Calculations of Free-Energy Contributions to Protein–RNA Complex Stabilization

Mark A. Olson
Molecular Modeling Laboratory, Department of Cell Biology and Biochemistry, United States Army Medical Research Institute of Infectious Diseases, Frederick, Maryland 21702 USA

ABSTRACT The problem of calculating binding affinities of protein–RNA complexes is addressed by analyzing a computational strategy of modeling electrostatic free energies based on a nonlinear Poisson–Boltzmann (NLPB) model and linear response approximation (LRA). The underlying idea is to treat binding as a two-step process. Solutions to the NLPB equation calculate free energies arising from electronic polarizability and the LRA is constructed from molecular dynamics simulations to model reorganization free energies due to conformational transitions. By implementing a consistency condition of requiring the NLPB model to reproduce the solute–solvent free-energy transitions determined by the LRA, a "macromolecule dielectric constant" (ϵ_m) for treating reorganization is obtained. The applicability of this hybrid approach was evaluated by calculating the absolute free energy of binding and free-energy changes for amino acid substitutions in the complex between the U1A spliceosomal protein and its cognate RNA hairpin. Depending on the residue substitution, ϵ_m varied from 3 to 18, and reflected dipolar reorientation not included in the polarization modeled by $\epsilon_m = 2$. Although the changes in binding affinities from substitutions modeled strictly at the implicit level by the NLPB equation with $\epsilon_m = 4$ reproduced the experimental values with good overall agreement, substitutions problematic to this simple treatment showed significant improvement when solved by the NLPB-LRA approach.

INTRODUCTION

Molecular interactions between proteins and RNAs are ubiquitous events critical to many cellular processes. With the increasing number of reported protein–RNA complexes determined by x-ray crystallography and NMR spectroscopy, computational methods are challenged to provide detailed understanding of the factors that govern molecular recognition of nucleic acids by proteins and the formation of stable complexes. Included in this challenge are quantitative calculations of the absolute binding free energies and the evaluation of the effects of amino acid substitutions on complex stabilization. Application of the latter is the structure-based design of protein ligands where it is necessary to anticipate the effect of residue substitution on binding the RNA molecule.

An attractive computational approach for estimating binding affinities of protein–RNA complexes is continuum models (Sharp and Honig, 1990; Olson, 1999). Rather than enumerating the configurations of the macromolecules and surrounding solvent by all-atom simulations, a mean-field treatment is implemented, in which the complex is modeled by a semimicroscopic description and the solvent by a dielectric continuum. Electrostatic contributions to binding affinities are calculated from numerical solutions to the Poisson–Boltzmann (PB) equation, and nonpolar interactions are modeled by using solute–solvent cavitation energies. Although several recent studies demonstrated the general applicability of PB models for analyzing protein interactions with nucleic acids (Misra et al., 1998; Olson and Cuff, 1999; Reyes and Kollman, 2000a,b), many issues remain to be clarified regarding the consistent treatment of the dipolar response underlying macromolecular associations.

Central to the application of PB models and a source of many of the unresolved issues is the so-called macromolecule dielectric constant (ϵ_m). It is generally accepted that, in a continuous approach, the solute ϵ_m is a scaling factor that represents all of the contributions that are not treated explicitly, rather than a true dielectric constant (King et al., 1991; Warshel and Åqvist, 1991). Although standard continuum methods commonly model relaxation and nonrelaxation free energies by the application of a single ϵ_m, macromolecular dielectric environments are in fact inhomogeneous. Dielectric constants for proteins calculated from simulations show dipole moment fluctuations that depend on the site considered (King et al., 1991; Simonson and Perahia, 1995; Simonson and Brooks, 1996). In the dehydrated core regions, electronic polarizability makes the largest contribution to ϵ_m with a value of ~2. Near ionizable groups at the periphery, ϵ_m reflects the reorientation of charged side chains and contains a much larger value of ~20–40. Because of this manifold of dielectric responses, the binding process reflects a combination of induced dipoles and reorganization from conformational transitions. The optimal choice of the dielectric constant to treat both contributions presents a challenge for continuum methods. This challenge arises principally from a lack of universality in ϵ_m and the weak connection to the physical macroscopic constant (King et al., 1991; Warshel and Papazyan, 1998). The development of accurate and reliable models ultimately requires a departure from a strictly implicit scheme to that of an explicit representation of the relaxation component.

The purpose of this study is to explore the application of combining a nonlinear PB (NLPB) model with the explicit

Received for publication 4 May 2000 and in final form 13 June 2001.

Address reprint requests to Mark A. Olson, USAMRIID, Molecular Modeling Lab, Dept. of Cell Biology & Biochemistry, 1425 Porter St., Frederick, MD 21702. Tel.: 301-619-4236; Fax: 301-619-2348; E-mail: molson@ncifcrf.gov.

0006-3495/01/10/1841/13 $2.00

determination of reorganization free energies. A computational strategy is developed that retains the simplicity of PB models for calculating the induced polarization component, whereas conformational transitions are modeled at the explicit level by a linear response approximation (LRA). The LRA method provides a powerful and convenient calculational framework for determining free energies of solvation by all-atom simulations. Recent applications of LRA include modeling protein–inhibitor interactions (Lee et al., 1992; Åqvist et al., 1994; Muegge et al., 1997; Jones-Hertzog and Jorgensen, 1997), protein–protein binding (Muegge et al., 1998), and charge insertion in proteins (Del Buono et al., 1994; Sham et al., 1997, 1998; Simonson et al., 1999). Here, a theoretical model is provided of the free-energy terms that contribute to complex formation between the U1A protein and a 21-nucleotide RNA hairpin. U1A is a component of the U1 small nuclear ribonucleoprotein and plays a functional role in initiating spliceosomal assembly. The complex with the RNA is well-characterized crystallographically (Oubridge et al., 1994) and experimental differences in binding affinities from amino acid substitutions have been reported (Jensen et al., 1991; Kranz et al., 1996; Kranz and Hall, 1998, 1999).

To establish the utility of a combined NLPB-LRA approach, a series of models were applied to the U1A–RNA complex in estimating the absolute binding free energy and the effects of mutations. By using molecular dynamics (MD) simulations, reorganization in solute–solvent free energies was evaluated by the LRA. These simulation results were combined with the nonrelaxation terms calculated from the NLPB equation to yield the final free energies. Calculations are also reported for the application of standard NLPB methods, followed by a nonstandard protocol of using the NLPB model with the protein dielectric environment treated implicitly as inhomogeneous (Olson, 1999; Olson and Reinke, 2000). The NLPB-LRA strategy differs from the recent works of Reyes and Kollman (2000a,b) and provides an approach by way of LRA that is less sensitive to the large calculational fluctuations in determining interaction energies (Muegge et al., 1998). Moreover, the NLPB-LRA model offers a computational framework to gauge the accuracy of implicit versus explicit schemes for modeling associations.

THEORY AND COMPUTATIONAL STRATEGIES

General Formulation

The molecular association between the U1A protein (denoted as $\mathcal{P}$) and the RNA receptor ($\mathcal{R}$) can be described as a two-step process given by (see, e.g., Misra and Honig, 1995; Misra et al., 1998)

$$\mathcal{P} + \mathcal{R} \overset{\Delta G_{\text{conf}}}{\Leftrightarrow} \mathcal{P}^* + \mathcal{R}^*, \tag{1}$$

$$\mathcal{P}^* + \mathcal{R}^* \overset{\Delta G_{\text{static}}}{\Leftrightarrow} \mathcal{P}^*\mathcal{R}^*, \tag{2}$$

where the first step accounts for the conformational transition from the native $\mathcal{P}$ and $\mathcal{R}$ conformers to the functional (i.e., bound) conformers $\mathcal{P}^*$ and $\mathcal{R}^*$, and the second step describes the binding of the two macromolecules in their functional state. The Gibbs free energy of binding is obtained by combining Eqs. 1 and 2,

$$\begin{aligned}\Delta G_{\text{bind}} &= \Delta G_{\text{conf}} + \Delta G_{\text{static}} \\ &= \Delta G_{\text{conf}} + \Delta W + \Delta G_{\text{int}} - T\Delta S,\end{aligned} \tag{3}$$

where ΔG_{conf} is the free-energy shift upon conformational transitions, and ΔG_{static} is the nonrelaxation free energy. This latter term will be referred to below as the "static" free energy that considers the single-conformer process commonly invoked in continuum models. A more complete description for ΔG_{static} is given by partitioning the term into independent energetic contributions, where ΔW represents the change in the potential of mean force for $\mathcal{P}^*$ and $\mathcal{R}^*$ due to solute–solvent interactions, ΔG_{int} is the interaction free energy between the bound molecules, and $T\Delta S$ is the change in nonelectrostatic entropy determined at temperature T. In Eq. 3, the concentration of the molecular species $\mathcal{P}$, $\mathcal{R}$, and $\mathcal{P}^*\mathcal{R}^*$, and their standard concentration are implicit in the given formula. A more detailed discussion of this point can be found in a recent review (Gilson et al., 1997).

