		−2.4 ± 0.4	−14.5 ± 1.3

[a] The average favorable solvated structures of the drugs are used as the reference structures for the site and hydration free energy calculations. The numbers are reported in kcal/mol.

degree, as the calculations with several sets of constraining constants produce similar binding free energies around −13.6 kcal/mol. Nevertheless, we picked $k_c = 10$ kcal/mol/Å^2, $k_t = 10$ kcal/mol/Å^2, and $k_a = 200$ kcal/mol/rad^2 as the set of constraining constants, as it generates relatively low statistical errors for each contribution of the free energies. This set of force constants is consistent with the one applied by Deng et al.[45] The combination of $k_c = 10$ kcal/mol/Å^2, $k_t = 1$ kcal/mol/Å^2, and $k_a = 200$ kcal/mol/rad^2 provides lower binding free energy error, but the free energy component for repulsive interaction is increased which indicates that the low error in binding free energy might not be sustainable. Interestingly, the hardest ($k_c = 100$, $k_t = 100$, and $k_a = 2000$) and softest ($k_c = 0$, $k_t = 0$, and $k_a = 0$) constraints applied lead to the largest statistical deviations in binding free energy (−18.3 and −15.4 kcal/mol, respectively), reflecting two major problems one may face with constrained FEP simulations: under-sampled conformational space for substrate dynamics and over-restricted decreased conformational space that prohibits substrate dynamics in the site.

D. Effect of the Reference Structure Choice and RMSD Constraint Scheme. Up to this point we have used the average bound structure as the reference structure for the configurational constraint. The free energy component due to this constraint has a positive sign which indicates that there is a free energy loss upon binding due to the fact that the configurational freedom is restricted when one drug is bound to the receptor (moving from solvent to the protein binding pocket). A less natural, but still reasonable, reference structure is the lowest-energy structure for the hydrated drug. To find out the preferred conformation in the bulk solution for three drugs, we have performed an extensive replica exchange MD simulation.[68,69] Briefly, the trajectories of the replica exchange MD simulation of solvated drugs were used to obtain a probability distribution of conformations according to RMSD. The solvated structures of drugs with the most probable RMSDs were averaged to give the referenced average-solvated structure. The free energies calculated using these reference structures are listed in Table 4. The binding free energy changes from −13.6 to −17.9 for CMI, from −13.4 to −13.2 for IMI, and from −12.0 to −14.5 kcal/mol for DSI. CMI is still stands out as the most potent inhibitor for LeuT. DSI becomes more favorable than IMI, but the difference is within the statistical uncertainty (2.6 kcal/mol), considering that IMI and DSI have very similar binding affinity. However, the free energy differences between CMI and IMI are more pronounced and are significantly higher than that of estimated IC50 values reported from the experiment.[2]

It is interesting to examine the variation of the free energy components due to the choice of reference structures for the drugs (Figure 5). The sum of nonbonding components of the relative free energy ($G_{rep} + G_{disp} + G_{elec}$) becomes unfavorable when using the average solvated structure as the reference structure for conformational constraints. This is due to the fact that the reference structure is not the average bound structure, and the bound state differs substantially from that found in the bulk. The free energy component from the configurational constraint changes its sign to a negative value, compensating for the unfavorable constraint of restraining the ligand to the bulk-optimized conformation. Thus, by using the average solvated structure as a reference structure, we calculate the G_{int}(site) with an unfavorable bound structure, and we must rely the G_{conf}(site) component to correct the result. This is reflected by the standard mean deviations of the G_{conf} (site). When the average bound

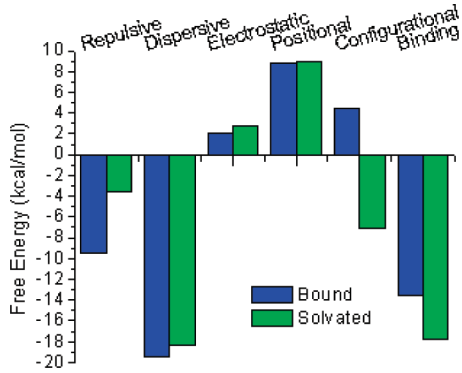


Figure 5. Column illustration of the binding free energy components of CMI binding to LeuT calculated with the average bound structure (blue) and the average favorable solvated structure (green) as the configurational reference of the RMSD constraint. From left to right: repulsive, dispersive, electrostatic, positional (including translational and rotational constraints), configurational, and total binding free energy.

