Research Article

Journal of Molecular Recognition

Received: 1 September 2009, Accepted: 7 September 2009, Published online in Wiley InterScience: 2010

(www.interscience.wiley.com) DOI:10.1002/jmr.996

Calculation of the standard binding free energy of sparsomycin to the ribosomal peptidyl-transferase P-site using molecular dynamics simulations with restraining potentials

Xiaoxia Ge[a] and Benoît Roux[b,c]*

The standard (absolute) binding free energy of the antibiotic sparsomycin with the 50S bacteria ribosomal subunit is calculated using molecular dynamics (MD) free energy perturbation (FEP) simulations with restraining potentials developed by Wang *et al.* [Biophys. J. 91:2798–2814 (2006)]. In the simulation protocol, restraining potentials are activated for the orientational and translational movements of the ligand relative to the binding site when it is decoupled from the binding pocket, and then released once the ligand fully interacts with the rest of the system. A reduced system is simulated to decrease the computational cost of the FEP/MD calculations and the effects of the surrounding atoms are incorporated using the generalized solvent boundary potential (GSBP) method. The loss of conformational freedom of the ligand upon binding is characterized using the potential of mean force (PMF) as a function of the root-mean-square deviation (RMSD) relative to the bound conformation. The number of water molecules in the binding pocket is allowed to fluctuate dynamically in response to the ligand during the calculations by combining FEP/MD with grand canonical Monte Carlo (GCMC) simulations. The calculated binding free energy is about −6 kcal/mol, which is in reasonable agreement with the experimental value. The information gleaned from this study provides new insight on the recognition of ribosome by sparsomycin and highlights the challenges in calculations of absolute binding free energies in these systems. Copyright © 2010 John Wiley & Sons, Ltd.

Keywords: simulation; electrostatics; counterions

INTRODUCTION

Ribosomes are large macromolecular assemblies comprising about 65% ribosomal RNAs and 35% ribosomal proteins. Their molecular mass is about 2.5 MDa in bacteria. Because their active sites are made of RNAs, ribosomes are classified as "ribozymes" (Rodnina *et al.*, 2007). The ribosome is the target of nearly half of antibiotics characterized to date. In recent years, the determination of high-resolution structures of ribosomal subunits and even of the entire ribosomes by X-ray crystallography has revolutionized our understanding of the role of the ribosome in translation and the mechanism underlying the inhibition of protein synthesis by antibiotics (Klein *et al.*, 2001; Schuwirth *et al.*, 2005; Korostelev *et al.*, 2006; Selmer *et al.*, 2006).

Ribosomes consist of two nonequivalent ribonucleoprotein subunits that associate with each other non-covalently. When peptide elongation proceeds, both subunits operate in a concerted fashion. The small subunit serves as the mRNA-binding machinery and the decoding center. It provides the path for mRNA progression and controls the translation fidelity (Stern *et al.*, 1989; Yusupova *et al.*, 2001). The large subunit performs the main ribosomal catalytic function in the peptidyl-transferase center (PTC) and provides the peptide-exit tunnel. The two subunits are joined together at the three tRNA binding sites, i.e. A (aminoacyl), P (peptidyl), and E (exit) sites, where tRNA molecules decode the genetic information and carry the corresponding amino acids to be incorporated into the growing peptide chain (Moazed and Noller, 1989; Noller, 2005). Several antibiotics, such as sparsomycin, are known to bind at one or two of the tRNA binding sites in either of the subunits (Woodcock *et al.*, 1991; Hansen *et al.*, 2003).

Some biochemical experiments (Vandenbroek *et al.*, 1987, 1989) have provided the evidence of sparsomycin binding to the peptidyl-transferase P-site. It was proposed that sparsomycin could enhance the affinity of tRNA for the P-site and consequently halt the protein synthesis by blocking the translocation. The binding site of sparsomycin in the 50S subunit has been determined both by the labs of Steitz (Hansen *et al.*, 2002) and Yonath (Bashan *et al.*, 2003) for different species at various levels

* *Correspondence to: B. Roux, Department of Biochemistry and Molecular Biology, The University of Chicago, 929 East 57th Street, Chicago, IL, 60637, USA.*
E-mail: roux@uchicago.edu

a *X. Ge*
Department of Physiology and Biophysics, Weill Medical College of Cornell University, New York, NY, USA

b *B. Roux*
Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, IL, USA

c *B. Roux*
Bioscience Division, Argonne National Laboratory, Argonne, IL, USA

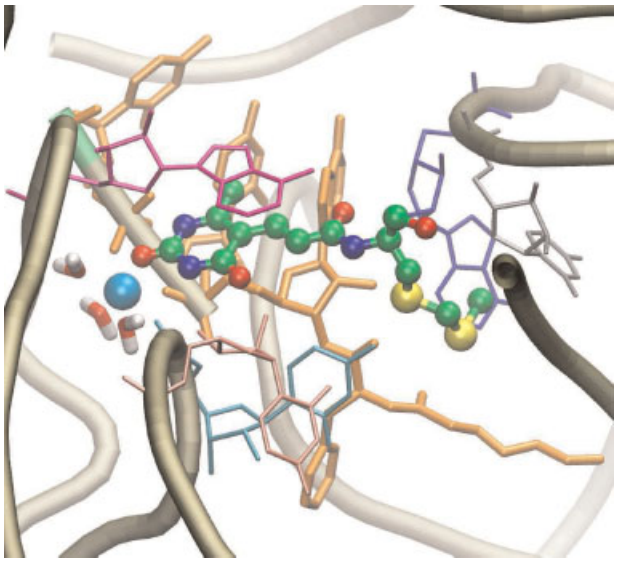

Figure 1. Binding site of the sparsomycin–ribosome complex in the crystal structure 1M90. The backbones of the ribosomal RNA are represented by tubes in tan. Sparsomycin of 1M90 (colored balls and sticks) is aligned in the center of the image, CCA end of P-site substrate analogue and phe–cap–biotin are represented with sticks in orange. Base A2637 (magenta lines) stacks on the top of the sparsomycin. Mg^{2+} (cyan sphere) interacts with three water molecules (CPK sticks) regenerated by the GCMC/MD equilibration. The active-site hydrophobic crevice is formed between A2486 (blue lines) and C2487 (gray lines), U2619 (pink lines), and U2620 (cyan lines) interact with the pseudouracil ring of sparsomycin through H-bonding.

of resolutions. They both found that sparsomycin was stacked onto the same rRNA base, although on opposite sides of the base. The implication of such finding is that sparsomycin's primary function relies on the interactions with the P-site tRNA. In the recently published high resolution structure (Schmeing *et al*., 2005), sparsomycin is packed along the backbone of C75 and A76 of the P-site tRNA and its peptide, and the uracil-like moiety of sparsomycin stacks directly on the universally conserved nucleobase A2637 (Figure 1). The crystal structures provide a structural basis for the ribosomal inhibition by antibiotics sparsomycin. Despite the extensive biochemical and structural studies, the driving forces for the sparsomycin-ribosome interactions are not elucidated.

The accurate calculation of absolute binding free energy is one of the most important goals of computational biophysics. The emerging successes reported in computing the binding free energy of small ligands to proteins using molecular dynamics (MD) simulations indicate that physics-based approaches hold the promise of expediting the drug discovery process (Deng and Roux, 2009). Here we report a free energy simulation study of the ribosome–sparsomycin binding interactions. The standard (absolute) binding free energy of sparsomycin to the bacterial ribosomal large subunit is calculated using alchemical free energy perturbation molecular dynamics (FEP/MD) simulations with restraining potentials (Wang *et al*., 2006a). Such restraining potentials are activated and released at different stages during the FEP/MD simulation, in order to more efficiently sample the changes in translational, rotational, and conformational freedom of the ligand and the receptor upon binding. As long as their effects are properly accounted for, the total reversible work between the end-points and the calculated free energy are unaffected by the restraints introduced in the intermediate stages of the FEP/MD simulations (Deng and Roux, 2009).

To decrease the prohibitive cost of FEP calculations, a reduced system of the sparsomycin–ribosome complex is simulated and the influence of the surrounding is incorporated using the generalized solvent boundary potential (GSBP, Im *et al*., 2001) and the ligand in the bulk solution is simulated using the spherical solvent boundary potential (SSBP, Beglov and Roux, 1994). While MD simulations of complete atomic models of a solvated ribosome are certainly possible (Sanbonmatsu and Tung, 2007), such models are extremely large for converged and accurate FEP/MD simulations. GSBP and SSBP are attractive strategies, which consist of treating only a small number of atoms in the vicinity of the ligand explicitly, while representing the influence of the surrounding via an effective potential. This type of hybrid MD/continuum approach is justified because the binding affinity is dominated by local interactions in the vicinity of the ligand and the remote regions are expected to contribute only in an average manner. The grand canonical Monte Carlo (GCMC) algorithm is incorporated into the FEP/MD simulation to allow the dynamical response of the number of the water molecules in the buried pocket of the ribosomal P-site during the calculations. To quantify the loss of conformational freedom of sparsomycin upon binding, a potential of mean force (PMF) is calculated as a function of the root-mean-square deviation (RMSD) of the ligand relative to its bound conformation.

The theories and methods for the FEP/MD simulations are given in the following sections. Then, the results are discussed in the light of the hydration of the binding pocket, the importance of translational, orientational, and conformational free energy contributions. Finally, the paper is concluded with a short summary of the main results and a perspective on future studies of antibiotic binding to the ribosome.

THEORY AND METHODS

Equilibrium binding constant

The methods for computing binding free energies from atomistic MD simulations are now clearly established (see Deng and Roux, 2009 for a recent review). The equilibrium constant K_b of the binding reaction $L + R \rightleftharpoons LR$ is defined as $K_b = [LR]/([L][R])$, where [L], [R], and [LR] are the equilibrium concentrations of the unbound ligand, unbound receptor, and bound complex, respectively. The standard binding free energy follows from the equilibrium constant by $\Delta G_b^\circ = -k_B T \ln(C^\circ K_b)$, where C° is the standard concentration ($1\,\text{mol/L} = 1/1660\,\text{Å}^3$), k_B is the Boltzmann constant, and T is the absolute temperature. Following the statistical mechanical formulation from Wang *et al*., (2006a) the macroscopic standard binding free energy and the microscopic term equilibrium constant are related by

$$K_b = \frac{\int_{site} dL \int dX e^{-U/k_B T}}{\int_{bulk} dL \delta(r_L - r^*) \int dX e^{-U/k_B T}} \quad (1)$$

where U is the total potential energy of the system, L and X are the coordinates of the ligand and the remaining atoms (solvent and receptor), respectively, r_L is the position of the center-of-mass of ligand L, and r^* is some arbitrary location in the bulk region (Deng and Roux, 2009). The delta function $\delta(r_L - r^*)$ in the denominator arises from the translational invariance of the ligand in the bulk volume. Therefore, K_b has the dimension of volume

(typically Å^3 in atomistic MD simulations). Multiplying K_b by the standard concentration C° yields the dimensionless quantity K_bC°, from which a meaningful binding free energy ΔG_b° can be defined.