Although the terms of Eq. 3 are an explicit function of the atomic coordinates for the solute molecules plus solvent, the computational difficulty in achieving numerical convergence in free energies from all-atom simulations requires the development of simplified models. An implicit model of the solvent environment can be derived from a statistical mechanical description of an explicit solute–solvent system (Pratt and Chandler, 1977; Ben-Naim, 1990; Gilson et al., 1997; Roux and Simonson, 1999). The free energy for the $\mathcal{P}^*\mathcal{R}^*$ complex is related to the classical partition function, $Z_{\mathcal{P}^*\mathcal{R}^*}$, and can be written as ($\beta = 1/k_{\text{B}}T$ and k_{B} is the Boltzmann constant),

$$Z_{\mathcal{P}^*\mathcal{R}^*} = \frac{\int\int_{\mathcal{V}} d\mathbf{x}_{\mathcal{P}^*\mathcal{R}^*}\, d\mathbf{y}_{\mathcal{P}^*\mathcal{R}^*} \exp[-\beta U_{\mathcal{P}^*\mathcal{R}^*}(\mathbf{x}_{\mathcal{P}^*\mathcal{R}^*}, \mathbf{y}_{\mathcal{P}^*\mathcal{R}^*})]}{\int_{\mathcal{V}} d\mathbf{y}_{\mathcal{P}^*\mathcal{R}^*} \exp[-\beta U_{\text{s,s}}(\mathbf{y}_{\mathcal{P}^*\mathcal{R}^*})]}, \tag{4}$$

where $\mathbf{x}_{\mathcal{P}^*\mathcal{R}^*}$ and $\mathbf{y}_{\mathcal{P}^*\mathcal{R}^*}$ denote the atomic coordinates of a bound conformational state in the ensemble and the surrounding solvent, respectively, in volume $\mathcal{V}$. The potential energy function is defined as

$$\begin{aligned}U_{\mathcal{P}^*\mathcal{R}^*}(\mathbf{x}_{\mathcal{P}^*\mathcal{R}^*}, \mathbf{y}_{\mathcal{P}^*\mathcal{R}^*}) &= U_{\text{int}}(\mathbf{x}_{\mathcal{P}^*\mathcal{R}^*}) + U_{\text{intra}}(\mathbf{x}_{\mathcal{P}^*\mathcal{R}^*}) \\ &\quad + U_{\text{m,s}}(\mathbf{x}_{\mathcal{P}^*\mathcal{R}^*}, \mathbf{y}_{\mathcal{P}^*\mathcal{R}^*}) + U_{\text{s,s}}(\mathbf{y}_{\mathcal{P}^*\mathcal{R}^*}),\end{aligned} \tag{5}$$

where U_{int} is the interaction energy represented by a sum of Coulombic and van der Waals (vdW) terms, U_{intra} is the intramolecular potential energy of the complex, $U_{\text{m,s}}$ is the solute–solvent potential interaction energy, and $U_{\text{s,s}}$ represents the potential energy of solvent–solvent interactions. Although the configuration integral in Eq. 4 extends over all conformations, it is straightforward to modify the formula to treat either molecule as partly rigid. This will be the case for modeling the U1A protein, and the statistical thermodynamic basis for such an approximation is well documented (see, e.g., Gilson et al., 1997).

The partition function can be rewritten as

$$Z_{\mathcal{P}^*\mathcal{R}^*} = \int_{\mathcal{V}} d\mathbf{x}_{\mathcal{P}^*\mathcal{R}^*} \exp\{-\beta[U_{\text{int}}(\mathbf{x}_{\mathcal{P}^*\mathcal{R}^*}) + U_{\text{intra}}(\mathbf{x}_{\mathcal{P}^*\mathcal{R}^*}) + W(\mathbf{x}_{\mathcal{P}^*\mathcal{R}^*})]\}, \quad (6)$$

where

$$\exp[-\beta W(\mathbf{x}_{\mathcal{P}^*\mathcal{R}^*})] = \frac{\int_{\mathcal{V}} d\mathbf{y}_{\mathcal{P}^*\mathcal{R}^*} \exp\{-\beta[U_{\text{m,s}}(\mathbf{x}_{\mathcal{P}^*\mathcal{R}^*}, \mathbf{y}_{\mathcal{P}^*\mathcal{R}^*}) + U_{\text{s,s}}(\mathbf{y}_{\mathcal{P}^*\mathcal{R}^*})]\}}{\int_{\mathcal{V}} d\mathbf{y}_{\mathcal{P}^*\mathcal{R}^*} \exp[-\beta U_{\text{s,s}}(\mathbf{y}_{\mathcal{P}^*\mathcal{R}^*})]}. \quad (7)$$

Eq. 6 indicates an implicit representation of the solvent acting as an external potential through $W(\mathbf{x}_{\mathcal{P}^*\mathcal{R}^*})$ on the complex with conformation $\mathbf{x}_{\mathcal{P}^*\mathcal{R}^*}$. Similar functions are obtained for the unbound states. One major problem with implicit solvent formulations, as the main results of this paper will illustrate, is extracting the dipolar reorganization from extending $W(\mathbf{x}_{\mathcal{P}^*}) \rightarrow W(\mathbf{x}_{\mathcal{P}^*})$ required of modeling Eq. 1.

The solute–solvent interactions that define ΔW from Eq. 7 applied to the bound and unbound states can be separated into individual terms,

$$\Delta W = \Delta G_{\text{cav}} + \Delta G_{\text{s,vdW}} + \Delta G_{\text{s,ele}}, \quad (8)$$

where ΔG_{cav} is the change in free energy required to form the solute-sized cavities in the solvent, $\Delta G_{\text{s,vdW}}$ is the free-energy change in vdW interactions of inserting solute molecules into the cavities, and $\Delta G_{\text{s,ele}}$ is the change in the free energy of solvent polarization. The cavitation term models the hydrophobic effect and is proportional to the change in solvent-exposed surface area upon association with a constant surface tension, γ,

$$\Delta G_{\text{cav}} = \gamma(A_{\mathcal{P}^*\mathcal{R}^*} - A_{\mathcal{P}^*} - A_{\mathcal{R}^*}), \quad (9)$$

where the A denote surface areas. The value of the proportionality constant γ for a water–vacuum interface depends on the definition of the surface area, either solvent-accessible surface area or molecular surface area (Jackson and Sternberg, 1994).

The solvent polarization term given in Eq. 8 can be conveniently calculated from classical electrostatics by using a thermodynamic cycle of charging and uncharging $\mathcal{P}^*$ and $\mathcal{R}^*$ on complex formation (Gilson and Honig, 1988; Muegge et al., 1997). The expression for $\Delta G_{\text{s,ele}}$ can be rewritten as

$$\Delta G_{\text{s,ele}} = (\Delta G_{\text{s,ele}}^{\mathcal{P}^*\mathcal{R}^{*\prime}} - \Delta G_{\text{s,ele}}^{\mathcal{P}^*}) + (\Delta G_{\text{s,ele}}^{\mathcal{P}^{*\prime}\mathcal{R}^*} - \Delta G_{\text{s,ele}}^{\mathcal{R}^*}) = \Delta\Delta G_{\text{s,ele}}^{\mathcal{P}^*\mathcal{R}^{*\prime}} + \Delta\Delta G_{\text{s,ele}}^{\mathcal{P}^{*\prime}\mathcal{R}^*}, \quad (10)$$

where $\mathcal{P}^{*\prime}$ and $\mathcal{R}^{*\prime}$ are uncharged. Each free energy is determined by applying the reference state of bringing the solvent boundary from infinity to the solvent-accessible surface of a given solute molecule, e.g.,

$$\Delta G_{\text{s,ele}}^{\mathcal{P}^*\mathcal{R}^{*\prime}} = G_{\text{s,ele}}^{\mathcal{P}^*\mathcal{R}^{*\prime}}(\epsilon_{\text{m}}, \epsilon_{\text{s}}) - G_{\text{s,ele}}^{\mathcal{P}^*\mathcal{R}^{*\prime}}(\epsilon_{\text{m}}, \epsilon_{\text{m}}), \quad (11)$$

where ϵ_{s} the solvent dielectric constant. From Eq. 10, the term $\Delta\Delta G_{\text{s,ele}}^{\mathcal{P}^*\mathcal{R}^{*\prime}}$ corresponds to the loss of solute–solvent interaction energy through the partial desolvation of the electrostatically charged $\mathcal{P}^*$ on binding the uncharged $\mathcal{R}^{*\prime}$; and $\Delta\Delta G_{\text{s,ele}}^{\mathcal{P}^{*\prime}\mathcal{R}^*}$ is the loss of solute–solvent interaction energy through the partial desolvation of the electrostatically charged $\mathcal{R}^*$ on binding the uncharged $\mathcal{P}^{*\prime}$. Early studies of using this particular thermodynamic cycle in computing binding free energies are given by Misra and Honig (1995), and Jackson and Sternberg (1995).

The contribution ΔG_{int} contains polar and nonpolar intermolecular interaction terms between $\mathcal{P}^*$ and $\mathcal{R}^*$,

$$\Delta G_{\text{int}} = \Delta G_{\text{m,vdW}} + \Delta G_{\text{m,ele}}, \quad (12)$$

where $\Delta G_{\text{m,vdW}}$ is the vdW contribution and $\Delta G_{\text{m,ele}}$ the electrostatic interaction component. The term $\Delta G_{\text{m,vdW}}$ combined with $\Delta G_{\text{s,vdW}}$ of Eq. 8 can be neglected as an approximation by invoking an enthalpy–entropy compensation phenomenon argument (Nicholls et al., 1991). The argument assumes that the change in dispersion energy between atoms making interactions at the interface of the complex and atoms contacting water in the dissociated state is equal to the loss in side-chain conformational entropy upon association.