structure is used, the standard deviation of G_{conf}(site) is 0.1 kcal/mol (Table 2), while when the favorable solvated structure is used, the standard deviation of G_{conf}(site) is about 0.3−0.4 kcal/mol (Table 4). Contrary to this, by using the average bound structure as a reference structure, we calculate G_{int}(solv) with an unfavorable solvated structure, and we must rely the G_{conf}(solv) component to correct the result. Since we usually have a more restricted space for the ligand in the binding pocket than in the bulk, G_{conf}(solv) should converge better than G_{conf}(site) (i.e., more overlapping between windows of umbrella sampling). Thus, using the average bound structure as a reference would usually be a better choice despite the fact that one could also get reasonable results with the average solvated structure as a reference.

To illustrate this compensation due to the use of different reference structures, we have evaluated the potential of mean force (PMF) as a function of the RMSD constraint. The PMFs of the configurational restraints for clomipramine are shown in Figure 6, where the reference structure of the restraint is the average bound structure in (a) and the average most probable solvated structure in (b). When the average bound structure is the reference structure, the free energy component resulting from the configurational restraint (4.4 kcal/mol) can be understood by a comparison of the site and solvent PMFs. The PMF of clomipramine in solvent goes up to about 4.4 kcal/mol at a RMSD corresponding to the minimum of the PMF in the binding site (∼0.2 Å). Similarly, when the average most probable solvated structure is the reference structure, the free energy component resulting from the configurational restraint (−7.2 kcal/mol) can also be intuitively understood. The PMF of clomipramine in the binding site goes up to about 7 kcal/mol at a RMSD, corresponding to the minimum of the PMF in solvent (∼0.7 Å). That is, in order to keep the configuration of bound structure around its most probable solvated structure, a free energy of 7 kcal/mol is required. Recently, Yang et al. incorporated the ligand reorganization free energy emphasizing its importance and even occasionally dominant contribution to high-affinity binding. It was shown that accounting for the free energy change for ligands between the free and bound states leads to better binding affinity prediction and enhanced the correlation coefficient for a number of studied complexes.[28] While they obtained the ligand reorganization free energy in the frame of MM-GBSA (molecular mechanics-generalized Born surface area), we want to point out here that a similar free energy component due to the reorganization of the ligand can also be robustly introduced within the FEP/MD absolute binding free energy calculation scheme.

E. Sensitivity to Potential Parameters. There are typically no molecular models of drugs available for the study of drug/protein binding free energy. For all-atom molecular simulations, it is a common practice to get the equilibrium structure from the crystal structure, and by using quantum mechanical calculations, determine the atomic charges and/or intramolecular potentials and obtain the remaining parameters from an available force field. An important concern is the assignment of atomic charges. Deng et al.[45] showed that with CHARMM, CHELPG, and AMSOL,[70] the calculated binding free energies could differ by about 3 kcal/mol for aromatic molecules binding to the T4 Lysozyme L99A mutant. In this report, we examine the dependence of calculated binding free energies on two charge sets: one initially guessed set based on CHARMM force field ($CS_{initial}$) and one charge set obtained from the RESP scheme described by Anisimov et al.[61] (CS_{fitted}). The method of charge fitting for CS_{fitted} is described in Section III. The atomic charges for the two charge sets are listed in Table S2 in the Supporting Information. The influence of using different charge sets (CS_{fitted} and $CS_{initial}$) was shown to be significant when calculating the hydration free energies of the TCAs (a difference of 6.2 kcal/mol for G_{tot}(solv) of DSI). On the other hand, both sets produced similar binding free energies for the ligands. Nevertheless, Table 2 shows that, from CS_{fitted} to $CS_{initial}$, the total binding free energy (in kcal/mol) changed from −13.6 to −11.0 for CMI, from −13.4 to −10.0 for IMI, and from −12.0 to −13.8 for DSI. For this particular binding site and these ligands, the different charge sets exhibit a difference up to 3.4 kcal/mol. Notably, the free energy order also changes, now contradicting the experimental results. Thus, for ligands without partial charge parameters in the available force field, the derivation of partial charges from high-level QM electrostatic potentials is highly preferred for molecular models, and validation against available experimental data is warranted.