In the case of ligand bound in a deeply buried pocket such as the sparsomycin–ribosome complex, treating the free energy computation on the basis of the reaction path for the bimolecular association can be very challenging. The most effective strategy for computing the binding free energy for such a system is alchemical FEP/MD, in which the interactions between the ligand and its surroundings are switched "on" and "off" (Deng and Roux, 2009). To control the translation of the non-interacting "decoupled" ligand relative to the receptor binding site, a restraining potential is introduced at one end-point to "confine" the uncoupled ligand within the binding site, and is then "released" at the other end-point (Hermans and Subramaniam, 1986; Roux *et al*., 1996; Boresch *et al*., 2003). This protocol is known as the "double decoupling method" (DDM, Gilson *et al*., 1997). For computational convenience, the total free energy of binding is expressed as the reversible thermodynamic work needed for sequentially switching on the interaction of the ligand with its surrounding (receptor and solvent) as well as various restraining potentials in a step-by-step process (Wang *et al*., 2006a; Deng and Roux, 2009). It follows that the standard binding free energy obtained from the stepwise transformations can be expressed as

$$\Delta G_b^\circ = \Delta\Delta G_{int} + \Delta\Delta G_t^\circ + \Delta\Delta G_r + \Delta\Delta G_c \quad (2)$$

where $\Delta\Delta G_{int} = [\Delta G_{int}^{site} - \Delta G_{int}^{bulk}]$ corresponds to the interaction (int) free energy difference associated with removing the ligand from the bulk and inserting it in the binding site, $\Delta\Delta G_c = [\Delta G_c^{bulk} - \Delta G_c^{site}]$ corresponds to the free energy cost for restricting the conformational (c) freedom of the ligand upon binding, while the free energy cost associated with the introduction and removal of the translational (t) and rotational (r) restraints on the ligand are $\Delta\Delta G_t^\circ = [-k_BT\ln(F_tC^\circ) - \Delta G_t^{site}]$ and $\Delta\Delta G_r = [-k_BT\ln(F_r) - \Delta G_r^{site}]$, respectively, (*n.b.* the standard concentration C° cancels the unit in F_t). The translational and orientational factors F_t and F_r correspond to simple numerical integrals that have been defined previously (Deng and Roux, 2006; Wang *et al*., 2006a) by Equations (3) and (4)

$$F_t = \int_0^\infty dr_L r_L^2 \int_0^\pi d\theta_L \sin(\theta_L) \int_{-\pi}^{\pi} d\phi_L e^{-u_t(r_L,\theta_L,\phi_L)/k_BT} \quad (3)$$

and

$$F_r = \frac{1}{8\pi^2}\int_0^\pi d\alpha_L \sin(\alpha_L) \int_{-\pi}^{\pi} d\beta_L \int_{-\pi}^{\pi} d\gamma_L e^{-u_r(\alpha_L,\beta_L,\gamma_L)/k_BT} \quad (4)$$

where

$$u_t = \frac{1}{2}[k_t(r_L - r_0)^2 + k_a(\theta_L - \theta_0)^2 + k_a(\phi_L - \phi_0)^2] \quad (5)$$

and

$$u_r = \frac{1}{2}[k_a(\alpha_L - \alpha_0)^2 + k_a(\beta_L - \beta_0)^2 + k_a(\gamma_L - \gamma_0)^2] \quad (6)$$

define quadratic translational and orientational restraining potentials, k_t is the force constant for the distance restraint, and k_a is the force constants for the angle and dihedral restraints; r_0, θ_0, ϕ_0, α_0, β_0, and γ_0 are the reference values of the distance, angle, and dihedral determined from an average of the equilibration trajectory and subsequently used to define the position and the orientation of the ligand.

This formulation decomposes the binding free energy into contributions from individual factors. Although the values of the individual contributions depend on the details of the restraint potentials, the total standard binding free energy is rigorously independent of the restraints. From this formulation of the standard binding free energy, approximations like linear interaction energy (LIE, Åqvist *et al*., 2002) and MMPB-SA (Kollman *et al*., 2000; Simonson *et al*., 2002; Swanson *et al*., 2004) can be made. The free energy components provide insights on what determines the binding process (Hermans and Wang, 1997; Wang *et al*., 2006a). For example, the contributions from the loss of translational, rotational, and conformational freedom, i.e. $\Delta\Delta G_t^\circ$, $\Delta\Delta G_r$, and $\Delta\Delta G_c$, respectively, are always positive and oppose the binding reaction.

Further insights into the microscopic factors driving ligand binding are provided when the binding free energy is separated in terms of the most salient intermolecular forces. The latter are typically characterized by electrostatic interactions, harsh short-range repulsion, and more slowly varying van der Waals attractions. For this reason, it is advantageous to stage the decoupling alchemical FEP/MD simulations accordingly with a proper choice of intermediate states by exploiting the distinct characters in spatial range and magnitude of these various contributions. The separation of the Lennard-Jones (LJ) 12–6 potential in terms of repulsive and attractive interactions is achieved by using the Weeks-Chandler-Andersen (WCA) scheme (Weeks *et al*., 1971). This has been shown to provide a clear separation of the competing interactions in the solvation process, and to robustly enhance the efficiency of FEP/MD calculations (Deng and Roux, 2004). Accordingly, the interaction free energies are further decomposed into three components,

$$\Delta\Delta G_{int} = \Delta\Delta G_{rep} + \Delta\Delta G_{dis} + \Delta\Delta G_{elec} \quad (7)$$

where $\Delta\Delta G_{rep} = [\Delta G_{rep}^{site} - \Delta G_{rep}^{bulk}]$, $\Delta\Delta G_{dis} = [\Delta G_{dis}^{site} - \Delta G_{dis}^{bulk}]$, and $\Delta\Delta G_{elec} = [\Delta G_{elec}^{site} - \Delta G_{elec}^{bulk}]$ correspond to the free energy differences in terms of repulsive (rep), dispersive (dis), and electrostatic (elec) interactions associated with removing the ligand from the bulk and inserting it in the binding site. The separation of the binding free energy according to this formulation helps clarify the important aspects of the binding process, which might otherwise remain hidden.

PMF treatment of ligand conformation

The ligand bound to the receptor is expected to be more restricted than when it is in the bulk. Such restriction is often associated with a loss of conformational entropy, which counteracts the favorable interactions between the ligand and the receptor. To quantitatively account for the loss of conformational freedom of ligand upon binding, we use a deliberate sampling strategy based on a restraining potential (Woo and Roux, 2005). A conformational restraining potential $u_c(\xi)$ is applied to sparsomycin. This restraint $u_c(\xi)$ is a harmonic potential with respect to ξ, the RMSD of the ligand relative to a reference conformation. The conformation of the ligand in the bound state is chosen as reference. To improve the accuracy, the free energies associated with the conformational restriction of the ligand near the reference conformation, ΔG_c^{bulk} and ΔG_c^{site}, are calculated by an explicit integration of the Boltzmann factor of the PMF as a function of RMSD obtained from umbrella sampling simulations (Woo and Roux, 2005). The free energy cost ΔG_c^{bulk} and ΔG_c^{site} associated with restraining the conformation of the

Journal of Molecular Recognition

ligand in the bulk or in the binding site by the potential u_c are calculated as

$$\Delta G_c^{\text{bulk}} = -k_B T \ln \left[\frac{\int d\xi e^{-W_c^{\text{bulk}}(\xi)/k_B T} e^{-u_c(\xi)/k_B T}}{\int d\xi e^{-W_c^{\text{bulk}}(\xi)/k_B T}} \right] \quad (8)$$

and

$$\Delta G_c^{\text{site}} = -k_B T \ln \left[\frac{\int d\xi e^{-W_c^{\text{site}}(\xi)/k_B T} e^{-u_c(\xi)/k_B T}}{\int d\xi e^{-W_c^{\text{site}}(\xi)/k_B T}} \right] \quad (9)$$

where $W_c^{\text{bulk}}(\xi)$ and $W_c^{\text{site}}(\xi)$ are the RMSD–PMFs of the ligand in the bulk and in the binding site, respectively (Woo and Roux, 2005; Deng and Roux, 2006; Wang *et al*., 2006a). They are calculated with biased simulations via umbrella sampling (see below).

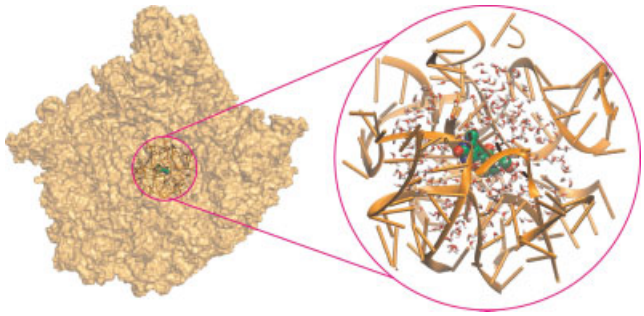

Figure 2. Reduced simulation system. The left panel shows the crystal structure 1M90, with the sparsomycin in CPK sphere, the RNAs and protein in the GSBP sphere in cartoon (orange) and the rest of the system in transparent contour surface for RNAs and proteins (orange). The right panel shows the simulation region, including the GSBP inner sphere (yellow cartoon) and explicit solvent water (sticks).

COMPUTATIONAL DETAILS

Atomic system and potential function

The initial input structure for simulation is the crystal structure of ribosome 50S in complex with sparsomycin from the Brookhaven Protein Data Bank (PDB ID code 1M90) (Hansen *et al*., 2003). 1M90 was the first available crystal structure of the 50S subunit of *H. marismortui* in complex with antibiotic sparsomycin and CCA–Phe–caproic acid–biotin (CAP) at 2.8 Å resolution. Another crystal structure (PDB ID code 1VQ9) was reported for the same system more recently (Schmeing *et al*., 2005). The N-terminal of Phe–CAP was patched to the 3′ end of CCA. The hydrogen atoms were added using the HBUILD (Brunger and Karplus, 1988) facility in CHARMM (Brooks *et al*., 1983). The complete 50S structure contains about 150 000 atoms, without counting the water molecules. Such an atomic model would render FEP/MD simulations computationally prohibitive.