To evaluate the net electrostatic contribution to binding, the mean electrical potential (φ) is solved by NLPB equation

$$\nabla \cdot \epsilon(\mathbf{r})\nabla\varphi(\mathbf{r}) - \epsilon(\mathbf{r})\kappa^2(\mathbf{r})\sinh(\varphi(\mathbf{r})) + 4\pi\rho_{\text{f}}(\mathbf{r}) = 0, \quad (13)$$

where ρ_{f} is the interior charge distribution of the fixed positions described by $\mathbf{r}$ of all charges in the solute molecules and κ is the Debye–Hückel parameter. The NLPB model was chosen due to the high charge density of the phosphodiester backbone of the RNA molecule. The electrostatic free energies are calculated from the integrals (Sharp and Honig, 1990)

$$\Delta G_{\text{s,ele}}^{\mathcal{P}^*\mathcal{R}^{*\prime}} = \int [\rho_{\text{f}}^{\mathcal{P}^*\mathcal{R}^{*\prime}} \varphi_{\text{pol}}^{\mathcal{P}^*\mathcal{R}^{*\prime}} - (\rho^{\mathcal{P}^*\mathcal{R}^{*\prime}} \varphi_{\text{pol}}^{\mathcal{P}^*\mathcal{R}^{*\prime}}/2 + \Delta\Pi(\varphi^{\mathcal{P}^*\mathcal{R}^{*\prime}}))] d\nu, \quad (14)$$

$$\Delta G_{\text{s,ele}}^{\mathcal{P}^{*\prime}\mathcal{R}^*} = \int [\rho_{\text{f}}^{\mathcal{P}^{*\prime}\mathcal{R}^*} \varphi_{\text{pol}}^{\mathcal{P}^{*\prime}\mathcal{R}^*} - (\rho^{\mathcal{P}^{*\prime}\mathcal{R}^*} \varphi_{\text{pol}}^{\mathcal{P}^{*\prime}\mathcal{R}^*}/2 + \Delta\Pi(\varphi^{\mathcal{P}^{*\prime}\mathcal{R}^*}))] d\nu, \quad (15)$$

$$\Delta G_{\text{m,ele}}^{\mathcal{P}^*\mathcal{R}^*} = \int [\rho_{\text{f}}^{\mathcal{P}^*\mathcal{R}^*} \varphi_{\text{int}}^{\mathcal{P}^*\mathcal{R}^*} - (\rho^{\mathcal{P}^*\mathcal{R}^*} \varphi_{\text{int}}^{\mathcal{P}^*\mathcal{R}^*}/2 + \Delta\Pi(\varphi^{\mathcal{P}^*\mathcal{R}^*}))] d\nu, \quad (16)$$

where ρ is total charge distribution, φ_{pol} is the potential generated from electronic and orientational polarization, φ_{int} is the intermolecular interaction potential between the solute molecules, and $\Delta\Pi \sim c_{\text{b}}[\cosh(\varphi) - 1]$, where c_{b} is the bulk salt concentration.

The numerical value of ϵ_{m} depends on the representation of the conformational term implemented in modeling Eq. 1. Typically, in an implicit approach, contributions from the native conformers of $\mathcal{P}$ and $\mathcal{R}$ are not explicitly accounted for by sampling conformational space via the distribution functions similar to Eq. 6. Rather, the contribution of dipolar fluctuations to ΔG_{conf} are treated by scaling ϵ_{m} greater than the implicit polarizability limit of 2. This scheme models only Eq. 2 and sets the electrostatic terms to include implicitly the structural reorganization. Caution must be used, however, in applying large values of $\epsilon_{\text{m}} \sim 20$ or higher in treating reliably the charge–charge reorganization (Warshel et al., 1997; Muegge et al., 1997, 1998; Sham et al., 1998). These scaling problems stem from the dielectric heterogeneity and can be removed in part by

partitioning the electrostatic terms based on ionized and polar residues (Olson and Reinke, 2000),

$$\Delta\Delta G_{\mathrm{s,ele}}^{\mathcal{P}^*\mathcal{R}^{*\prime}} = \Delta\Delta G_{\mathrm{s,ele}}^{\mathcal{P}^*(\mathrm{ion})-\mathcal{R}^{*\prime}} + \Delta\Delta G_{\mathrm{s,ele}}^{\mathcal{P}^*(\mathrm{dipole})-\mathcal{R}^{*\prime}} \quad (17)$$

$$\Delta\Delta G_{\mathrm{s,ele}}^{\mathcal{P}^{*\prime}\mathcal{R}^*} = \Delta\Delta G_{\mathrm{s,ele}}^{\mathcal{P}^{*\prime}-\mathcal{R}^*(\mathrm{ion})} + \Delta\Delta G_{\mathrm{s,ele}}^{\mathcal{P}^{*\prime}-\mathcal{R}^*(\mathrm{dipole})} \quad (18)$$

$$\Delta G_{\mathrm{m,ele}}^{\mathcal{P}^*\mathcal{R}^*} = \Delta G_{\mathrm{m,ele}}^{\mathcal{P}^*(\mathrm{ion})-\mathcal{R}^*(\mathrm{ion})} + \Delta G_{\mathrm{m,ele}}^{\mathcal{P}^*(\mathrm{ion})-\mathcal{R}^*(\mathrm{dipole})} + \Delta G_{\mathrm{m,ele}}^{\mathcal{P}^*(\mathrm{dipole})-\mathcal{R}^*(\mathrm{ion})} + \Delta G_{\mathrm{m,ele}}^{\mathcal{P}^*(\mathrm{dipole})-\mathcal{R}^*(\mathrm{dipole})}, \quad (19)$$

where, generally,

$$G_{\mathrm{ele}} = G_{\mathrm{ele}}(\epsilon_{\mathrm{m}}^{\mathrm{ion}}, \epsilon_{\mathrm{m}}^{\mathrm{dipole}}, \epsilon_{\mathrm{s}}), \quad (20)$$

and $\epsilon_{\mathrm{m}}^{\mathrm{ion}}$ and $\epsilon_{\mathrm{m}}^{\mathrm{dipole}}$ are the dielectric constants for ionized residues and polar residues, respectively, $\Delta\Delta G_{\mathrm{s,ele}}^{\mathcal{P}^*(\mathrm{ion})-\mathcal{R}^{*\prime}}$ is the free-energy change where only the ionized residues of $\mathcal{P}^*$ contain atomic charges while $\mathcal{R}^*$ is uncharged, and $\Delta\Delta G_{\mathrm{s,ele}}^{\mathcal{P}^*(\mathrm{dipole})-\mathcal{R}^{*\prime}}$ is the free-energy change where only polar residues of $\mathcal{P}^*$ are charged and $\mathcal{R}^*$ is uncharged. The remaining free energies of Eqs. 18–19 are similarly defined. The parameter $\epsilon_{\mathrm{m}}^{\mathrm{ion}}$ is scaled to large values, whereas $\epsilon_{\mathrm{m}}^{\mathrm{dipole}}$ is restricted to values corresponding to the internal regions of globule.

An alternative approach to the implicit schemes is to model explicitly by all-atom simulations the contribution of ΔG_{conf} as a combination of free-energy terms,

$$\Delta G_{\mathrm{conf}} = \Delta G_{\mathrm{reorg}} + \Delta G_{\mathrm{intra}}, \quad (21)$$

where $\Delta G_{\mathrm{reorg}}$ represents the solute–solvent reorganization free energies and $\Delta G_{\mathrm{intra}}$ is the free energy due to intramolecular strain of $\mathcal{P}^*$ and $\mathcal{R}^*$. Each free-energy term contains nonbonded interactions and, in addition, $\Delta G_{\mathrm{intra}}$ models covalent interactions. We will focus strictly on the electrostatic terms and neglect the remaining contributions. The structural reorganization of $\mathcal{P} \rightarrow \mathcal{P}^*$ (and similarly for $\mathcal{R} \rightarrow \mathcal{R}^*$) can be expressed as a sum over small perturbations from the native unbound structure to the final functional conformer

$$\mathcal{P}(\mathbf{x}_{\mathcal{P}}) \xrightarrow{\lambda_1} \mathcal{P}(\mathbf{x}_1) \xrightarrow{\lambda_2} \mathcal{P}(\mathbf{x}_2) \cdots \xrightarrow{\lambda_{n-1}} \mathcal{P}(\mathbf{x}_{n-1}) \xrightarrow{\lambda_n} \mathcal{P}^*(\mathbf{x}_{\mathcal{P}^*}), \quad (22)$$

where λ_i is the order parameter for describing the transition along the conformational path defined by $\mathbf{x}_i$. The free-energy difference for each change is given by (Zwanzig, 1954)

$$\Delta G_{\lambda_i}^{\mathcal{P}\rightarrow\mathcal{P}^*} = -\beta^{-1}\ln\langle e^{-\beta\Delta U_{\lambda_i}}\rangle_{\lambda_i}, \quad (23)$$