F. Comparison between GSBP and PBC Simulations. Relative Binding Free Energies of TCAs Binding to LeuT. To compare possible artifacts due to reduction of the system with the GSBP scheme, we performed atomistic free energy simulations for the full system embedded into lipid bilayer using the FEP technique. The relative binding free energies can be calculated between two pairs: CMI/IMI and IMI/DSI from FEP simulations or simply from differences in absolute binding free energies obtained with GSBP simulations. The relative binding free energies between these compounds are known, and thus it is possible to correlate performance of two methods (GSBP and PBC) to experimental data. The structural difference between CMI and IMI is that the chlorine atom in CMI is replaced by a hydrogen

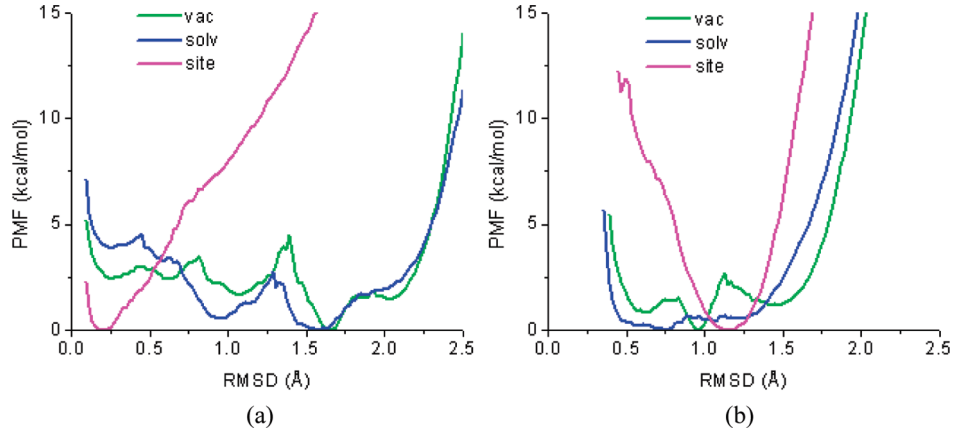


Figure 6. PMF of the configurational restraints on clomipramine. The PMFs of clomipramine in the binding site (site), solvent (solv), and vacuum (vac) are shown in magenta, blue, and green, respectively. The reference structure of clomipramine is the average bound structure (a) and the average favorable solvating structure (b).

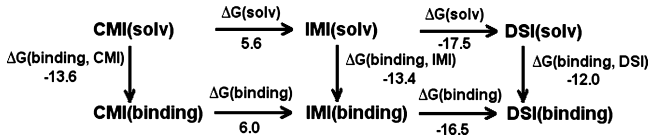


Figure 7. Free energy cycle of the binding of CMI, IMI, and DSI to LeuT. The binding free energies of CMI (ΔG(binding, CMI)), IMI (ΔG(binding, IMI)), and DSI (ΔG(binding, DSI)) are obtained from absolute binding free energy calculations. The relative free energies (ΔG(solv) and ΔG(binding)) are obtained from relative binding free energy calculations. The numbers are reported in kcal/mol.

atom in IMI. The difference between IMI and DSI is an addition of a methyl group on the tail (see Figure 2). The results are shown in Figure 7. The relative free energy for CMI to IMI is 6.0 kcal/mol for the bound state and 5.6 kcal/mol for the unbound state (in bulk), leading to a relative binding free energy of about 0.4 kcal/mol. The result is remarkably comparable to the difference of the absolute binding free energy of CMI to IMI, 0.2 kcal/mol obtained from absolute binding free energy simulations. The relative free energy for IMI to DSI is −16.5 and −17.5 kcal/mol for the bound and unbound states, respectively, leading to a relative binding free energy of 1.0 kcal/mol, which is also comparable to the difference of the absolute binding free energy of IMI to DSI of 1.4 kcal/mol. This, to some degree, justifies the use of GSBP for the absolute binding free energy calculations compared to the periodical boundary conditions used in the relative binding free energy calculations. The results are very interesting as one can use the more mature relative binding free energy calculation to check the results of absolute binding free energies when experimental binding affinities are not available. It also suggests that the reduced system provides an excellent and, more importantly, a relatively cheap test ground for rapid evaluation of relative free energies (for ranking of substrates). The thermodynamic cycle can be applied as an assessment tool to further the development of absolute binding free energy calculation methodologies.