In the present study, a reduced system representing sparsomycin bound to the ribosome was simulated using GSBP (Im *et al*., 2001). A reduced system representing sparsomycin in bulk solution was simulated using SSBP (Beglov and Roux, 1994). SSBP and GSBP are hybrid MD/continuum approaches in which the system is divided into inner and outer regions. Both methods incorporate the influence of the remote regions by using continuum electrostatic approximations. For the bound complex, the atoms within a sphere of 15 Å radius centered on the ligand were defined as the inner GSBP region, yielding a reduced system of about 4100 atoms. The reduced simulation system, constructed from the 50S subunit (1M90), is shown in Figure 2. Although the complete ribosome structures are available (Schuwirth *et al*., 2005; Korostelev *et al*., 2006; Selmer *et al*., 2006), the functional sites are similar, as revealed by cross-crystal averaging (Simonovc and Steitz, 2008). Since the closest atoms of the 30S subunit are more than 25 Å away from the center of the sparsomycin binding site, building the reduced GSBP model from the 50S subunit is not expected to affect the results. For the ligand in bulk solution, a system of 1000 explicit water molecules was simulated with SSBP (Beglov and Roux, 1994), for a total of about 3000 atoms.

The inner region was extended by 3 Å to define a smooth spherical dielectric cavity. The ribosome atoms located in the 3 Å shell, and those within the extended inner spherical region ($\leq$18 Å) linked via 1–3 bonds with the outer region ($>$18 Å) were fixed according to a group-based criterion. All the other atoms including the ribosome, sparsomycin, and solvent molecules in the inner region were allowed to move dynamically. To avoid a straightforward truncation of the large system, which would lead to inaccurate results due to the neglect of long-range electrostatic effects, the influence of the surrounding outer region on the atoms of the inner region was incorporated in the form of a solvent-shielded static field, and a solvent-induced reaction field. The reaction field caused by changes in charge distribution of the dynamic inner region is expressed in terms of a basis set expansion of the inner simulation region charge density. The basis set coefficients correspond to generalized electrostatic multipoles. Here, a basis set of nine spherical harmonic functions was used for the sparsomycin–ribosome complex. The solvent-shielded static field and the reaction field matrix, representing the couplings between the generalized multipoles, were both invariant with respect to the configuration of the explicit atoms in the inner simulation region. They were calculated once with finite-difference Poisson–Boltzmann (PB), assuming dielectric constants of four inside the ribosome and 80 outside. The atomic Born radii for protein and nucleic acid atoms used to setup the dielectric boundaries in the PB calculations were determined by free energy simulations with explicit solvent (Nina *et al*., 1997; Banavali and Roux, 2002). To mimic the physiological condition, a salt concentration of 0.2 M was used in the PB calculations. The water molecules within the inner region are confined by a nonpolar cavity potential to prevent entry into the surrounding dielectric continuum.

The CHARMM27 force field was used for the ribosome structure, including RNAs, proteins, and ion species (MacKerell *et al*., 1998; Foloppe and MacKerell, 2000; MacKerell and Banavali, 2000). The water molecules were represented by the TIP3P model (Jorgensen *et al*., 1983). The topology and parameter files used to represent the potential function of sparsomycin were generated using the program ANTECHAMER1.27 within the framework of the Generalized AMBER force field (GAFF, Wang *et al*., 2006b; Wang *et al*., 2004). Atomic partial charges were determined using the Fitcharge module in CHARMM (Anisimov *et al*., 2005) obtained by optimizing the charge fitting to the *ab initio* quantum mechanical (QM) electrostatic potential. The QM electrostatic map for sparsomycin was calculated at the B3LYP level with the *aug-cc-p*VDZ basis set (Lee *et al*., 1988; Becke, 1993; Hertwig and Koch, 1995) using Gaussian 03 (Frisch *et al*., 2004). The atomic charges and nonbonded parameters of the

Table 1. Sparsomycin atomic partial charges

Atom ID	GAFF type	Partial charge	Atom ID	GAFF type	Partial charge	Atom ID	GAFF type	Partial charge
C1	C	0.614	C8	C3	−0.192	H13	HO	0.396
O1	O	−0.506	H8	HC	0.143	C14	C3	−0.501
N2	N	−0.665	C9	C3	−0.258	H141	HC	0.226
H2	HN	0.376	H9	HC	0.119	H142	HC	0.226
C3	C	0.778	C10	C	0.622	S15	S4	0.671
O3	O	−0.539	O10	O	−0.540	O15	O	−0.557
N4	N	−0.661	N11	N	−0.663	C16	C3	−0.641
H4	HN	0.380	H11	HN	0.302	H161	HC	0.249
C5	CA	0.355	C12	C3	0.055	H162	HC	0.249
C7	C3	−0.536	H12	HC	0.271	S17	SS	−0.058
H71	HC	0.178	C13	C3	−0.026	C18	C3	−0.597
H72	HC	0.178	H131	HC	0.117	H181	HC	0.235
H73	HC	0.178	H132	HC	0.117	H182	HC	0.235
C6	CA	0.022	O13	OH	−0.587	H183	HC	0.235

sparsomycin ligand used in the present simulations are given in Tables 1 and 2.

The reduced sparsomycin–ribosome system was hydrated with 10 cycles of MC and MD (10 000 MC moves followed by 10 000 MD steps with 2 fs time step). To ensure the convergence of hydration, 20 cycles of such MC/MD simulations were performed before generating the equilibrated structure for the FEP calculations. To hydrate the inner region, the three types of MC moves (GCMC insertion/deletion, rigid body translation, and rotations) were chosen randomly for the water molecules with equal probability. The solvated sparsomycin–ribosome complex was then equilibrated for 3 ns. All the simulations were carried out using Langevin dynamics at 300 K with 2 fs time step. A friction constant corresponding to a relaxation time of 5 ps was applied to all the non-hydrogen atoms. The bonds involving hydrogens, and the TIP3 water geometry, were kept rigid using SHAKE (Ryckaert *et al*., 1977). All electrostatic interactions beyond 12 Å in the inner region were treated on the basis of dipolar and quadrupolar expansions using the Extended Electrostatic method (EXTE ELEC, Stote *et al*., 1991). This treatment of the nonbonded interactions reduces the computational time by about a factor of two relative to a no-cutoff scheme for the entire inner region while avoiding the artifacts caused by a truncation of electrostatic interactions. All MD simulations are summarized in Table 3.

Table 2. Lennard-Jones parameters for sparsomycin[1]

GAFF type	E_{min} (kcal/mol)	$R_{min}/2$ (Å)
C	−0.0860	1.9080
O	−0.2100	1.6612
N	−0.1700	1.8240
HN	−0.0157	0.6000
C3	−0.1094	1.9080
HC	−0.0157	1.4870
CA	−0.0860	1.9080
OH	−0.2104	1.7210
HO	−0.0460	0.2245
S4	−0.2500	2.0000
SS	−0.2500	2.0000

[1] GAFF non-bonded parameters assigned by ANTECHAMER1.27 (Wang *et al*., 2004; Wang *et al*., 2006b).

Mg^{2+} counterions

Because highly charged RNA polyions are the main constituent of the ribosome, it is imperative to introduce counterions in order to maintain the structural stability of the macromolecular structure. More specifically, Mg^{2+} ions are needed to stabilize pairs of negatively charged phosphate groups that are in proximity of one another within the GSBP inner region. A model of the binding site strictly based on the crystal structure 1M90 without any additional counterions is observed to distort rapidly during MD. As shown in Figure 3A, the RMS deviation of the ligand goes up to 3.0 Å during a 2 ns simulation (MD1 in Table 3). Such deviations are indicative of considerable structural distortions in the ribosome binding site. This is confirmed upon a closer look at the structure at the end of the 2 ns simulation. In principle, the plausible locations of explicit counterions needed to stabilize the macromolecular structure could be estimated from the electrostatic potential calculated by PB (Cheatham and Kollman, 1996; Young *et al*., 1997; Filizola *et al*., 2006). In practice, the most critical positions where Mg^{2+} counterions are needed can be determined directly by monitoring the stability of the neighboring phosphates (P–P pairs). As shown in Figure 5A, several P–P pairs

Table 3. Simulations

Simulations	GSBP systems[1]	MD
MD1	1M90	2 ns
MD2	1M90 + 9 Mg^{2+}	6 ns
MD3[2]	1M90 + 9 Mg^{2+}/harm	3 ns

[1] All systems were hydrated and equilibrated with GCMC/MD and simulated with GSBP for a spherical region of 15 Å radius.
[2] Translational/orientational restraint and average structure of bound ligand were calculated from the last 1 ns of MD3.

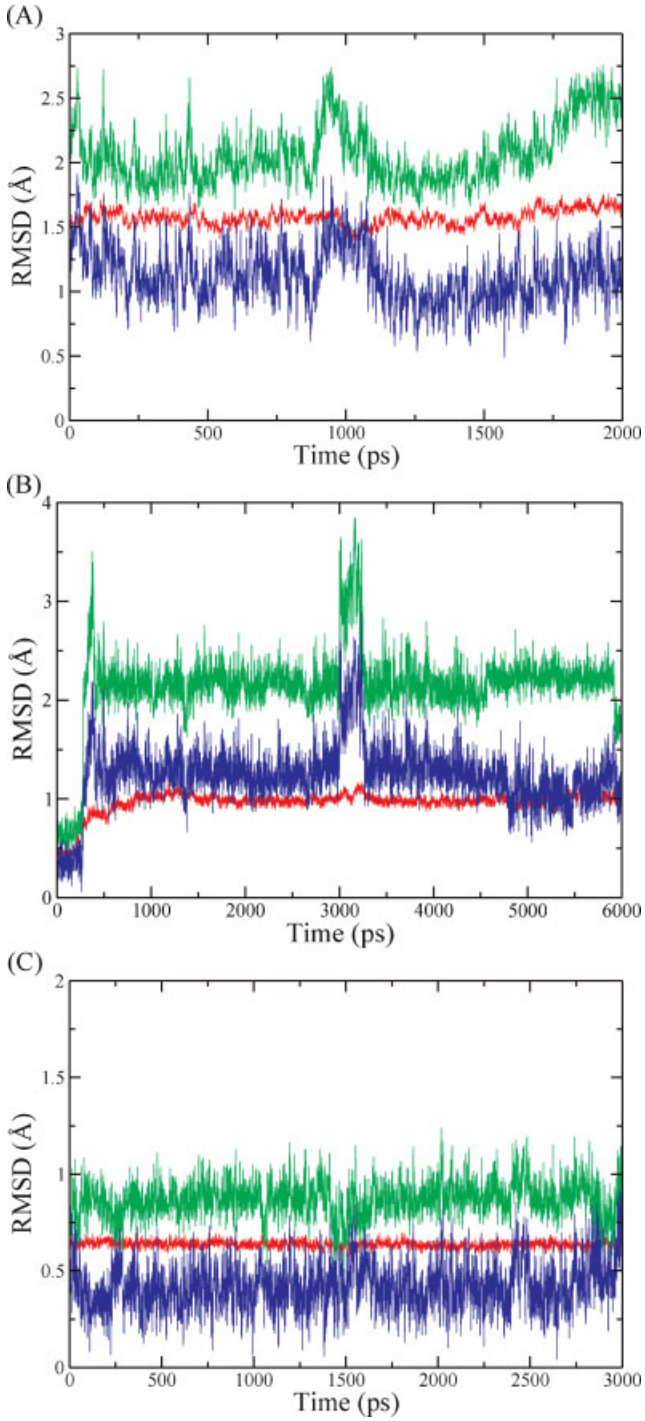

Figure 3. Structural deviations for different MD simulations of sparsomycin bound to the ribosomal P-site. RMSD of non-hydrogen atoms of the binding site (red) and sparsomycin (green), and the center-of-mass of the ligand (blue) are shown. (A) Simulation of the binding site based on the crystal structure 1M90 without additional counterions; (B) simulation of the binding site based on the crystal structure 1M90 with 9 additional Mg^{2+} counterions to stabilize P-P pairs; (C) simulation of a system identical to (B) with additional harmonic restraints applied to the non-hydrogen atoms of the ribosome.