where $\langle\cdots\rangle$ denotes a canonical ensemble average over the potential energy change ΔU_{λ_i} from the perturbation $\lambda_i \rightarrow \lambda_{i+1}$. Eq. 23 is a rigorous treatment of relaxation and can be evaluated by free-energy perturbation methods using MD simulations; however, calculations for macromolecular assemblies are generally computationally prohibitive. Alternatively, an LRA treatment can be constructed for modeling reorganization in the solvation free energies by using the two endpoint structures $\mathcal{P}$ and $\mathcal{P}^*$,

$$\Delta\Delta G_{\mathrm{s,reorg}}^{\mathcal{P}\rightarrow\mathcal{P}^*} = \tfrac{1}{2}(\langle U_{\mathrm{m,s}}^{\mathcal{P}^*} - U_{\mathrm{m,s}}^{\mathcal{P}^{*\prime}}\rangle_{\mathcal{P}^*} + \langle U_{\mathrm{m,s}}^{\mathcal{P}^*} - U_{\mathrm{m,s}}^{\mathcal{P}^{*\prime}}\rangle_{\mathcal{P}^{*\prime}}) - \tfrac{1}{2}(\langle U_{\mathrm{m,s}}^{\mathcal{P}} - U_{\mathrm{m,s}}^{\mathcal{P}^{\prime}}\rangle_{\mathcal{P}} + \langle U_{\mathrm{m,s}}^{\mathcal{P}} - U_{\mathrm{m,s}}^{\mathcal{P}^{\prime}}\rangle_{\mathcal{P}^{\prime}}), \quad (24)$$

where the ensemble averages are over MD trajectories using the potential surface representing interactions between the solute and the explicit solvent. The thermodynamic average corresponding to $\langle\cdots\rangle_{\mathcal{P}^*}$ is conditional on restricting $\mathcal{P}$ to the functional conformation. Eq. 24 can be reduced by $U_{\mathrm{m,s}}^{\mathcal{P}^{\prime}} = U_{\mathrm{m,s}}^{\mathcal{P}^{*\prime}} = 0$ and noting that the solvent in the uncharged state of $\mathcal{P}$ and $\mathcal{P}^*$ does not experience the charge distribution of the solute. This gives

$$\Delta\Delta G_{\mathrm{s,reorg}}^{\mathcal{P}\rightarrow\mathcal{P}^*} = \tfrac{1}{2}(\langle U_{\mathrm{m,s}}^{\mathcal{P}^*}\rangle_{\mathcal{P}^*} - \langle U_{\mathrm{m,s}}^{\mathcal{P}}\rangle_{\mathcal{P}}). \quad (25)$$

The change in solvation of $\mathcal{R}$ from its relaxed unbound conformation to the functional form is similarly determined from

$$\Delta\Delta G_{\mathrm{s,reorg}}^{\mathcal{R}\rightarrow\mathcal{R}^*} = \tfrac{1}{2}(\langle U_{\mathrm{m,s}}^{\mathcal{R}^*}\rangle_{\mathcal{R}^*} - \langle U_{\mathrm{m,s}}^{\mathcal{R}}\rangle_{\mathcal{R}}). \quad (26)$$

If both solute molecules are relaxed sufficiently, Eqs. 25 and 26 provide good estimates of reorganization on solvent polarization due to the reorientation of the dipoles of $\mathcal{P}$ and $\mathcal{R}$.

The internal electrostatic interaction of $\mathcal{P}$ is reorganized to the functional $\mathcal{P}^*$ conformer through the potential energy surface,

$$\Delta U_{\mathrm{intra,ele}}^{\mathcal{P}\rightarrow\mathcal{P}^*} = \Delta U_{\mathrm{qq}} + \Delta U_{\mathrm{q}\mu} + \Delta U_{\mu\mu}, \quad (27)$$

representing interactions between charge–charge, charge–dipole, and dipole–dipole, respectively. Evaluation of Eq. 27 for discriminating between $\mathcal{P}$ and $\mathcal{P}^*$ conformations is particularly challenging by the application of simulation models because of computational instabilities in determining reliable potential energies. As a simplification, we have constructed for the Coulomb term a semimacroscopic scaling relationship based on a modification (Warshel and Åqvist, 1991; Lee et al., 1993; Muegge et al., 1998) of the generalized Born (GB) approximation. The underlying idea of the scaling relationship is a partial compensation between $\Delta\Delta G_{\mathrm{s,reorg}}$ and $\Delta G_{\mathrm{intra,ele}}$. Specifically, the change in self-energy from structural reorganization is opposed by the change in charge distribution of the solute. This offset involves detailed balance of large charge–charge interactions and is nicely illustrated by microscopic simulations of an ion pair in solvent water (Lee et al., 1993). Here, we model the competing energies as $\Delta G_{\mathrm{intra,ele}} \approx -f\Delta\Delta G_{\mathrm{s,reorg}}$, where f is a smooth scaling function that is assumed to depend only upon the reorganization of the protein permanent dipoles described by ϵ_{m}. The functional form of f is motivated by considering the limit $\epsilon_{\mathrm{m}} \rightarrow \epsilon_{\mathrm{s}}$, which yields a free-energy balance of $\Delta G_{\mathrm{conf}} \approx 0$ and in the limit of small $\epsilon_{\mathrm{m}} \sim 2$, ΔG_{conf} reaches its maximal value. The ansatz for treating the interpolation of f between the two extremes is Born's formula. We approximate the energy change of charging the functional conformer from the nonfunctional form and bringing the charges from a gas-phase dielectric to a uniform ϵ_{m} medium by scaling the interaction energy by the factor $(1 - 1/\epsilon_{\mathrm{m}})$. Applying this scaling factor, the free-energy contribution of $\Delta G_{\mathrm{intra,ele}}$ is estimated by the expression,

$$\Delta G_{\mathrm{intra,ele}}^{\mathcal{P}\rightarrow\mathcal{P}^*} \approx \left(\frac{\epsilon_{\mathrm{m}}}{1 - \epsilon_{\mathrm{m}}}\right)\Delta\Delta G_{\mathrm{s,reorg}}^{\mathcal{P}\rightarrow\mathcal{P}^*}, \quad (28)$$

where ϵ_{m} contains the lower bound of the electronic polarizability limit. As noted above, we must again state that the dielectric constant represents all contributions that are not considered explicitly. However, estimates for ϵ_{m} are not selected arbitrarily, but rather are obtained from applying the NLPB equation to conformations extracted from the MD simulation trajectories and by finding an ϵ_{m} value that brings the continuum model in agreement with the LRA, via Eq. 25. Although Eq. 28 is a nonrigorous approach, the term $\Delta\Delta G_{\mathrm{s,reorg}}$ is calculated at the explicit level and the precision of determining $\Delta G_{\mathrm{intra,ele}}$ by a scaling relationship should be no worse than that of computing the difference between two very large microscopic energy terms (see, e.g., Reyes and Kollman, 2000b). In fact, the scaling of microscopic energy terms reduces the absolute value of the energy contribution and thus increases the precision of the calculations (Lee et al., 1993; Sham et al., 1998). A similar formulation is constructed for the $\mathcal{R} \rightarrow \mathcal{R}^*$ transition.

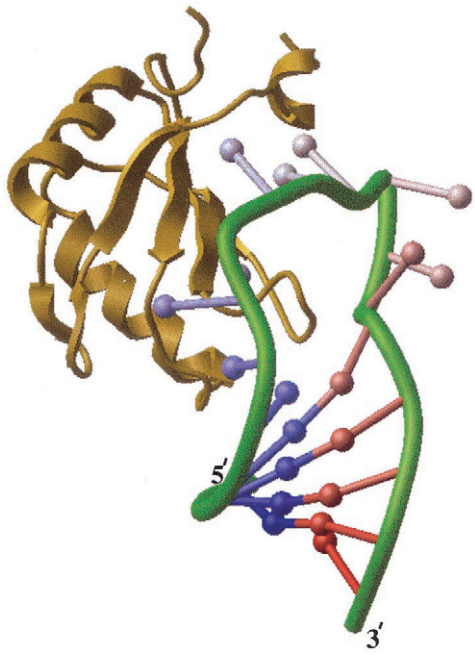

FIGURE 1 Schematic illustration of the x-ray crystallographic structure of the U1A–RNA complex. The U1A protein is depicted in a ribbon representation colored yellow and the RNA hairpin is an oval tube colored green. The RNA bases are depicted as ball-and-stick structures. The figure was constructed with the program ICMlite (R. A. Abagyan and coworkers).

Calculations

Atomic coordinates of the U1A–RNA hairpin complex were extracted from the PDB file 1URN (Oubridge et al., 1994). Crystallographic waters were ignored in the calculations. The initial x-ray structure (Fig. 1) contained two amino acid substitutions, Y31H and Q36R, neither of which interacts with the RNA structure. The RNA molecule consisted of the sequence 5′AAUCCAUUGCACUCCGGAUUU3′. Partial atomic charges for U1A and the RNA were derived from the AMBER force field (Weiner et al., 1986). Single-residue mutants (Y13F, N16V, N18A, E19D, S46A, S48A, L49A, F56Y, R52K, and R52Q) were constructed from the wild-type (WT) complex and, in the case of non-alanine substitutions, the side chains were replaced with the most favorable rotamers.