G. Implications for the Molecular Mechanism of Antidepressant Binding to LeuT. Crystal structures of TCA-bound LeuT[2,3] show that the substrate and the drug binding sites are quite close to each other, mainly separated by a charged pair Arg 30 and Asp 404. Singh et al. demonstrated that the binding of the substrate and CMI might be thermodynamically coupled.[2] To explore this possibility, we calculated the absolute binding free energies of TCA's binding to LeuT without the bound leucine substrate. As crystal structures for such systems are not directly available, we obtained the starting structures by removing the substrates from the TCA-bound LeuT (2Q6H, 2Q72, 2QB4 after step B in section III) and allowing a further 2 ns relaxation and equilibration of the structures. The results are presented in Table 5. Generally, the binding free energies for all three drugs decreases when the substrates are removed: from −13.4 to −16.9 for CMI, from −13.2 to −15.0 for IMI, and from −12.0 to −13.4 kcal/mol for DSI. For the most bulky drug, CMI, the repulsive van der Waals (vdW) interaction is the main reason for the more favorable binding free energies. The free energy contribution due to repulsive vdW interactions changes from −9.4 for the substrate-bound LeuT (Table 2) to −17.0 kcal/mol for the substrate-free LeuT (Table 5). One explanation is that the removal of the substrate relaxes the protein structure, allowing for a more flexible drug binding pocket, facilitating the binding of CMI. The results here immediately infer that it is very difficult to compare the computed free energies to their corresponding experimental values, since experimental binding affinities reflect the binding free energies of a combination of LeuT with and without substrate.

Asp 401 is one of the key residues in the TCA binding pocket of LeuT.[2,3] The presence of the charge in this position is preserved in several NSS transporters. It is sometimes switched to a positively charged lysine or arginine and is thought to be functionally important. Its negatively charged side-chain carboxylate forms a salt bridge with the protonated side-chain nitrogen atom (Figure 2) in the "tail" of the bound TCAs.[2,3] To quantify the contribution of this interaction, we mutated the negatively charged Asp 401 to a positively charged lysine residue using the SCWRL3.0 program[71]

Table 5. Absolute Free Energy of Binding for Clomipramine (C), Imipramine (I), and Desipramine (D) to Substrate-Free LeuT in Reduced GSBP System

					G_{pos}				
		G_{rep}	G_{disp}	G_{elec}	G_{const}	$-k_BT\ \ln(F_tC^o)$	$-k_BT\ \ln(F_r)$	G_{conf}	G_{tot}
site	C	23.1 ± 1.0	-51.7 ± 0.5	-49.4 ± 0.2	-1.7 ± 0.1	5.3	6.5	-0.5 ± 0.1	-68.5 ± 1.1
	I	31.6 ± 1.8	-48.5 ± 0.6	-57.2 ± 0.3	-2.2 ± 0.1	5.5	6.4	-1.6 ± 0.1	-66.0 ± 1.7
	D	25.9 ± 0.9	-47.8 ± 0.6	-53.8 ± 0.3	-2.1 ± 0.1	5.4	6.4	-1.9 ± 0.3	-67.9 ± 0.4
binding	C	-17.0 ± 1.6	-16.8 ± 0.7	1.0 ± 0.3		10.0 ± 0.1		5.9 ± 0.8	-16.9 ± 1.5
	I	-6.9 ± 2.2	-14.8 ± 0.7	-5.8 ± 0.3		9.7 ± 0.1		2.8 ± 0.7	-15.0 ± 2.1
	D	-10.9 ± 0.7	-15.7 ± 0.4	2.2 ± 0.2		9.7 ± 0.1		1.3 ± 0.5	-13.4 ± 0.8

starting from the equilibrated, membrane embedded 2Q6H structure. We thus obtained the D401K mutant of LeuT with bound ions, substrate, and clomipramine. The calculated absolute binding free energy for CMI binding to LeuT−D401K is −9.9 kcal/mol with contributions of −17.2, −19.7, 13.2, 10.2, and 3.6 kcal/mol from the repulsive, dispersive, electrostatic, translational, rotational, and configurational components, respectively. Comparing CMI binding between the mutant and wild-type systems, the loss in the electrostatic interaction is 11.2 kcal/mol, primarily due to lack of interaction between CMI's fully protonated side-chain nitrogen atom and the negative side-chain of D401 residue in the wild-type LeuT. This loss is only partially compensated by the gains of −7.8 kcal/mol from the repulsive term. It may be due to the fact that the loss of the salt bridge allows CMI to rearrange in the binding pocket and avoid strong repulsive interactions. It should be noted that the overall effect of the mutation is a 3.7 kcal/mol less favorable for binding. This result has recently received surprising support from the combination of electrophysiological and biochemical studies on related GABA transporters. Cherubino et al. have reported that charge-switching mutations of lysine at the position 448 (K448E and K448D) that corresponds to 401 in LeuT lead to significant increase in the efficacy of desipramine,[72] further supporting functional role of negative charge in this location for high-affinity antidepressant binding.