that are proximal in the crystal structure rapidly come apart during MD simulations. On the basis of this observation, the unstable P–P pairs (those with more than 1 Å increase in distances) were identified as potential insertion sites for Mg^{2+}. A water molecule near the unstable P–P pairs was replaced by Mg^{2+}, and then hydration and equilibration with GCMC/MD were repeated. A total of nine Mg^{2+} counterions, illustrated in Figure 4, were inserted in the inner GSBP region according to this procedure. As shown in Figures 3B and 5B, the structure of the binding pocket is stabilized after the insertion of the nine Mg^{2+} counterions (MD2 in Table 3). The P–P pairs now remain closer to the initial distance seen in the crystallographic structure (Figure 5B), and the overall deviation of the phosphates involved in the P–P pairs is less than 1.5 Å RMSD. Nevertheless, the deviations of the RNA components lining the binding pocket remain fairly large and the RMS deviation of the ligand is about 2.0 Å (green line in Figure 3B). To prevent unrealistic conformational changes of the binding pocket away from the X-ray crystallographic structure, a harmonic restraining potential with a force constant of 1.0 kcal/mol/Å^2 was applied to the ribosome non-hydrogen atoms throughout the FEP/MD simulations (MD3 in Table 3). As shown in Figure 3C, the RMS deviation of sparsomycin ligand is now about 0.8 Å RMS with this restraint (green line).

FEP and PMF simulations

The FEP/MD simulation for the sparsomycin in the binding site comprises 18 windows for the ligand repulsion, 5 windows of ligand dispersion, 10 windows of ligand electrostatics, and 15 windows for removing the potential restraints. Two sets of trajectories were generated for each window (forward and backward). For the repulsion, dispersion, and electrostatics stages, each trajectory consisted of six cycles of GCMC/MD run (10 000 MC steps followed by MD with 10 ps equilibration and 10 ps data collection). An additional 20 ps FEP/MD and 10 ps data collection was done after each GCMC/MD cycle (for a total sampling time of 140 ps for each window). For the translational/rotational restraint potential, two sets of trajectories of 140 ps were generated starting from different initial velocities with a series of coupling parameters. The set of all coupling parameters is given in Table 4. The last 100 ps trajectory was collected and processed with the weighted histogram analysis method (WHAM, Kumar *et al*., 1992; Souaille and Roux, 2001). For the FEP/MD simulation of sparsomycin in bulk solvent, the coupling parameters for repulsion, dispersion, and electrostatics stages were the same as those for the bound ligand simulations. For interaction free energy, the length of each simulation window was 70 ps equilibration followed by 70 ps of sampling (140 ps in total for each window). All the FEP calculations were carried out with the PERT module of the program CHARMM (Brooks *et al*., 1983) version c35a1. To account for the missing atoms in the outer region during the free energy calculation with GSBP, a long-range correction to the van der Waals dispersion of the ligand with the rest of the system (ribosome or bulk water) was calculated using a large cutoff of 30 Å with a complete atomic system (Shirts *et al*., 2003; Deng and Roux, 2004). The free energy associated with this long-range correction is −2.1 kcal/mol. The free energy decomposition for the repulsive, dispersive, and electrostatic contributions is shown in Figure 6 and summarized in Table 5.

A set of atoms were selected randomly on the ribosome and sparsomycin ligand to define the anchoring points needed for

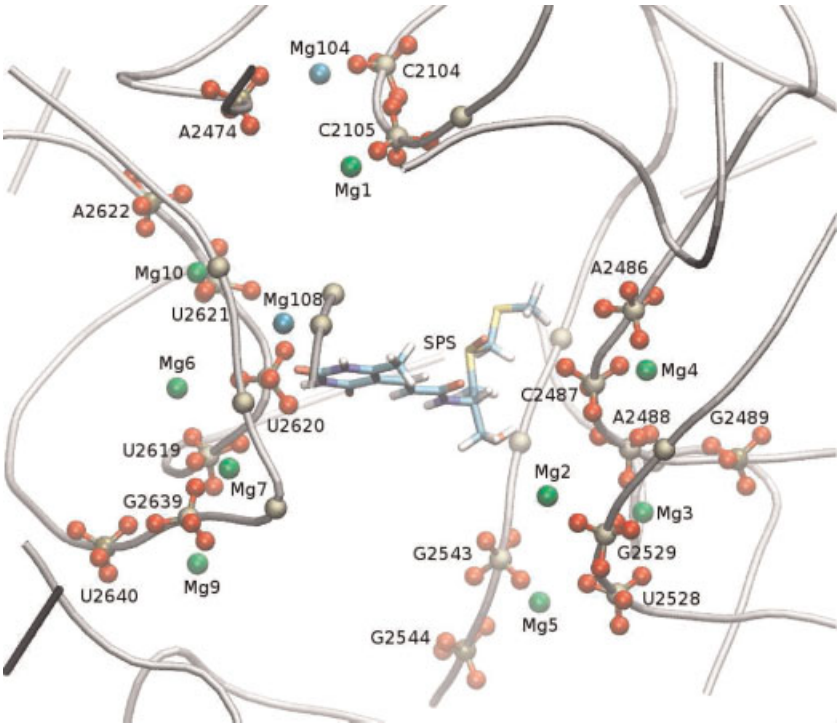

Figure 4. Final locations of the nine Mg^{2+} counterions added to stabilize the P–P pairs in the inner GSBP region (from Figure 3B). The network of phosphate groups (ball and stick) bridged by Mg^{2+} ions are shown in the context of ribosome backbone (silver tubes). The positions of the remaining phosphates in the simulation region are marked by tan balls representing the phosphorus atom. The nine Mg^{2+} counterions (green balls) and the two crystal Mg^{2+} ions (lightblue balls) in the inner simulation region are displayed. Sparsomycin (CPK sticks) is shown in the center.

the translational and rotational restraints used in the staged FEP calculations. They are illustrated in Figure 7A. As shown in Figure 7B, the position and orientation of the ligand are well defined by those internal coordinates. The average of the last 1 ns MD trajectory was used as the reference values for the distance, angle, and dihedral restraints. For the FEP calculations, the force constant was set to 10 kcal/mol/Å^2 for the translational restraints, and to 200 kcal/mol/rad^2 for the orientational restraints. The translational and rotational restraints were combined in one stage with a single coupling parameter $\lambda_{t,r}$, yielding, $\Delta\Delta G_{t,r} = \Delta\Delta G_t^\circ + \Delta\Delta G_r$. In Figure 6D, the progression of the translational and rotational free energy cost is plotted as a function of $\lambda_{t,r}$.

The PMFs of the ligand in the binding site and in the bulk solution were calculated with umbrella sampling as a function the RMSD relative to a reference conformation using 21 biasing windows centered on RMSD offsets increasing from 0.0 to 5.0 Å in steps of 0.25 Å (RMSD-PMF). During the equilibration of sparsomycin in the binding site, the dihedrals N11-C12-C13-O13 involving the hydroxyl group samples two states, the dominant one being similar to the conformation in 1VQ9 and the other close to 1M90. The two crystal structures 1M90 and 1VQ9 are very similar, except for slight differences in the conformation of sparsomycin and Phe-CAP. The reference structure of sparsomycin used for $u_c(\xi)$ was calculated as the average from the last 1 ns of an MD trajectory of sparsomycin bound to the ribosome followed by energy minimization. This conformation is similar to that of the crystal structure 1M90, with some difference regarding the isomerization of the hydroxyl group, which is observed in the other crystal structure 1VQ9 (Schmeing *et al.*, 2005).

The initial configurations for the 21 windows were rapidly generated using a short initial simulation with a strong force constant of 500 kcal/mol/Å^2 to enable a distributed calculation of all the simulations on separate compute nodes. For each window, the system was equilibrated with a force constant of 10 kcal/mol/Å^2 for 100 ps, followed by 400 ps of sampling. No translational or orientational restraining potentials were applied during the RMSD-PMF umbrella sampling simulations. WHAM was used to unbias the results and compute the PMFs (Kumar *et al.*, 1992; Souaille and Roux, 2001). To monitor the convergence of the PMF calculations, 10 successive runs of umbrella sampling simulations (400 ps per window for each) were conducted. During the umbrella sampling simulations, the pseudouracil ring on sparsomycin remains very stable, in contacting with the CCA, base A2637, U2619, U2620, and the ion Mg118. In the early stages of the calculation, the PMFs were changing progressively as a result of the increased sampling. The final PMFs are based on the last 3.6 ns of umbrella sampling simulations.