The finite-difference NLPB method (Reiner and Radke, 1990; Sharp and Honig, 1990) implemented in the program DelPhi (Gilson and Honig, 1988) was applied for continuum calculations of the electrostatic free energies. Electrostatic potentials for each molecule were calculated by using the solvent-accessible surfaces to define regions of low dielectric medium embedded in high dielectric solvent water of ionic strength set at 0.145 M. The AMBER parameter set was used to represent the charges and atomic radii. DelPhi calculations were carried out on a cubic grid of 121^3 grid points. Ionization states were set corresponding to a neutral pH. Full Coulombic boundary conditions were applied for all calculations and the number of nonlinear iterations for solving the NLPB equation was set at 250. Dummy atoms were used to retain an identical scale and position on the grid for the complex and individual structures.

The dielectric constant for the protein and RNA was initially set at either 2 or 4, and, for modeling bulk water, a dielectric of 80 was used. Because DelPhi treats dielectric environments as homogeneous by allowing only one dielectric constant for the protein and RNA, approximations were used in modeling a non-uniform dielectric response. We partitioned only the protein reaction field into ionized and neutral residues by applying ϵ_m^{ion} and ϵ_m^{dipole} set at 25 and 4, respectively (see, e.g., Muegge et al., 1998). Independent DelPhi calculations were carried out by turning off the atomic charges of ionized or polar residues, and the free energies were combined to approximate the true final contribution. Partitioning the reaction fields into ionic and polar components is a reasonable first approach for treating heterogeneity and quantitative errors in the net free energies are acceptably small due the their cancellation calculated from the thermodynamic cycle.

MD simulations for evaluating the LRA were performed using the program DISCOVER (Molecular Simulations, Inc.) on structures representing the bound and unbound conformers starting from the WT complex. Conformational sampling of the U1A structure in the bound state and the native unbound conformer used an active-site region, in which amino acid residues positioned within 8 Å of the RNA in the complex were allowed to move freely. The remaining U1A residues lying in the outer shell were rigidly fixed to their initial positions. Dynamics of the RNA bound and native conformers were treated as unrestrained structures, allowing the molecules to be completely flexible. All unrestrained atomic regions were modeled with a 16-Å layer of explicit solvent by using the TIP3P water model (Jorgensen et al., 1983). The simulations consisted of the WT and mutant protein–RNA complexes bound with 21 sodium counterions and 2965 solvent water molecules; the unbound WT and mutant U1A structures with 1858 water molecules; the unbound RNA structure with 21 sodium counterions and 2633 water molecules; and unfolded U1A in an extended conformation with 8095 water molecules. The sodium counterions were placed adjacent to the charged phosphate groups to neutralize the simulation systems.

Simulations were initiated with 100 cycles of minimization by a conjugate gradient method, followed by 25-ps MD equilibration phase. The initial atomic velocities were assigned from a Boltzmann distribution corresponding to temperature of 298 K. Coulombic interactions were modeled by a cell multipole method (Ding et al., 1992) with a constant dielectric of $\epsilon = 1$, and the vdW interactions were calculated using a group-based approach with a cutoff radius of 12.0 Å. Constraints were applied via the RATTLE algorithm (Andersen, 1983) to bond lengths of the macromolecules plus the solvent during dynamics runs. The integration timestep was set at 2.0 fs and ensemble averages were determined from 100 ps with coordinates saved every 500 time steps for further analysis.

The functional conformers of U1A and the RNA were extracted from the MD simulations of the complexes. Evaluation of the LRA terms of Eqs. 25 and 26 corresponding to $\mathcal{P}^*$ and $\mathcal{R}^*$ used single conformations taken from the trajectories based by a selection criteria of yielding the smallest root-mean-square deviation (rmsd) from the average MD structures of the bound state. The interaction between a functional conformer and solvent was sampled with the solute restricted to its conformation as determined from the MD simulation, while the water molecules were allowed to move freely. The approximation of using an average conformer based on rmsd rather than a Boltzmann average over all MD conformers is reasonable and the sensitivity of the results to the specific conformation was observed to be no greater than ~1 kcal/mol difference. A simulation protocol was applied similar to that described above for the bound and native conformers.

Multiconformers for the WT and mutant bound structures were generated by using a series of short MD simulations. Each structure was analyzed from trajectories extracted from 10 independent simulations, where each MD calculation was carried out to 20 ps of dynamics. Delphi calculations were applied to each saved trajectory. The dielectric constant for the electronic polarization component was set at 2. All individual free-energy contributions were averaged over the simulation runs to yield a final determined value.

Cavitation energies were determined from the molecular surfaces using the Connolly algorithm (Connolly, 1981) with the solvent probed radius set at 1.4 Å. The numerical value of γ was set at 69 cal/mol/Å^2 (Jackson and Sternberg, 1994).

TABLE 1 Conformational reorganization free energies (kcal/mol) for wild-type U1A and RNA hairpin structures

Free Energy Term	LRA	Continuum model				
	$\epsilon_m = 1$	$\epsilon_m = 2$	$\epsilon_m = 4$	$\epsilon_m = 6$	$\epsilon_m = 8$	$\epsilon_m = 10$
$\Delta\Delta G_{s,reorg}^{\mathcal{P}\rightarrow\mathcal{P}^*}$	−6.8	−24.5	−11.8	−10.5	−7.6	−3.4
$\Delta\Delta G_{intra,ele}^{\mathcal{P}\rightarrow\mathcal{P}^*}$	7.8–13.6*	48.9	15.7	12.6	8.7	3.8
$\Delta\Delta G_{conf}^{\mathcal{P}\rightarrow\mathcal{P}^*}$	1.0–6.8	24.4	3.9	2.1	1.1	0.4
$\Delta\Delta G_{s,reorg}^{\mathcal{R}\rightarrow\mathcal{R}^*}$	−2.5	−21.6	−10.3	−6.4	−4.6	−3.5
$\Delta\Delta G_{intra,ele}^{\mathcal{R}\rightarrow\mathcal{R}^*}$	2.7–5.0[a]	43.2	13.7	7.7	5.3	3.9
$\Delta\Delta G_{conf}^{\mathcal{R}\rightarrow\mathcal{R}^*}$	0.2–2.5	21.6	3.4	1.3	0.7	0.4
ΔG_{conf}	1.2–9.3	46.0	7.3	3.4	1.8	0.8

*Calculated using a range of ϵ_m values from 2 to the value obtained from fitting the NLPB calculation to the results determined by the LRA model. Statistical errors ~1–3 kcal/mol.

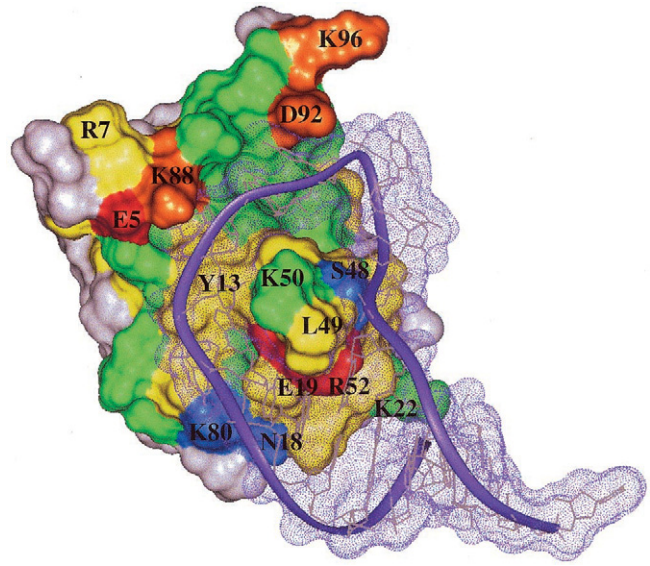

FIGURE 2 The change in free energy of solvation for the U1A transition from the native to functional conformer projected onto the protein molecular surface. The free energy scale (kcal/mol) is as follows: surfaces colored red depict residues that contribute free energies < -2 kcal/mol; $-2 <$ orange < -1; $-1 <$ yellow < 0; $0 <$ green $< +1$; blue $> +1$. The gray-colored surface regions represent no free-energy changes. Bound RNA hairpin is shown colored purple.

RESULTS AND DISCUSSION

Absolute binding free energy

To apply a solute dielectric constant of 2 in the NLPB model to capture the correct physics of protein–ligand binding, structural relaxation at the explicit level must be taken into account (Sham et al., 1998; Warshel and Papazyan, 1998; Muegge et al., 1998). From LRA calculations of relaxation combined with an NLPB model analysis of the simulation trajectories, the electrostatic components to ΔG_{conf} are summarized in Table 1. For the transition $\mathcal{P} \rightarrow \mathcal{P}^*$, $\Delta\Delta G_{s,reorg} = \sim -7$ kcal/mol and indicates that the functional form of U1A exhibits a more favorable solvation free energy than that corresponding to the native conformer. To put this number in proper perspective, the change in self-energy for the interfacial residues from the native conformation to a simulation model of the unfolded state is $\Delta\Delta G_{s,reorg} = \sim -59$ kcal/mol. The interfacial surface of U1A consists primarily of loop structural elements, and ligand binding selects a dipolar orientation more disordered than that of the native conformer. Yet, an ordered state is achieved in which the protein dipoles are orientated with the electrostatic potential of the RNA molecule.