V. Conclusion

In this report we used the free energy perturbation/molecular dynamics (FEP/MD) method with constraints to calculate the standard binding free energies for tricyclic antidepressant (TCA) binding to LeuT. The computed binding free energies are comparable to the experimental results. We showed that restraining potentials are essential and robust for enhancing sampling in studies of drugs binding to membrane proteins. The choice of the magnitude of the restraining forces on translation, rotation, and configuration are relatively robust (within 2 kcal/mol), as long as extreme values are avoided. For the configurational constraint, we tried two kinds of reference structures: the average bound structure and the average favorable solvated structure. It was shown that the use of the bound structure as the reference structure produced better results. Thus it is recommended for future applications of the FEP/MD method for the evaluation of absolute binding free energies. The use of the generalized solvent boundary potential (GSBP) approximation for studies of drug binding to membrane proteins also appears to be justified and accurate. We also showed that developing the molecular mechanics models (charge sets) for drugs using quantum mechanical electrostatic potential maps provided better accuracy. Interestingly, there is a notable compensation observed in the binding free energies between both molecular mechanics models. The absolute free energies (such as hydration and binding free energies) may vary significantly across models, while the relative binding free energies between drugs differ only by a small amount across models. This result is encouraging and in good agreement with similar conclusions from discussions on ion−protein interactions and force field development. We also showed that with the current FEP/MD method for absolute binding free energy, the results are compatible with those from the relative binding free energy calculations. We propose that the application of the free energy cycle can be applied to assess new methods of absolute binding free energy calculations. Finally, we showed that the absolute binding free energies for TCAs binding to LeuT are slightly different in the substrate-bound and substrate-free situations, indicating substrate-drug coupling as proposed by Singh et al.[2,3] Consistent with experimental indication, our FEP/MD calculation of the absolute binding free energy showed that the D401K mutation impairs the binding of clomipramine to LeuT, proving the essential role of the salt bridge between D401 of LeuT and the protonated nitrogen in the "tail" of the TCAs. In conclusion, this report shows that the use of the FEP/MD method for calculating absolute free energies of drugs bound to membrane proteins is a promising tool that can be used for drug design.

Acknowledgment. We are gratefully acknowledged discussions with Drs. Benoit Roux, Julia Subbotina, and Yuqing Deng. We are greatly indebted to Satinder Singh for guidance and discussion of the experimental work on LeuT and LeuT complexations with antidepressants. This work was supported by a Discovery Grant from the Natural Sciences and Engineering Council of Canada (NSERC) to S.N. S.N. is a CIHR New Investigator, an Alberta Heritage Foundation for Medical Research Scholar, and an Alberta Ingenuity New Scholar. The computational support for this work was provided by the West-Grid Canada through a resource allocation award to S.N. C.F.Z. is an AHFMR Post-Doctoral Fellow.

Supporting Information Available: Steps for the calculation of different components of the absolute free energy of a ligand binding to a receptor site. Molecular model for partial-charge assignment and atom labeling in clomi-

pramine. RESP (CS_{fitted}) and initial ($CS_{initial}$) partial charges for clomipramine, imipramine, and desipramine. Figure shows the binding pocket before and after the interactions between clomipramine (CMI) and the environment are fully turned off. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Yamashita, A.; Singh, S. K.; Kawate, T.; Jin, Y.; Gouaux, E. Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters. *Nature* **2005**, *437*, 215.

(2) Singh, S. K.; Yamashita, A.; Gouaux, E. Antidepressant binding site in a bacterial homologue of neurotransmitter transporters. *Nature* **2007**, *448*, 952.

(3) Zhou, Z.; Zhen, J.; Karpowich, N. K.; Goetz, R. M.; Law, C. J.; Reith, M. E. A.; Wang, D. N. LeuT-desipramine structure reveals how antidepressants block neurotransmitter reuptake. *Science* **2007**, *317*, 1390.

(4) Zhou, Z.; Zhen, J.; Karpowich, N. K.; Law, C. J.; Reith, M. E. A.; Wang, D. N. Antidepressant specificity of serotonin transporter suggested by three LeuT-SSRI structures. *Nat. Struct. Mol. Biol.* **2009**, *16*, 652.