RESULTS AND DISCUSSION

The calculated standard binding free energy and its decomposition into various contributions are given in Table 5 for the sparsomycin–ribosome complex. Experimentally, the dissociation constant K_d ($=1/K_b$) is estimated from the measured ED_{50}, the concentration of antibiotic causing 50% inhibition of radiolabeled phenol–alanine sparsomycin binding to the ribosome. Values in the range from 1 μM (Lazaro *et al.*, 1991a,b) to 10 μM (Pestka, 1974) can be estimated for *E. coli* and *S. cerevisae* ribosomes. This corresponds to a ΔG_b° of about -8.2 to -6.9 kcal/mol. The present calculations, based on the structure of the archaea strain *H. marismortui*, yield a binding free energy of about

Journal of Molecular Recognition

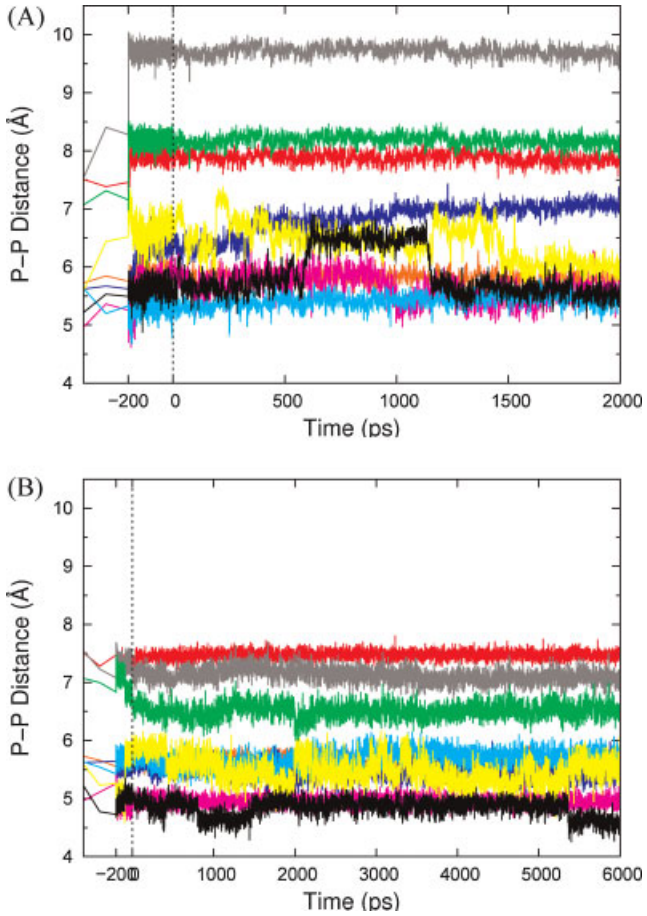

Figure 5. Time evolution of the distances of P–P pairs during initial equilibration. The distances of the P–P pair stablized by the nine Mg^{2+} counterions are plotted by single-colored lines corresponding to nine residue pairs C2105-A2474 (red), U2528-G2529 (orange), A2488-G2489 (blue), A2486-C2487 (magenta), G2543-G2544 (lightblue), U2619-U2620 (yellow), U2619-G2639 (gray), U2619-U2640 (green), and U2621-A2622 (black). The distances measured in the crystal and minimized structures are plotted at the beginning of each time series. (A) Starting with the minimized crystal structure, the system was equilibrated for 2 ns following 200 ps dynamics during the GCMC hydration process. (B) Starting with the minimized crystal structure neutralized by nine Mg^{2+} counterions, the system was equilibrated for 6 ns following 200 ps dynamics during the hydration process.

−5.8 kcal/mol, which is slightly less favorable than the experimental values. To ascertain the convergence of the FEP calculations, we conducted a series of sequential runs following the protocol described in "FEP and PMF simulations" section (each run was restarted from the final configurations of the previous run). This is shown in Figure 8. For both the solvation and binding free energy calculations, the results approach stable values after a few runs. Although the number of repeated runs is limited, the convergence enhances our confidence that the FEP protocol is sound and the results are meaningful. The last five runs were used to calculate the values in Table 5.

Importance of charge neutralization by counterions

The structural integrity of the binding site is documented throughout Figures 3 and 5. As shown in Figure 5A, some phosphate pairs that are initially in proximity of one another come apart with an increase in P-P distance during MD strictly based on the X-ray structure 1M90 without any additional counterions. Some RNA bases flip at a certain point, opening up space for the ligand to move in the binding pocket. Without additional counterions, the backbone of the ribosomal RNA becomes highly unstable and displays significant deviations from the initial X-ray structure. To further illustrate the importance of counterions, two crystallographic Mg^{2+} ions, Mg104 and Mg118, were removed in a separate simulation. As a result, the ligand became very unstable with a considerable increase of RMSD to 3.5 Å upon removal of the two Mg^{2+}. This shows that the latter are critical for sparsomycin binding, consistent with the experimental results from structure-function studies (Nissen *et al*., 2000).

The instability observed in these initial simulations is hardly surprising. There are 22 negatively charged phosphate groups in the GSBP inner region, giving rise to a total charge of −20.765 in the 15 Å spherical simulation region (*n.b.*, the total charge in the inner GSBP region is not necessarily an integer). *In vitro* and *in vivo* studies suggest that Mg^{2+} is essential for the assembly and structural integrity of the ribosome (McCarthy, 1962). Among all the physiologically available ions, Mg^{2+} is especially suitable for neutralizing the RNA phosphate backbone due to its high charge density (Pyle, 2002). About 116 Mg^{2+} ions are included in the available crystal structures of the 50S subunit such as 1M90 and 1S72. Assuming a neutralizing ratio of one Mg^{2+} per two nucleotides, the 116 Mg^{2+} can only serve to neutralize about 8% of the nucleotides, far less than the reported level of 20% (Weiss *et al*., 1973). These considerations suggest that the considerable distortion of the structure is directly caused by the missing Mg^{2+} counterions. To generate a realistically stable simulation system, nine Mg^{2+} were added into the inner region. As shown in Figure 3B, the addition of nine Mg^{2+} counterions results in a substantial stabilization of the binding pocket. The non-hydrogen atoms of the ribosome are stabilized within 1.2 Å RMS during a 6 ns simulation. All the nine counterions added to the system remain coordinated to their P–P pairs, fluctuating around their average position by less than 1 Å. Despite the increase in stability due to the counterions, the RNA bases and Phe-CAP retained a considerable flexibility. To ensure that the conformation of the binding pocket does not deviate significantly from the X-ray crystallographic structure, an artificial harmonic restraint was applied to the non-hydrogen ribosome atoms.

Table 4. Values of the coupling parameters in the FEP calculations

λ_{rep}	0.0	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0					
λ_{dis}	0.0	0.25	0.5	0.75	1.0										
λ_{elec}	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0				
λt,r	0.0	0.003	0.005	0.007	0.01	0.02	0.04	0.06	0.08	0.1	0.2	0.4	0.6	0.8	1.0

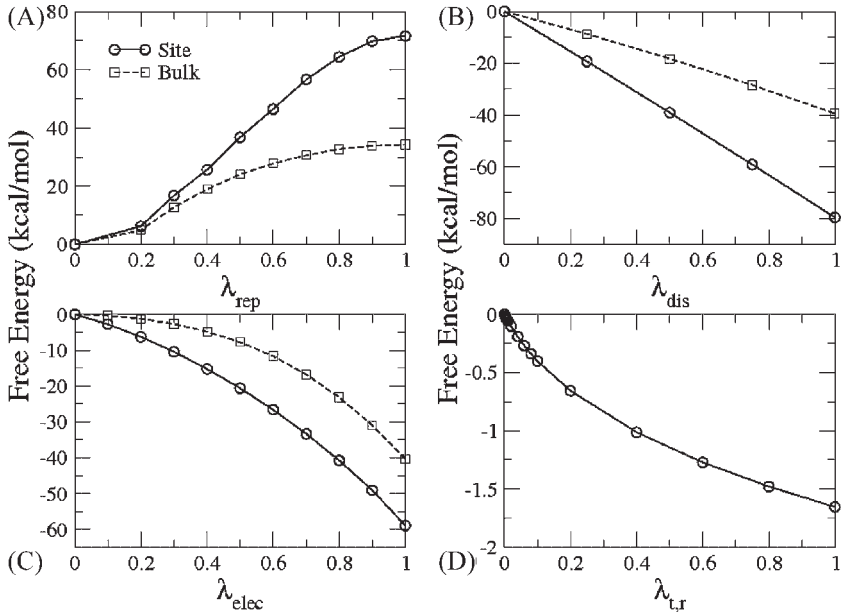

Figure 6. Progression of the free energy components with respect to the coupling parameters for sparsomycin in the binding site (solid lines) or in the bulk solution (dashed lines).

Hydration of the binding pocket during the free energy simulations

Due to its bulky size, sparsomycin is expected to push several water molecules out of the ribosome P-site upon binding. To generate solvent configurations that are consistent with a thermodynamic system in open equilibrium with an infinite bulk reservoir of water molecules, the FEP/MD were combined with the GCMC algorithm (Deng and Roux, 2008). This strategy allows the number of water molecules in the binding pocket to fluctuate and respond dynamically during the various stages of the FEP calculations. The variation in the average number of water molecules within the binding site during the FEP/MD simulation is shown in Figure 9. The number of water molecules steadily decreases when sparsomycin is gradually materialized in the binding pocket. The change is clearly more prominent during the initial stages of the solute insertion, when the purely repulsive WCA-LJ potential is switched on (stages 1–10). About 17 water molecules are expelled from the binding site when the repulsive potential of sparsomycin is introduced. When the dispersive/attractive potential is switched on (stages 10–14), the number of water molecules increases slightly consistent with a weak stabilization resulting from the van der Waals attractive interaction (Vaitheeswaran *et al.*, 2004; Deng and Roux, 2008). Due to the polarity of the ligand, switching on the electrostatic interactions (stages 14–24) also affects the number of water molecules. The considerable fluctuation in the number of water molecules during the free energy simulation justifies the approximation that the number of water molecules is not re-equilibrated at each perturbed λ during the staged FEP. A similar observation was made in the case of campher binding to the P450 target (Deng and Roux, 2008).

Decomposition of interaction free energy

The free energy components $\Delta\Delta G_{rep}$, $\Delta\Delta G_{dis}$, and $\Delta\Delta G_{elec}$ provide an interesting view of the importance of the different microscopic interactions affecting the binding of sparsomycin. Nevertheless, it should be noted that such a decomposition is path-dependent. Each of the components $\Delta\Delta G_{rep}$, $\Delta\Delta G_{dis}$, and $\Delta\Delta G_{elec}$, is a "conditional" free energy, which depends on the applied restraining potentials and on the context of the step-by-step sequence used to compute the reversible work (e.g. electrostatics is switched on in the presence of repulsion and dispersion, etc.). In spite of the path dependence, such a detailed analysis helps clarify the origin of the free energy changes responsible for ligand binding.

Table 5. Binding Free Energy of Sparsomycin to the Ribosome (kcal/mol)[1]

System	ΔG_{rep}	ΔG_{dis}	ΔG_{elec}	$\Delta G_t^\circ + \Delta G_r$	ΔG_c	ΔG_b°
site	71.2 ± 1.3	-81.6 ± 0.2	-58.6 ± 0.3	1.6 ± 0.02	3.1	
bulk	34.3 ± 0.1	-40.7 ± 0.2	-40.4 ± 0.2	$6.6 + 6.4$	8.2	
$\Delta\Delta G$	36.9 ± 1.4	-40.9 ± 0.4	-18.2 ± 0.5	11.4 ± 0.02	5.1	
		-22.2 ± 2.3		16.4 ± 0.2		$\mathbf{-5.8 \pm 1.4}$

[1] The free energies for the site and the bulk are computed with GSBP and SSBP treatment, respectively. The errors are the standard deviation of the last five successive runs, each starting at random velocities from the last configuration of the previous run.