The end effect of the $\mathcal{P} \rightarrow \mathcal{P}^*$ structural reorganization is molecular complementarity with the RNA; however, not all protein residues contribute favorably to the transition. Illustrated in Fig. 2 are residue contributions to $\Delta\Delta G_{s,reorg}$ for the transition projected onto the U1A molecular surface. Surfaces colored red depict residues that strongly favor solvation of the functional conformation ($\Delta\Delta G_{s,reorg} < -2$ kcal/mol) and blue-colored surfaces strongly favor the native conformation ($\Delta\Delta G_{s,reorg} > 1$ kcal/mol). Intermediate are residues moderately influenced by the solvent in either promoting (yellow/orange) or hindering (green) the transition. An interesting feature of the free-energy decomposition at the residue level is the dispersal of "hot" and "cold" regions. A possible significance of the solute–solvent "fingerprint" is cooperativity, where a recent study of protein–protein complexes (Freire, 1999) suggested that low-stability regions might be involved in the transmission of binding information to regions other than the interface. An implication of this, as noted by Nussinov and coworkers (Tsai et al., 1999), is that stable regions contribute to conformational specificity and less stable regions are more flexible in accommodating binding affinity.

The role of protein–solvent reorganization as a determinant of which residues contribute to specificity or binding affinity relates to the magnitude in individual free-energy transitions. In other words, significant changes in solvation generally arise from conformational specificity, whereas highly flexible residues show marginal transitions. Charged residues E19 and R52, critical for the formation of specific U1A side-chain interactions with RNA bases, are strongly facilitated by the solvent in obtaining the functional conformer. For these residues, the solvent acts as a facilitator in shifting the conformer populations. Although both E19 and R52 are positioned in loop regions and show free-energy preferences for hydration of the unfolded state versus the native state, MD simulation of the native structure indicates stable conformations for the side chains (Fig. 3 *a*). Not all conformationally specific ionic-charge interactions are necessarily favorable in solvation changes. For example, K80 forms a hydrogen bond with the base U8 in the simulation structure and prefers the native conformer. Residues important solely for binding affinity of the RNA phosphodiester backbone (e.g., K20, K22, and K23) generally show large

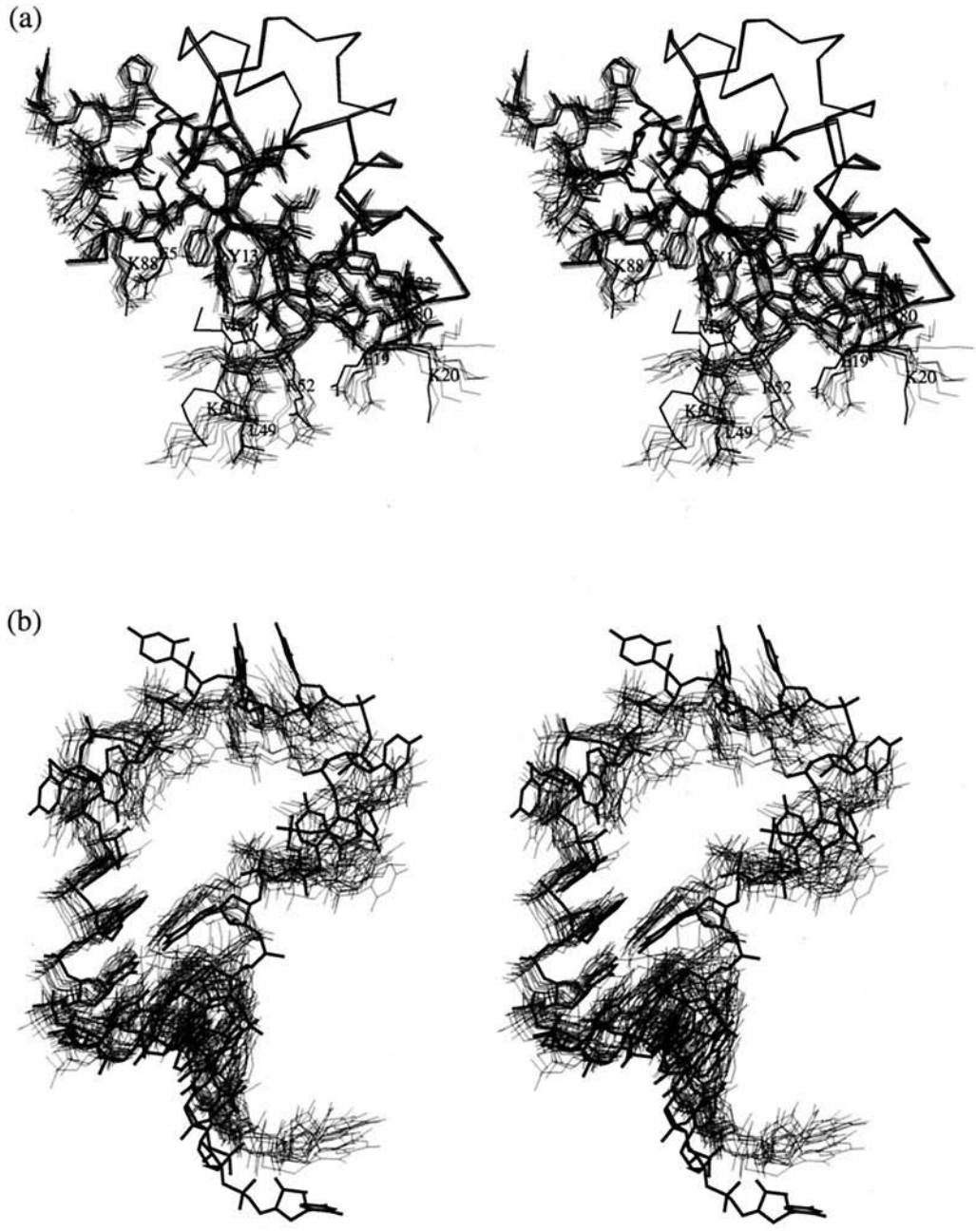

FIGURE 3 Stereo views of the conformational transitions between the native and functional conformers for (*a*) the U1A protein and (*b*) RNA structure. Structures shown as thin lines were extracted from MD simulation trajectories of the native conformers and structures shown as thick lines were taken from the x-ray crystal structure of the bound conformers.

side-chain orientational fluctuations in the MD structures, whereas the local shifts in solvation between functional and native conformers are insignificant. A noted exception is S48, which is part of a flexible loop in the free structure and shows a large unfavorable $\Delta\Delta G_{s,reorg}$. Hydrophobic contacts at the interface from Y13 and L49 display favorable values for $\Delta\Delta G_{s,reorg}$, albeit the transitions as expected are much smaller than ionic interactions.

Although solvation favors the $\mathcal{P}$* conformation, the free-energy barrier to sampling this conformational state for the native protein is the strain energy incurred at the bound conformation. Estimates of this penalty can be obtained from the scoring function of Eq. 28; however, one must first resolve the problem of choosing an optimal value for the protein dielectric constant describing the energetics of the internal transition. One approach to tackling this problem is to require consistency between calculations of an NLPB analysis of the conformational ensemble generated from the MD trajectory and the LRA results. This consistency condition yields a linear relationship between the explicit and

TABLE 2 Absolute free energy of binding (kcal/mol) without explicit relaxation for the U1A–RNA complex*

ϵ_m	$\Delta\Delta G^{\mathcal{P}*\mathcal{R}*\prime}_{s,ele}$	$\Delta\Delta G^{\mathcal{P}*\prime\mathcal{R}*}_{s,ele}$	$\Delta G^{\mathcal{P}*\mathcal{R}*}_{m,ele}$	$\Delta G^{\mathcal{P}*\mathcal{R}*}_{ele}$	$\Delta G^{\mathcal{P}*\mathcal{R}*}_{cav}$	ΔG_{conf}	$-T\Delta S$	ΔG_{bind}	ΔG_{expt}
2	57.6	42.8	−47.6	52.8	−74.7	1.2–9.3†	~10	−11–−2	−14.2
4	28.1	23.4	−28.8	22.7	−74.7	10.5–18.6‡	~10	−32–−23	
4, 25	9.9	23.4	−20.3	13.0	−74.7	10.5–18.6‡	~10	−41–−33	

*Experimental data taken from Williams and Hall (1996).
†Combination of solvation free energy and strain energy.
‡Strain energy only.

implicit models, and leads to a natural description of molecular association, namely, multiple dielectric constants reflecting different aspects of binding. Here, the result is the application of two dielectric constants for the implicit model, one treating the static component of electronic polarizability and the second treating relaxation involving the motion of charged or polar atoms. This idea of a multistep component analysis is particularly applicable to modeling mutations and is conceptually similar to a recent study of charge insertion in an enzyme-active site (Simonson et al., 1999). From Table 1, $\Delta\Delta G_{s,reorg}$, calculated from the NLPB model, is consistent with the free energy determined by the LRA for a value of $\epsilon_m \sim 8$. This value for the dielectric constant correctly accounts for relaxation (i.e., >2) and indicates that the residues responsible for the conformational transition are moderately polarizable. Using this value of ϵ_m in Eq. 28, the electrostatic "strain energy" is estimated at $\Delta G_{intra,ele} = \sim 8$ kcal/mol. Alternatively, using an ϵ_m of 2 or 4 in the scoring function yields $\Delta G_{intra,ele} = \sim 14$ kcal/mol and 9 kcal/mol, respectively. Combining the strain energy term with the transition in solvation energy, the shift in the static electrostatic free energy for U1A is $\Delta G_{conf} \sim$ 1–7 kcal/mol. The calculated range for ΔG_{conf} is physically reasonable, showing a value less than the free energy of unfolding for the U1A protein, which is 9 kcal/mol at neutral pH (Lu and Hall, 1997).