(5) Rudnick, G. Serotonin transporters - Structure and function. *J. Membr. Biol.* **2006**, *213*, 101.

(6) Schafer, W. R. How do antidepressants work? Prospects for genetic analysis of drug mechanisms. *Cell* **1999**, *98*, 551.

(7) Quick, M.; Winther, A. M. L.; Shi, L.; Nissen, P.; Weinstein, H.; Javitch, J. A. Binding of an octylglucoside detergent molecule in the second substrate (S2) site of LeuT establishes an inhibitor-bound conformation. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 5563.

(8) Shi, L.; Quick, M.; Zhao, Y. F.; Weinstein, H.; Javitch, J. A. The mechanism of a neurotransmitter: sodium symporter - Inward release of Na+ and substrate is triggered by substrate in a second binding site. *Mol. Cell* **2008**, *30*, 667.

(9) Quick, M.; Javitch, J. A. Monitoring the function of membrane transport proteins in detergent-solubilized form. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 3603.

(10) Beveridge, D. L.; Dicapua, F. M. Free-Energy Via Molecular Simulation - Applications to Chemical and Biomolecular Systems. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431.

(11) Jorgensen, W. L. Free-Energy Calculations - a Breakthrough for Modeling Organic-Chemistry in Solution. *Acc. Chem. Res.* **1989**, *22*, 184.

(12) Sneddon, S. F.; Tobias, D. J.; Brooks, C. L. Thermodynamics of Amide Hydrogen-Bond Formation in Polar and Apolar Solvents. *J. Mol. Biol.* **1989**, *209*, 817.

(13) Jorgensen, W. L.; Severance, D. L. Aromatic Aromatic Interactions - Free-Energy Profiles for the Benzene Dimer in Water, Chloroform, and Liquid Benzene. *J. Am. Chem. Soc.* **1990**, *112*, 4768.

(14) Tobias, D. J.; Brooks, C. L. The Thermodynamics of Solvophobic Effects - a Molecular-Dynamics Study of Normal-Butane in Carbon-Tetrachloride and Water. *J. Chem. Phys.* **1990**, *92*, 2582.

(15) Kollman, P. Free-Energy Calculations - Applications to Chemical and Biochemical Phenomena. *Chem. Rev.* **1993**, *93*, 2395.

(16) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. The statistical-thermodynamic basis for computation of binding affinities: A critical review. *Biophys. J.* **1997**, *72*, 1047.

(17) Simonson, T.; Archontis, G.; Karplus, M. Free energy simulations come of age: Protein-ligand recognition. *Acc. Chem. Res.* **2002**, *35*, 430.

(18) Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. Absolute binding free energies: A quantitative approach for their calculation. *J. Phys. Chem. B* **2003**, *107*, 9535.

(19) Deng, Y. Q.; Roux, B. Computations of Standard Binding Free Energies with Molecular Dynamics Simulations. *J. Phys. Chem. B* **2009**, *113*, 2234.

(20) Zhou, H. X.; Gilson, M. K. Theory of Free Energy and Entropy in Noncovalent Binding. *Chem. Rev.* **2009**, *109*, 4092.

(21) Guvench, O.; MacKerell, A. D. Computational evaluation of protein-small molecule binding. *Curr. Opin. Struct. Biol.* **2009**, *19*, 55.

(22) Villoutreix, B. O.; Bastard, K.; Sperandio, O.; Fahraeus, R.; Poyet, J. L.; Calvo, F.; Deprez, B.; Miteva, M. A. In silico-in vitro screening of protein-protein interactions: Towards the next generation of therapeutics. *Curr. Pharm. Biotechnol.* **2008**, *9*, 103.

(23) Gane, P. J.; Dean, P. M. Recent advances in structure-based rational drug design. *Curr. Opin. Struct. Biol.* **2000**, *10*, 401.

(24) Gilson, M. K.; Zhou, H. X. Calculation of protein-ligand binding affinities. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21.

(25) Gohlke, H.; Hendlich, M.; Klebe, G. In *Knowledge-based scoring function to predict protein-ligand interactions*. *J. Mol. Biol.* **2000**, *295*, 337.

(26) Sondergaard, C. R.; Garrett, A. E.; Carstensen, T.; Pollastri, G.; Nielsen, J. E. Structural Artifacts in Protein-Ligand X-ray Structures: Implications for the Development of Docking Scoring Functions. *J. Med. Chem.* **2009**, *52*, 5673.