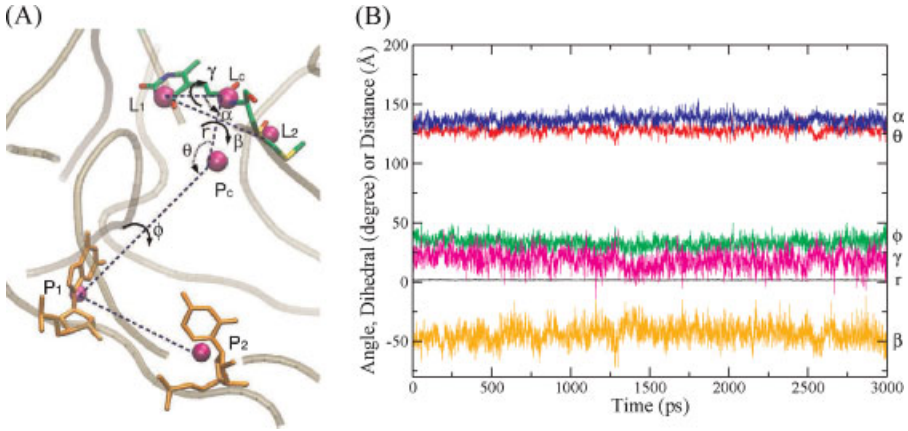

Figure 7. Translational and orientational internal coordinates used for the energy restraints of the ligand in the bound complex. (A) The six restraining internal coordinates are r, and angles θ and ϕ (ligand position) α, β, γ (ligand orientation). P_c and L_c are the center-of-mass of the binding site and the ligand, respectively. The four random point positions L_1 and L_2 are atom-groups in the ligand, P_1 and P_2 are residues in the binding site. Specifically, L_1 consists of C1, O2, and N2 three atoms in the pseudouracil ring of sparsomycin, L_2 is composed of S15, O15, and C16 three atoms in sulfur-containing tail of sparsomycin. P_1 and P_2 are represented by the non-hydrogen atoms of residue G2643 and U2645, respectively, in the fixed boundary of the GSBP sphere. (B) Time series showing the fluctuations of the six internal coordinates. The reference values used for the restraints are the average from the last 1 ns of simulation.

In Figure 6A, the progression of the repulsive free energy is shown as a function of the coupling parameter λ_{rep}. It is clear that the variation of the free energy with respect to λ_{rep} is quite different in the binding site and in bulk water. A point of inflexion at $\lambda_{rep} = 0.2$ in the solvent environment is not observed in the binding site environment. Although the PTC cavity of the ribosome has an appropriate size for sparsomycin, the hydration structure within the ribosome binding site is more dense than bulk water, and the WCA-LJ repulsion free energy pushes to expel the sparsomycin out of the pocket. As a result, the free energy contribution from the repulsive part of the potential disfavors ligand binding. The overall free energy cost is about 36.9 kcal/mol for the repulsive interaction, in disfavor of ligand binding.

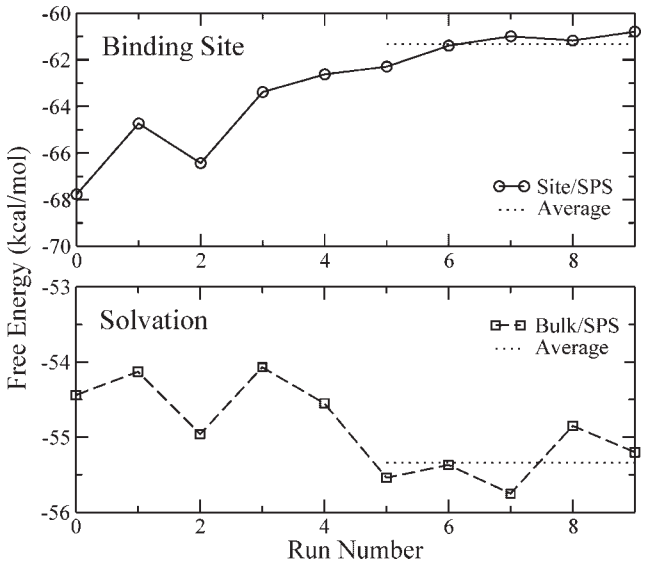

Figure 8. Convergence of the free energy calculations. The top panel shows the binding of sparsomycin to the ribosome. The bottom one shows the solvation of sparsomycin in the bulk. Each window of a free energy calculation is launched from the last configuration of the corresponding window from the previous run. The last five simulations were used for calculating the binding free energy. SPS represents sparsomycin.

In Figure 6B, the progression of the dispersive free energy is plotted as a function of the coupling parameter λ_{dis}. In both bulk solvent and the binding site, the dispersive free energy increases linearly with λ_{dis}, although at different slopes. The linear progression of the dispersive free energy with λ_{dis} reflects the fact that there is no significant changes occurring in the average structure of the ribosome and solvent as this interaction is switched on during the simulation. This difference in slopes originates directly from the number density of van der Waals interaction centers per unit volume surrounding the ligand. Invariably, this number is larger in the binding site of a receptor macromolecule compared with bulk water, which is a fairly loose and extended liquid (Deng and Roux, 2006; Wang *et al*., 2006a; Deng and Roux, 2008). Here, the free energy contribution from the dispersive van der Waals interaction favors ligand binding by

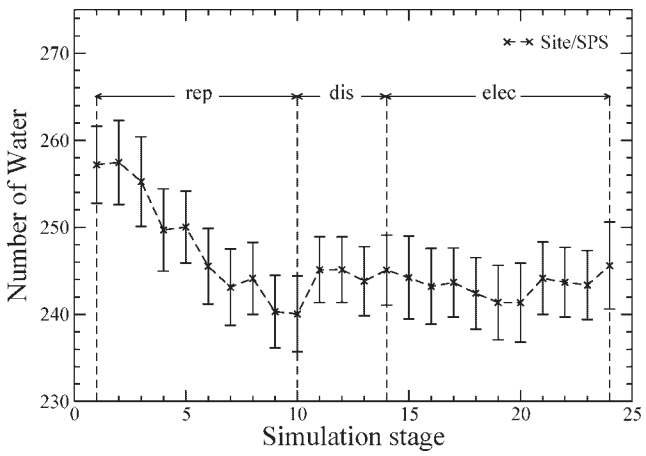

Figure 9. Progression of the number of water molecules in the binding site during the FEP/MD calculation. The simulation is divided into 24 stages, starting from non-interacting stage 1 and progressing to full-interacting sparsomycin in stage 24. In stage 1–10, the repulsion is turned on gradually. In stage 10–14, the attraction is added. In the final stages 14–24, the charges of sparsomycin are switched on. The errors are standard deviation of the five complete runs of FEP calculation.

about −40.9 kcal/mol. This is also consistent with the trends observed with the results from LIE (Marelius *et al*., 1998). Although the repulsion and dispersion have large unfavorable or favorable energy contributions, respectively, their opposite effects nearly cancel out, and the net nonpolar interaction (repulsion and dispersion) is only about −4.1 kcal/mol in favor of binding. When the component from the electrostatic interactions is included, this yields a total interaction free energy contribution of −22.2 kcal/mol, in favor of binding. As will be discussed in the next sections, this strongly favorable free energy is opposed by 16.4 kcal/mol from the unfavorable contribution due to the loss of translational, orientational, and conformational freedom of the ligand.

In Figure 6C, the progression of the electrostatic free energy of sparsomycin is plotted with respect to the coupling parameter, λ_{elec}. The free energy varies quadratically with the coupling parameter for sparsomycin in the binding site as well as in bulk solvent. The trend, characteristic of dielectric media (Lee *et al*., 1988), is indicative that both environments respond electrostatically to the charging of sparsomycin. The ribosome receptor provides a favorable net electrostatic static field for ligand binding, and the polarity of sparsomycin induces an electrostatic reaction field response to the ribosome binding site. In this case, the free energy of charging varies like $Aq + Bq^2$, and both the static field (Aq) and the reaction field (Bq^2) contributions play an important role in sparsomycin binding to the ribosome. In the solvent, there is no average static field and the charging free energy varies only quadratically, without any linear component.

Loss of translational and orientational freedom

Rotational and translational movements of the ligand are significantly restricted once it is confined within the binding pocket. The restriction leads to a nontrivial free energy cost that counteracts the favorable interaction free energy. This free energy cost is often ignored in widely used docking/scoring schemes (Shoichet, 2004). A rigorous treatment of entropy is challenging (Brunger and Karplus, 1988; Jakalian *et al*., 2002), and its partitioning into specific motions of the ligand is even more difficult (Chang *et al*., 2007). The step-by-step FEP/MD simulation protocol adopted here allows an estimate of the free energy cost arising from the loss of translational and rotational freedom of the ligand. The latter is calculated directly from the reversible work $\Delta\Delta G_t^\circ$ and $\Delta\Delta G_r$ associated with imposing and releasing the restraining potentials; this interpretation is particularly clear in the limit of strong restraints (Wang *et al*., 2006a). In the present system, $\Delta\Delta G_t^\circ + \Delta\Delta G_r$ amounts to about 11.4 kcal/mol, in disfavor of binding. The loss of translational and rotational freedom of the ligand cancels about half of the contribution from the favorable sparsomycin-ribosome interactions (−22.2 kcal/mol).

PMF calculations and loss of conformational freedom

Sparsomycin is a highly flexible molecule and the wide range of conformations that are available when it is in bulk solution is inevitably reduced when it is in the binding pocket environment. It is an important challenge to quantify the impact of such a loss of conformational freedom in FEP/MD simulations on a large and complex ligand. To address this issue, we utilize a deliberate sampling strategy, which was first introduced by Woo and Roux in a study of a highly flexible phosphotyrosine peptide binding to a SH2 domain (Woo and Roux, 2005). The restraining potential $u_c(\xi)$ defined in terms of the RMSD of the ligand with respect to a reference state is introduced at one end-point to "confine" the ligand, and then "released" at the other end-point. The latter step is carried out via an explicit numerical integration of the Boltzmann factor of the unbiased PMFs, according to Equations (8) and (9). Dill and co-workers have used a similar "confine-and-release" procedure to treat the effect of sidechain rotamers in T4 Lysozyme (Mobley *et al*., 2007).