For the $\mathcal{R} \rightarrow \mathcal{R}*$ transition of the RNA, $\Delta\Delta G_{s,reorg} = -2.5$ kcal/mol from the LRA model and the NLPB calculation requires an $\epsilon_m \sim 14$ to bring the self-energy in agreement with the LRA. This yields a dielectric constant for the RNA that provides for significant dipolar reorientation and is notably higher than that obtained for modeling the U1A protein reorganization. From the simulation models, the rmsd between the bound and unbound RNA conformers is greater than 3 Å (Fig. 3 *b*), whereas for the protein, the rmsd is ~1 Å. Longer simulations of the free RNA show an rmsd > 5 Å (Tang and Nilsson, 1999). The conformational flexibility and the modeled dielectric relaxation of the RNA is consistent with the observation from NMR studies of the free U1 snRNA hairpin that the loop is largely unstructured in free RNA (Allain and Varani, unpublished; referenced by Oubridge et al., 1994). The calculated shift in the RNA free energy upon binding U1A is in the range of 0.2–2.5 kcal/mol. The small magnitude of ΔG_{conf} suggests a rugged free-energy topology with a large ensemble of RNA conformers easily accessible to the functional conformation by thermal fluctuations and without any significant electrostatic penalty.

Given the caveats of the GB-modeled scoring function of Eq. 28 and the approximation that the neglected terms of Eq. 21 offset their net contribution to reorganization, the total ΔG_{conf} for association is in the range of 1.2–9.3 kcal/mol. Conformational sampling implemented in arriving at the free-energy shift was limited to relaxation of local structural changes, rather than the computationally more difficult problem of calculating global changes. We should note, in particular, that the MD calculations were carried out to relatively short times, and full convergence may not have been achieved in sampling complete torsional rotations. Nevertheless, the length of the simulations are typical of LRA calculations (Åqvist et al., 1994; Muegge et al., 1998) and extending the simulation time suggests that convergence in ϵ_m was adequate. An important point is that the calculated variation in ϵ_m from the two structures is clearly much larger than the fluctuations in ϵ_m from sampling a single structure. A recent application of a more global approach using a combination of molecular mechanics and a PB model with ϵ_m set to unity similarly reports a net unfavorable electrostatic contribution to the conformational transition for U1A–RNA hairpin binding (Reyes and Kollman, 2000b). Although the magnitude of the individual electrostatic terms calculated by the two computational approaches differ, both show the fundamental result of solvation promoting the shift in populations on the energy landscape favoring the functional conformers.

Table 2 presents the individual free-energy terms that contribute to the molecular association between the U1A protein and the RNA hairpin. The static terms were determined from the single-conformer model describing Eq. 2. Each calculation displays highly favorable electrostatic intermolecular interactions, yet the net electrostatic component is unfavorable due to the desolvation cost of releasing the interfacial waters from the protein and RNA surfaces. Unlike other modeled protein–nucleic acid complexes (Misra et al., 1998; Olson and Cuff, 1999), where the desolvation cost of the nucleic acid is much larger than the protein, displacement of the waters at the RNA surface for a single dielectric model is more favorable than the desolvation cost of the U1A protein. The lack of significant buried surface of the RNA phosphodiester backbone in the

low-dielectric molecular interface provides for the low desolvation penalty. Structurally, the ten-nucleotide RNA loop binds to the surface of U1A as an open structure (see Fig. 1) and the nucleotide sequence of the loop interacts extensively through stacking of RNA bases with side chains of the protein. The solute–solvent cavitation offsets the electrostatic penalty and stabilizes complex formation. This compensation from hydrophobicity mirrors results calculated for protein–protein complexes using various PB approaches (see, e.g., Olson, 1998) and is similar to that previously noted for proteins binding nucleic acids (Misra et al., 1998; Olson and Cuff, 1999).

The calculated net value of $\Delta\Delta G_{s,reorg}$ at ~ -9 kcal/mol for U1A–RNA complex formation indicates that the implicit models of $\epsilon_m = 4$ and the non-uniform approach of using the combination of $\epsilon_m = 4$ and 25 overestimated the contribution due to dipolar transitions. Consequently, an empirical value of ϵ_m for the WT complex is 2, a value equivalent to that used a priori in the two reported studies of protein–nucleic acid complexes (Misra et al., 1998; Olson and Cuff, 1999). This leaves an entropic cost to be calculated in reconciliation of ΔG_{bind} with ΔG_{expt}. Although the total entropic change on macromolecular association can be determined by calorimetric experiments, interpretation of the results in terms of individual energetic contributions is not clear. The loss of translational and rotational conformational degrees of freedom is difficult to calculate by rigorous methods and their magnitude on protein–ligand interactions has led to considerable debate (Karplus and Janin, 1999; Privalov and Tamura, 1999). Nevertheless, some progress has been made recently in calculations of binding entropy (Hermans and Wang, 1997; Sham et al., 2000). Here, we will take theoretical estimates of the translational and rotational loss for protein–protein assemblies to be in the range of 7–15 kcal/mol at room temperature (Karplus and Janin, 1999; and references cited therein). Using a value of $\sim$10 kcal/mol, we obtained a ΔG_{bind} of ~ -2 to -11 kcal/mol (reported in Table 2), whereas the experimental binding affinity is -14.2 kcal/mol (Williams and Hall, 1996). In comparison with other reported continuum model predictions of absolute binding affinities (Jackson and Sternberg, 1995; Froloff et al., 1997; Olson, 1998; Olson and Cuff, 1999), the predictions outlined here are quite reasonable for the U1A–RNA complex. It should be mentioned that the estimated entropy used in the calculations is much less than that determined by Reyes and Kollman (2000b), and the latter might be difficult to reconcile with values suggested by other researchers (Karplus and Janin, 1999; Hermans and Wang, 1997; Sham et al., 2000).

Differential binding free energies

Because of the general computational simplicity of PB models, these approaches are currently popular for evaluating energetic changes on macromolecular complexes due to amino acid substitutions. Table 3 presents the relative free-energy changes determined by the NLPB model for 10 single-residue protein substitutions without any explicit reorganization. The calculations are compared with experimental data taken from Nagai and coworkers (Jensen et al., 1991). Note that, for several of the mutants, somewhat different values for $\Delta\Delta G_{expt}$ have been reported, particularly Y13F, where the relative change is 2.4 kcal/mol (Kranz and Hall, 1998). Nonetheless, the general conclusions of the calculations remain essentially unaffected. A scatter plot of the calculated change in free energy versus $\Delta\Delta G_{expt}$ is presented in Fig. 4 *a*.

Results for the two uniform ϵ_m models show mean absolute errors of ± 1.4 kcal/mol for $\epsilon_m = 2$ and ± 1.1 kcal/mol for $\epsilon_m = 4$. For mutants Y13F, N16V, and S46A, modeling only static polarization effects of the electrostatic contribution appears to be adequate for obtaining reasonable agreements with experiments without the need of including conformational transitions. Treating the protein dielectric heterogeneity yields no significant improvement in the mean error from the $\epsilon_m = 4$ model (typically the standard protocol in applying continuum models), although, in several cases, better quantitative agreement is achieved. For example, the two-state dielectric model for R52K reduces the error from 0.8 to 0.1 kcal/mol. More promising overall results for partitioning the dielectric environment were obtained in modeling protein–protein complexes (Olson and Reinke, 2000). However, improvement may be made if a field correction is incorporated in the hydrophobic effect (Muegge et al., 1998), thus reducing ΔG_{cav}. Further examples are needed to gauge the applicability of this computational strategy. Altogether, the calculations for U1A–RNA complex were able to discriminate between many of the mutational effects and to detect the significance of Arg^{52} in binding the RNA. The errors are quite encouraging for modeling mutations of protein–RNA complexes by using continuum methods and are comparable with similar modeling studies of other macromolecular complexes (Novotony et al., 1997; Olson, 1998; Sharp, 1999; Olson and Cuff, 1999; Olson and Reinke, 2000).

To explore the application of the NLPB–LRA formalism in modeling the free-energy change underlying conformational reorganization, five mutants were selected: Y13F, N18A, E19D, L49A, and R52Q. From the single-conformer (nonrelaxation) results of Table 3, two of these mutants, E19D and L49A, show significant changes due to hydrophobic effects at the interface, and a simple modification of the total electrostatics by scaling ϵ_m will not reconcile the differences between theory and experiment. The mutant R52Q was chosen to examine the performance of the NLPB–LRA approach in evaluating mutations where binding is significantly reduced, as well as the effect of charge deletions on calculating ϵ_m.