(27) Klebe, G. Recent developments in structure-based drug design. *J. Mol. Med.* **2000**, *78*, 269.

(28) Yang, C. Y.; Sun, H. Y.; Chen, J. Y.; Nikolovska-Coleska, Z.; Wang, S. M. Importance of Ligand Reorganization Free Energy in Protein-Ligand Binding-Affinity Prediction. *J. Am. Chem. Soc.* **2009**, *131*, 13709.

(29) Wang, J. M.; Morin, P.; Wang, W.; Kollman, P. A. Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. *J. Am. Chem. Soc.* **2001**, *123*, 5221.

(30) Bashford, D.; Case, D. A. Generalized born models of macromolecular solvation effects. *Annu. Rev. Phys. Chem.* **2000**, *51*, 129.

(31) Swanson, J. M. J.; McCammon, J. A. Applying the statistical mechanics behind ligand-receptor binding affinities. *Biophys. J.* **2003**, *84*, 341A.

(32) Minh, D. D. L.; Bui, J. M.; Chang, C. E.; Jain, T.; Swanson, J. M. J.; McCammon, J. A. The entropic cost of protein-protein association: A case study on acetylcholinesterase binding to fasciculin-2. *Biophys. J.* **2005**, *89*, L25.

(33) Wright, J. D.; Noskov, S. Y.; Lim, C. Factors governing loss and rescue of DNA binding upon single and double mutations in the p53 core domain. *Nucleic Acids Res.* **2002**, *30*, 1563.

(34) Noskov, S. Y.; Lim, C. Free energy decomposition of protein-protein interactions. *Biophys. J.* **2001**, *81*, 737.

(35) Wang, J. Y.; Deng, Y. Q.; Roux, B. Absolute binding free energy calculations using molecular dynamics simulations with restraining potentials. *Biophys. J.* **2006**, *91*, 2798.

(36) Irudayam, S. J.; Henchman, R. H. Entropic Cost of Protein-Ligand Binding and Its Dependence on the Entropy in Solution. *J. Phys. Chem. B* **2009**, *113*, 5871.

(37) Kirkwood, J. G. Statistical mechanics of fluid mixtures. *J. Chem. Phys.* **1935**, *3*, 300.

(38) Roux, B. The Calculation of the Potential of Mean Force Using Computer-Simulations. *Comput. Phys. Commun.* **1995**, *91*, 275.

(39) Frenkel, D.; Smit, B. *Understanding molecular simulation: from algorithms to applications*, 2nd ed.; Academic Press: San Diego, CA, 2002; p xxii.

(40) Darve, E.; Pohorille, A. Calculating free energies using average force. *J. Chem. Phys.* **2001**, *115*, 9169.

(41) Gullingsrud, J. R.; Braun, R.; Schulten, K. Reconstructing potentials of mean force through time series analysis of steered molecular dynamics simulations. *J. Comput. Phys.* **1999**, *151*, 190.

(42) Jarzynski, C. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **1997**, *78*, 2690.

(43) Ensing, B.; De Vivo, M.; Liu, Z. W.; Moore, P.; Klein, M. L. Metadynamics as a tool for exploring free energy landscapes of chemical reactions. *Acc. Chem. Res.* **2006**, *39*, 73.

(44) Christi, C. D.; Mark, A. E.; van Gunsteren, W. F. Basic ingredients of free energy calculations: A review. *J. Comput. Chem.* **2010**, *31*, 1569.

(45) Deng, Y. Q.; Roux, B. Calculation of standard binding free energies: Aromatic molecules in the T4 lysozyme L99A mutant. *J. Chem. Theory Comput.* **2006**, *2*, 1255.

(46) Brandsdal, B. O.; Osterberg, F.; Almlof, M.; Feierberg, I.; Luzhkov, V. B.; Aqvist, J. Free energy calculations and ligand binding. *Protein Simul.* **2003**, *66*, 123.

(47) Nervall, M.; Hanspers, P.; Carlsson, J.; Boukharta, L.; Aqvist, J. Predicting binding modes from free energy calculations. *J. Med. Chem.* **2008**, *51*, 2657.

(48) Archontis, G.; Watson, K. A.; Xie, Q.; Andreou, G.; Chrysina, E. D.; Zographos, S. E.; Oikonomakos, N. G.; Karplus, M. Glycogen phosphorylase inhibitors: A free energy perturbation analysis of glucopyranose spirohydantoin analogues. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 984.