The RMSD–PMFs of sparsomycin are plotted in Figure 10A. The PMF of sparsomycin in the binding site displays four free energy minima, labeled i–iv. The global minimum is ii, corresponding to the conformation observed in the crystal structure 1M90, at a RMSD of 1.1 Å relative to the reference conformation. There are also three additional local free energy minima, at 0.5, 1.5, and 1.7 Å, labeled i, iii, and iv, respectively. The small barriers suggest the possibility of interconversion between those conformations. In general, sparsomycin in the bound state is confined to a RMSD less than 2 Å. In contrast, the PMF calculated in bulk solution shown in Figure 10A exhibits a much broader minimum centered around a RMSD of 2.3 Å. The PMFs suggest that conformations differing from the bound state up to 3.5 Å RMSD should be thermally accessible in bulk solution. Representative instantaneous snapshots taken from the simulations are illustrated in Figure 10B. The local free energy minima in the PMF calculations correspond to the conformational differences seen in the 1M90 and 1VQ9 crystal structures. Sparsomycin samples a broad configurational space with various rotameric combinations of the hydroxyl group and tail methyl group. In bulk solvent, the most populated conformations of sparsomycin differ mostly by the rotation of tail groups. As a consequence, a considerable thermodynamic penalty is incurred to restrict the ligand from its ensemble of accessible conformations in the bulk to the much

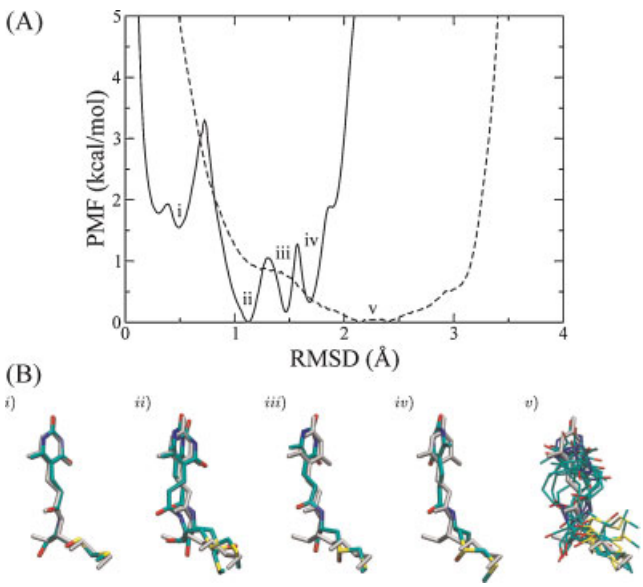

Figure 10. PMF calculations on the conformational restraints for sparsomycin in different environments. (A) PMF profiles of sparsomycin in the ribosome binding site (solid line) and in the bulk solution (dashed lines). The minima of the PMFs for the ligand in the binding site are labeled i, ii, iii, and iv. The minimum of the PMF for the ligand in the bulk solution is labeled v. (B) Instantaneous dynamic configurations (sticks) at the minima are aligned with the minimized average structure (gray lines) of the ligand in the binding site after 3 ns. The RMSD of the non-hydrogen atoms is used for the alignments.

narrower ensemble of possible conformations allowed in the binding site. Following Equations (8) and (9), this translates into a free energy of about 5 kcal/mol (Table 5).

CONCLUSIONS

An MD/FEP simulation methodology has been successfully applied to investigate the molecular recognition phenomenon in one of the largest known system, the binding of antibiotic sparsomycin to the ribosome. The computed binding free energy value, -5.8 ± 1.4 kcal/mol, agrees reasonably well with experimental data. Convergence of the result was tested by repeated FEP calculations. Of particular interest, it is noted that the total interaction free energy, -22.2 kcal/mol, favorable to ligand binding, is strongly opposed by a total of 16.4 kcal/mol arising from the loss of translational, orientational, and conformational freedom of the ligand. Such unfavorable contributions are often ignored in conventional docking/scoring schemes used in drug design (Shoichet, 2004), as well as approximations such as LIE (Åqvist *et al*., 2002) and MMPB-SA (Kollman *et al*., 2000; Simonson *et al*., 2002; Swanson *et al*., 2004). The present results underline the importance of accounting for those contributions properly.

The study illustrates that calculations of the absolute binding free energy provide insight into the molecular interactions of antibiotics in complex with the ribosome, allowing to go beyond the static view of the X-ray crystal structure. Nevertheless, it is clear that computational studies and free energy simulations involving the ribosome are confronted with outstanding challenges. Those regard (1) the size of the simulation system, (2) the magnitude of electrostatic interactions, (3) the hydration of the buried binding pocket, and (4) the considerable flexibility of the sparsomycin ligand. In this study, we have strived to specifically tackle each of those four challenges.

(1) The large size of the sparsomycin–ribosome complex renders extensive FEP simulations computationally prohibitive. Here, the system was reduced by adopting a hybrid MD/continuum approach based on GSBP (Im *et al*., 2001). Only the receptor atoms in proximity to the ligand binding region were simulated explicitly, while the influence from the remote atoms was included implicitly. The complete structure of the 50S subunit, comprising about 150 000 atoms (excluding solvent), was reduced to a simulation system of about 4100 atoms. The hybrid MD/continuum GSBP strategy enables the sufficient and time-efficient sampling necessary for FEP/MD simulations without compromising the accuracy of the calculation considerably.

(2) With its high density of polyion RNA, the strong electrostatic interactions represents a critical aspect of the ribosome structure. Here, ionic screening corresponding to the physiological salt concentration was used in the Poisson–Boltzmann calculations of the static field arising from the distant atoms in the outer region. In addition, the presence of several Mg^{2+} counterions within the inner region was found to be essential to maintain the stability of the phosphate pairs of the RNA molecules, as well as to model accurately the interaction with the ligand.

(3) A careful treatment of hydration and water occupancy within the binding pocket is particularly important for MD/FEP. The large ligand sparsomycin expels about 12 water molecules from the ribosome pocket upon binding. Diffusion of water molecules in and out of the pocket is far too slow to be handled accurately via brute force simulations. Those problems are further compounded by the reduction of the simulation system with GSBP. To compute free energies consistent with solvent molecules in open exchange with a thermodynamic reservoir, the GCMC algorithm was coupled to FEP/MD. The results imply that this is an efficient strategy for treating hydration during free energy computations.

(4) Sparsomycin displays considerable flexibility, allowing multiple stable conformations in solution. This flexibility is severely restricted when sparsomycin is in the binding pocket. In principle, brute force simulations could be able to sample those conformations. In practice, however, this may require prohibitively long trajectories. A strategy based on a deliberate sampling protocol is potentially more efficient. First the PMF of the sparsomycin was calculated using umbrella sampling as a function of the RMSD relative to a reference structure, taken as the average conformation in the bound state. The PMF of sparsomycin in the binding site displays multiple local free energy minima closely related to the bound conformation, whereas the PMF of sparsomycin in bulk water reveals large conformational fluctuations. Then, the alchemical FEP/MD simulations were performed in the presence of a conformational energy restraint, u_c, applied to confine the ligand around its bound conformation. The effect of the restraint was rigorously unbiased in post-analysis, revealing that the free energy cost due to the loss of conformational freedom of sparsomycin upon binding is about 5.0 kcal/mol (Table 5).

The present study illustrates how alchemical FEP/MD methodologies can be applied to compute the absolute binding free energy of small ligands to ultra-large macromolecular complexes such as the ribosome. On the downside, it should be noted that the model adopted, including force field, placement of ions, and MD/continuum approach, was not able to maintain the ribosome close to its crystal structure. Although the situation was improved by the presence of Mg^{2+} counterions, some additional harmonic restraints were needed to stabilize the structure of the binding site. It is likely that the presence of such structural restraints helped, in part, the convergence of the free energy calculations. Despite this shortcoming, the calculated binding free energy is within 1.1 to 2.4 kcal/mol of the experimental values. Clearly, the present study serves largely to identify and illustrate the numerous difficulties that are encountered when trying to carry out rigorous FEP calculations on this system and it will be necessary to consider additional ligands for a robust validation of this methodology. Although binding affinity is not the sole criterion for therapeutic efficiency of an antibiotic drug, an accurate calculation of the absolute binding free energy is a pre-requisite to achieve a clear understanding of the physical mechanism governing the binding process. It is our hope that the insight gained here will facilitate the discovery and design of novel lead compounds, and that free energy methods will play an increasing role in structure-based drug design.

Acknowledgements

The authors are grateful to Yuqing Deng, Haibo Yu, and Albert Pan for helpful discussions and suggestions. This work was supported by grant MCB-0630140 from the National Science Foundation.

REFERENCES

Anisimov VM, Lamoureux G, Vorobyov I, Huang N, Roux B, MacKerell A. 2005. Determination of electrostatic parameters for a polarizable force field based on the classical drude oscillator. *J. Chem. Theor. Comput.*, **1**: 153–168.

Åqvist J, Luzhkov VB, Brandsdal BO. 2002. Ligand binding affinities from MD simulations. *Account. Chem. Res.* **35**: 358–365.

Banavali N.K., Roux B. 2002. Atomic radii for continuum electrostatics calculations on nucleic acids. *J. Phys. Chem. B* **106**: 11026–11035.

Bashan A, Agmon I, Zarivach R, Schluenzen F, Harms J, Berisio R, Bartels H, Franceschi F, Auerbach T, Hansen HAS, Kossoy E, Kessler M, Yonath A. 2003. Structural basis of the ribosomal machinery for peptide bond formation, translocation, and nascent chain progression. *Mol. Cell* **11** (1): 91–102.

Becke AD. 1993. Density-functional thermochemistry .3. the role of exact exchange. *J. Chem. Phys.* **98** (7): 5648–5652 .

Beglov D, Roux B. 1994. Finite representation of an infinite bulk system: solvent boundary potential for computer simulations. *J. Chem. Phys.* **100**: 9050–9063.

Boresch S, Tettinger F, Leitgeb M, Karplus M. 2003. Absolute binding free energies: a quantitative approach for their calculation. *J. Phys. Chem. B* **107**: 9535–9551.

Brooks B, Bruccoleri R, Olafson B, States D, Swaminathan S, Karplus M. 1983. Charmm: a program for macromolecular energy minimization and dynamics calculations. *J. Comput. Chem.* **4**: 187–217.

Brunger AT, Karplus M. 1988. Polar hydrogen positions in proteins - empirical energy placement and neutron-diffraction comparison. *Prot. Struct. Funct. Genet.* **4** (2): 148–156.

Chang EA, Chen W, Gilson MK. 2007. Ligand configurational entropy and protein binding. *Proc. Natl Acad. Sci. USA* **104** (5): 1534–1539.

Cheatham TE, Kollman PA. 1996. Observation of the a-dna to b-dna transition during unrestrained molecular dynamics in aqueous solution. *J. Mol. Biol.* **259** (3): 434–444.

Deng Y, Roux B. 2004. Hydration of amino acid side chains: non-polar and electrostatic contributions calculated from staged molecular dynamics free energy simulations with explicit water molecules. *J. Phys. Chem. B* **108**: 16567–16576.

Deng Y, Roux B. 2006. Calculation of standard binding free energies: aromatic molecules in the t4 lysozyme l99a mutant. *J. Chem. Theor. Comput.* **2** (5): 1255–1273.

Deng Y, Roux B. 2008. Computation of binding free energy with molecular dynamics and grand canonical monte carlo simulations. *J. Chem. Phys.* **128** (11): 115103.

Deng Y, Roux B. 2009. Computations of standard binding free energies with molecular dynamics simulations. *J. Phys. Chem. B* **113**: 2234–2246.

Filizola M, Wang SX, Weinstein H. 2006. Dynamic models of g-protein coupled receptor dimers: indications of asymmetry in the rhodopsin dimer from molecular dynamics simulations in a popc bilayer. *J. Comput. Aided Mol. Des.* **20** (7–8): 405–416.

Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA. 2004. Gaussian 03, revision c.02. Gaussian, Inc., Wallingford CT.

Foloppe N, MacKerell AD Jr. 2000. All-atom empirical force field for nucleic acids: 1) parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comp. Chem.* **21**: 86–104.

Gilson MK, Given JA, Bush BL, Andrew McCammon J. 1997. The statistical-thermaldynamic basis for computation of binding affinities: a critical review. *Biophys. J.* **72**: 1047–1069.

Hansen JL, Moore PB, Steitz TA. 2003. Structures of five antibiotics bound at the peptidyl transferase center of the large ribosomal subunit. *J. Mol. Biol.* **330** (5): 1061–1075 .

Hansen JL, Schmeing TM, Moore PB, Steitz TA. 2002. Structural insights into peptide bond formation. *Proc. Natl Acad. Sci. USA* **99** (18): 11670–11675.

Hermans J, Subramaniam S. 1986. The free energy of xenon binding to myoglobin from molecular dynamics simulation. *Isr. J. Chem.* **27**: 225–227.

Hermans J, Wang L. 1997. Inclusion of loss of translational and rotational freedom in theoretical estimates of free energies of binding. Application to a complex of benzene and mutant T4 lysozyme. *J. Am. Chem. Soc.* **119**: 2707–2714.

Hertwig RH, Koch W. 1995. On the accuracy of density functionals and their basis-set dependence - an extensive study on the main-group homonuclear diatomic-molecules li-2 to br-2. *J. Comput. Chem.* **16** (5): 576–585.

Im W, Berneche S, Roux B. 2001. Generalized solvent boundary potential for computer simulations. *J. Chem. Phys.* **114** (7): 2924–2937.

Jakalian A, Jack DB, Bayly, Fast CI. 2002. efficient generation of high-quality atomic charges. am1-bcc model: Ii. parameterization and validation. *J. Comput. Chem.* **23** (16): 1623–1641 .

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79** (2): 926–935.

Klein DJ, Schmeing TM, Moore PB, Steitz TA. 2001. The kink-turn: a new rna secondary structure motif. *EMBO J.* **20** (15): 4214–4221.

Kollman PA, Massova I, Reyes C, Kuhn B, Huo SH, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE. 2000. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Account. Chem. Res.* **33** (12): 889–897.

Korostelev A, Trakhanov S, Laurberg M, Noller HF. 2006. Crystal structure of a 70S ribosome-trna complex reveals functional interactions and rearrangements. *Cell* **126** (6): 1065–1077.

Kumar S, Bouzida D, Swendsen R, Kollman P, Rosenberg J. 1992. The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method. *J. Comp. Chem.* **13**: 1011–1021.

Lazaro E, Van Den Broek LAGM, Felix AS, Ottenheijm HCJ, Ballesta JPG. 1991a. Biochemical and kinetic characteristics of the interaction of the antitumor antibiotic sparsomycin with prokaryotic and eukaryotic ribosomes. *Biochemistry* **30**: 9642–9648.

Lazaro E, Vandenbroek LAGM, Sanfelix A, Ottenheijm HCJ, Ballesta JPG. 1991b. Chemical, biochemical and genetic endeavors characterizing the interaction of sparsomycin with the ribosome. *Biochimie* **73**: 1137–1143.

Lee CT, Yang WT, Parr RG. 1988. Development of the colle-salvetti correlation-energy formula into a functional of the electron-density. *Phys. Rev. B* **37** (2): 785–789.

MacKerell AD Jr, Banavali NK. 2000. All-atom empirical force field for nucleic acids: 2) application to molecular dynamics simulations of dna and rna in solution. *J. Comp. Chem.* **21**: 105–120.

MacKerell AJ, Bashford D, Bellot M, Dunbrack R, Evanseck J, Field M, Fischer S, Gao J, Guo H, Ha DJ-MS, Kuchnir L, Kuczera K, Lau F, Mattos C, Michnick S, Ngo T, Nguyen D, Prodhom B, Reiher W III, Roux B, Schlenkrich M, Smith J, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Karplus M. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**: 3586–3616.

Marelius J, Hansson T, Aqvist J. 1998. Calculation of ligand binding free energies from molecular dynamics simulations. *Int. J. Quant. Chem.* **69** (1): 77–88.

McCarthy BJ. 1962. Effects of magnesium starvation on ribosome content of escherichia coli. *Biochimica Et Biophysica Acta* **55**: 880.

Moazed D, Noller HF. 1989. Interaction of transfer-rna with 23s ribosomal-rna in the ribosomal a-sites, p-sites, and e-sites. *Cell* **57**: 585–597.

Mobley DL, Chodera JD, Dill KA. 2007. Confine-and-release method: obtaining correct binding free energies in the presence of protein conformational change. *J. Comp. Theor. Chem.* **3**: 1231–1235.

Nina M, Beglov D, Roux B. 1997. Atomic radii for continuum electrostatics calculations based on molecular dynamics free energy simulations. *J. Phys. Chem. B* **101**: 5239–5248.

Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. 2000. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289** (5481): 920–930.

Noller HF. 2005. RNA structure: reading the ribosome. *Science* **309** (5740): 1508–1514.
Pestka S. 1974. Antibiotics as probes of ribosome structure: binding of chloramphenicol and erythromycin to polyribosomes; effect of other antibiotics. *Antimicrob. Agent. Chemother.* **5**: 255–267.
Pyle AM. 2002. Metal ions in the structure and function of rna. *J. Biol. Inorg. Chem.* **7** (7–8): 679–690.
Rodnina MV, Beringer M, Wintermeyer W. 2007. How ribosomes make peptide bonds. *Trend. Biochem. Sci.* **32** (1): 20–26.
Roux B, Nina M, Pomès R, Smith JC. 1996. Thermodynamic stability of water molecules in the bacteriorhodopsin proton channel: a molecular dynamics free energy perturbation study. *Biophys. J.* **71**: 670–681.
Ryckaert J, Ciccotti G, Berendsen H. 1977. Numerical Integration of the cartesian equation of motions of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comp. Chem.* **23**: 327–341.
Sanbonmatsu KY, Tung CS. 2007. High performance computing in biology: multimillion atom simulations of nanoscale systems. *J. Struct. Biol.* **157**: 470–480.
Schmeing TM, Huang KS, Kitchen DE, Strobel SA, Steitz TA. 2005. Structural insights into the roles of water and the 2′ hydroxyl of the p site trna in the peptidyl transferase reaction. *Mol. Cell* **20** (3): 437–448.
Schuwirth BS, Borovinskaya MA, Hau CW, Zhang W, Vila-Sanjurjo A, Holton JM, Cate JHD. 2005. Structures of the bacterial ribosome at 3.5 angstrom resolution. *Science* **310** (5749): 827–834.
Selmer M, Dunham CM, Murphy FV, Weixlbaumer A, Petry S, Kelley AC, Weir JR, Ramakrishnan V. 2006. Structure of the 70s ribosome complexed with mrna and trna. *Science* **313** (5795): 1935–1942.
Shirts MR, Pitera JW, Swope WC, Pande VS. 2003. Extremely precise free energy calculations of amino acid side chain analogs: comparison of common molecular mechanics force fields for proteins. *J. Chem. Phys.* **119**: 5740–5761.
Simonovc M, Steitz TA. 2008. Cross-crystal averaging reveals that the structure of the peptidyl-transferase center is the same in the 70s ribosome and the 50s subunit. *Proc. Natl Acad. Sci. USA* **105**: 500–505.
Simonson T, Archontis G, Karplus M. 2002. Free energy simulations come of age: protein-ligand recognition. *Account. Chem. Res.* **35** (6): 430–437.
Shoichet B. Virtual screening of chemical libraries. 2004. *Nature* **432**: 862–865.
Souaille M, Roux B. 2001. Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Comp. Phys. Comm.* **135**: 40–57.
Stern S, Powers T, Changchien LM, Noller HF. 1989. Rna-protein interactions in 30s ribosomal-subunits - folding and function of 16s ribosomal-rna. *Science* **244**: 783–790.
Stote R, States D, Karplus M. 1991. On the treatment of electostatic interactions in biomolecular simulation. *J. Chim. Phys.* **88**: 2419–2433.
Swanson JMJ, Henchman RH, McCammon JA. 2004. Revisiting free energy calculations: a theoretical connection to MM/PBSA and direct calculation of the association free energy. *Biophys. J.* **86**: 67–74.
Vaitheeswaran S, Yin H, Rasaiah JC, Hummer G. 2004. Water clusters in nonpolar cavities. *Proc. Natl Acad. Sci. USA* **101** (49): 17002–17005.
Van Den Broek LAGM, Fennis PJ, Arevalo MA, Lazaro E, Ballesta JPG, Lelieveld P, Ottenheijm HCJ. 1989. The role of the hydroxymethyl function on the biological-activity of the antitumor antibiotic sparsomycin. *Eur. J. Med. Chem.* **24**: 503–510.
Van Den Broek LAGM, Liskamp RMJ, Colstee JH, Lelieveld P, Remacha M, Vazquez D, Ballesta JPG, Ottenheijm HCJ. 1987. Structure-activity-relationships of sparsomycin and its analogs - inhibition of peptide-bond formation in cell-free systems and of l1210 and bacterial-cell growth. *J. Med. Chem.* **30**: 325–333.
Wang JY, Deng YQ, Roux B. 2006a. Absolute binding free energy calculations using molecular dynamics simulations with restraining potentials. *Biophys. J.* **91** (8): 2798–2814.
Wang JM, Wang W, Kollman PA, Case DA. 2006b. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **25** (2): 247–260.
Wang JM, Wolf RM, Caldwell JW, Kollman PA, Case DA. 2004. Development and testing of a general amber force field. *J. Comput. Chem.* **25** (9): 1157–1174.
Weeks JD, Chandler D, Andersen HC. 1971. Role of repulsive forces in determining equilibrium structure of simple liquids. *J. Chem. Phys.* **54**: 5237.
Weiss RL, Kimes BW, Morris DR. 1973. Cations and ribosome structure .3. effects on 30s and 50s subunits of replacing bound mg2+ by inorganic cations. *Biochemistry* **12**: 450–4456.
Woo HJ, Roux B. 2005. Calculation of absolute protein-ligand binding free energy from computer simulations. *Proc. Natl Acad. Sci. USA* **102** (19): 6825–6830.
Woodcock J, Moazed D, Cannon M, Davies J, Noller HF. 1991. Interaction of antibiotics with a-site-specific and p-site-specific bases in 16s ribosomal-rna. *EMBO J.* **10**: 3099–3103.
Young MA, Jayaram B, Beveridge DL. 1997. Intrusion of counterions into the spine of hydration in the minor groove of b-dna: fractional occupancy of electronegative pockets. *J. Am. Chem. Soc.* **119** (1): 59–69.
Yusupova GZ, Yusupov MM, Cate JHD, Noller HF. 2001. The path of messenger rna through the ribosome. *Cell* **106**: 233–241.