(49) Im, W.; Berneche, S.; Roux, B. Generalized solvent boundary potential for computer simulations. *J. Chem. Phys.* **2001**, *114*, 2924.

(50) Noskov, S. Y. Molecular mechanism of substrate specificity in the bacterial neutral amino acid transporter LeuT. *Proteins: Struct., Funct., Bioinf.* **2008**, *73*, 851.

(51) Straatsma, T. P.; Zacharias, M.; Mccammon, J. A. Holonomic Constraint Contributions to Free-Energy Differences from Thermodynamic Integration Molecular-Dynamics Simulations. *Chem. Phys. Lett.* **1992**, *196*, 297.

(52) Boresch, S. The role of bonded energy terms in free energy simulations - Insights from analytical results. *Mol. Simul.* **2002**, *28*, 13.

(53) Deng, Y. Q.; Roux, B. Hydration of amino acid side chains: Nonpolar and electrostatic contributions calculated from staged molecular dynamics free energy simulations with explicit water molecules. *J. Phys. Chem. B* **2004**, *108*, 16567.

(54) Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method 0.1. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22*, 1420.

(55) Straatsma, T. P.; Mccammon, J. A. Computational Alchemy. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407.

(56) Deng, Y. Q.; Roux, B. Computation of binding free energy with molecular dynamics and grand canonical Monte Carlo simulations. *J. Chem. Phys.* **2008**, 128.

(57) Weeks, J. D.; Chandler, D.; Andersen, H. C. Role of Repulsive Forces in Determining Equilibrium Structure of Simple Liquids. *J. Chem. Phys.* **1971**, *54*, 5237.

(58) Caplan, D. A.; Subbotina, J. O.; Noskov, S. Y. Molecular Mechanism of Ion-Ion and Ion-Substrate Coupling in the Na+-Dependent Leucine Transporter LeuT. *Biophys. J.* **2008**, *95*, 4613.

(59) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. Charmm - a Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187.

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

(61) Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V.; Huang, N.; Roux, B.; MacKerell, A. D. Determination of electrostatic parameters for a polarizable force field based on the classical Drude oscillator. *J. Chem. Theory Comput.* **2005**, *1*, 153.

(62) Connolly, M. L. Solvent-Accessible Surfaces of Proteins and Nucleic-Acids. *Science* **1983**, *221*, 709.

(63) Cantu, M. D.; Hillebrand, S.; Carrilho, E. Determination of the dissociation constants (pK(a)) of secondary and tertiary amines in organic media by capillary electrophoresis and their role in the electrophoretic mobility order inversion. *J. Chromatogr., A* **2005**, *1068*, 99.

(64) Shalaeva, M.; Kenseth, J.; Lombardo, F.; Bastinz, A. Measurement of dissociation constants (pK(a) values) of organic compounds by multiplexed capillary electrophoresis using aqueous and cosolvent buffers. *J. Pharm. Sci.* **2008**, *97*, 2581.

(65) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules 0.1. The Method. *J. Comput. Chem.* **1992**, *13*, 1011.

(66) Beglov, D.; Roux, B. Finite Representation of an Infinite Bulk System - Solvent Boundary Potential for Computer-Simulations. *J. Chem. Phys.* **1994**, *100*, 9050.

(67) Mobley, D. L.; Dill, K. A. Binding of Small-Molecule Ligands to Proteins: "What You See" Is Not Always "What You Get. *Structure* **2009**, *17*, 489.

(68) Swendsen, R. H.; Wang, J. S. Replica Monte-Carlo Simulation of Spin-Glasses. *Phys. Rev. Lett.* **1986**, *57*, 2607.

(69) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141.

(70) Li, J. B.; Zhu, T. H.; Cramer, C. J.; Truhlar, D. G. New class IV charge model for extracting accurate partial charges from wave functions. *J. Phys. Chem. A* **1998**, *102*, 1820.

(71) Canutescu, A. A.; Shelenkov, A. A.; Dunbrack, R. L. A graph-theory algorithm for rapid protein side-chain prediction. *Protein Sci.* **2003**, *12*, 2001.

(72) Cherubino, F.; Miszner, A.; Renna, M.; Sangaletti, R.; Giovannardi, S.; Bossi, E. GABA transporter lysine 448: a key residue for tricyclic antidepressants interaction. *Cell. Mol. Life Sci.* **2009**, *66*, 3797.

CT9006597