TABLE 3 Relative free-energy changes (kcal/mol) without explicit relaxation for U1A mutants*

Mutant	ϵ_m	$\Delta\Delta G^{\mathcal{P}*\mathcal{R}*\prime}_{s,ele}$	$\Delta\Delta G^{\mathcal{P}*\prime\mathcal{R}*}_{s,ele}$	$\Delta\Delta G^{\mathcal{P}*\mathcal{R}*}_{m,ele}$	$\Delta\Delta G^{\mathcal{P}*\mathcal{R}*}_{ele}$	$\Delta\Delta G^{\mathcal{P}*\mathcal{R}*}_{cav}$	$\Delta\Delta G_{bind}$	$\Delta\Delta G_{expt}$
Y13F	2	1.2	−0.1	0.0	1.1	0.0	1.1	1.9
	4	0.6	0.1	−0.2	0.5	0.0	0.5	
	4, 25	0.5	0.1	0.1	0.7	0.0	0.7	
N16V	2	0.0	0.0	0.9	0.9	0.0	0.9	1.6
	4	0.1	0.1	0.2	0.4	0.0	0.4	
	4, 25	−0.3	0.1	0.4	0.2	0.0	0.2	
N18A	2	−0.1	−0.1	0.0	−0.2	0.0	−0.2	1.0
	4	0.0	0.1	−0.2	−0.1	0.0	−0.1	
	4, 25	−0.1	0.1	0.0	0.0	0.0	0.0	
E19D	2	−4.9	−0.9	6.2	0.4	2.0	2.4	0.4
	4	−2.4	−0.3	2.8	0.1	2.0	2.1	
	4, 25	−0.2	−0.3	0.5	0.0	2.0	2.0	
S46A	2	0.0	−0.2	0.5	0.3	0.1	0.4	0.4
	4	0.0	0.0	0.1	0.1	0.1	0.2	
	4, 25	−0.1	0.0	0.3	0.2	0.1	0.3	
S48A	2	−0.9	−0.3	0.8	−0.4	−0.1	−0.5	0.4
	4	−0.4	1.5	0.2	1.3	−0.1	1.2	
	4, 25	−0.4	1.5	0.5	1.6	−0.1	1.5	
L49A	2	−0.5	1.1	0.1	0.7	1.8	2.5	0.4
	4	−0.2	1.3	−0.1	1.0	1.8	2.8	
	4, 25	−0.1	1.3	0.0	1.2	1.8	3.0	
F56Y	2	1.1	0.2	0.8	2.1	−0.3	1.8	0.7
	4	0.5	0.2	0.2	0.9	−0.3	0.6	
	4, 25	0.5	0.2	0.4	1.1	−0.3	0.8	
R52K	2	0.0	−1.2	2.8	1.6	0.6	2.2	0.7
	4	0.1	−0.5	1.3	0.9	0.6	1.5	
	4, 25	0.1	−0.5	0.4	0.0	0.6	0.6	
R52Q	2	−3.5	−1.7	11.3	6.1	1.6	7.7	>4.0
	4	−1.9	−0.7	6.9	4.3	1.6	5.9	
	4, 25	0.4	−0.7	4.2	3.9	1.6	5.5	

*Experimental data taken from Jensen et al. (1991).

Summarized in Table 4 are the conformational transition free energies for the mutants. Because of the smaller net change in the calculated ΔG_{conf} of the RNA in the WT complex, we applied the LRA method only to the protein interactions that contribute to the free energy of binding. The results show that, on account of the structural perturbations, the protein–solvent reorganization varies significantly among the different mutants. Each calculated reorganization term by the LRA method directly reflects the electrostatic potential felt by the protein and explicit solvent atoms. Consequently, a physical basis is provided for understanding the solvation free energies of incorporating amino acid substitutions. The large increase in the R52Q $\Delta G_{s,reorg}$ from the WT structure indicates a considerable change in solvent polarization due to the charge deletion, whereas, in contrast, the change for Y13F is much smaller. None of the mutants modified the solvation preference for the functional conformation, and two of the amino acid substitutions increased the equilibrium shift greater than the WT transition.

A variation is similarly observed for the reorganization dielectric constant calculated from fitting the NLPB model to the LRA result. Approximate values obtained for ϵ_m are 7 for Y13F, 3 for N18A, 4 for E19D, 6 for L49A, and 18 for R52Q. Values of the dielectric constants are comparable to the dielectric constants approximated from MD simulations (King et al., 1991; Simonson and Perahia, 1995; Simonson and Brooks, 1996). Moreover, the calculations correctly showed that, accounting for conformational flexibility, an ϵ_m value is necessarily greater than the static limit of 2. Consistent with the WT complex, the continuum model leads to a representation with two solute dielectric constants. The NLPB calculations further indicate that ϵ_m is very different for charged and uncharged deletions. This result is not very surprising and reflects the current trend in continuum modeling of applying large ϵ_m values of ~20 for handling the overestimation of charge–charge interactions (see, e.g., Olson, 1998). The NLPB results also suggest a correspondence between the magnitude of ϵ_m and the extent of conformational flexibility, although caution must be used

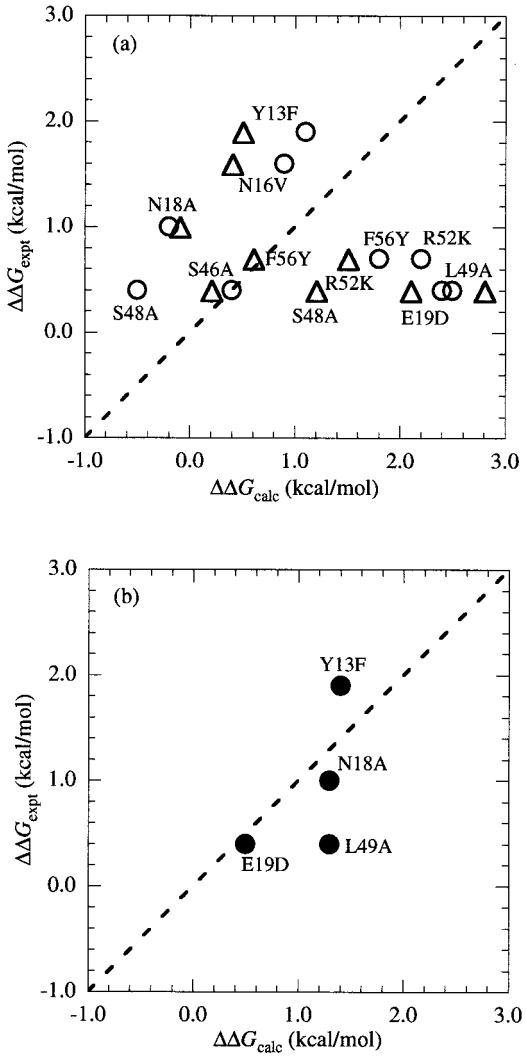

FIGURE 4 Scatter plots of $\Delta\Delta G_{calc}$ versus $\Delta\Delta G_{expt}$ for single-residue substitutions. (*a*) Results of the NLPB equation with the macromolecule dielectric constant modeled by using a value of 2 (*circles*) or 4 (*triangles*). (*b*) Results of the NLPB–LRA computational approach with the individual free energies calculated as a sum of the static polarizability limit by using a dielectric constant of 2 combined with the structural transition term determined by the LRA. The calculated points for R52Q are not included in the plots.

because of the implicit formulation of ϵ_m. With the dielectric constants, each protein–solvent reorganization term is offset by the transition in the internal energy, with N18A showing the greatest conformational penalty.

An important result of linking the NLPB model with the LRA is a clear demonstration of the difficulty in deriving a consistent value for ϵ_m in applying a strictly implicit scheme to modeling conformational flexibility. This indicates that, although MD simulations applied to the structural relaxation of $\mathcal{P}^*$ can sample significant local reorganization, a meaningful NLPB analysis of the conformations from the trajectories requires the correct a priori choice of ϵ_m. This is of practical concern for continuum models, because the results of Table 4 show that there is no single optimum ϵ_m that treats all substitutions. A direct implication of this is the difficulty in applying the hydration free energy calculated by standard continuum methods to the problem of distinguishing between native and non-native conformations, particularly the selection of loop conformations (e.g., Pellequer and Chen, 1997). The results further show that applying the standard protocol of $\epsilon_m = 4$ to calculating the change in $\Delta G_{s,reorg}$ relative to the WT U1A yields free energies that, when added to the static polarization term, does not improve the overall predictions. For example, calculations for Y13F indicate $\Delta\Delta G_{s,reorg} = -0.9$ kcal/mol, and, when combined with $\Delta\Delta G_{bind}$ of 1.1 kcal/mol (Table 3 calculated with $\epsilon_m = 2$), the net result is 0.2 kcal/mol; a value in the wrong direction when compared with experiments. Clearly, adding the $\Delta\Delta G_{intra,ele}$ contribution has little effect on improving the overall accuracy of the predictions when ϵ_m is incorrect.

Rather than relaxing only the unbound state, a multiconformer model must be applied to the complete thermodynamic cycle. Table 5 and Fig. 4 report the results obtained with full relaxation of the individual energetic components that contribute to ΔG_{static}, combined with the conformational free energies taken from the LRA with ϵ_m determined via the NLPB model. Several observations can be made of the calculated results. The first is that MD sampling of only the complex configuration and using the generated bound $\mathcal{P}^*$ and $\mathcal{R}^*$ conformers in NLPB calculations of the thermodynamic cycle failed to help the reconciliation of $\Delta\Delta G_{static}$ with experimental data for all mutants. This is particularly evident for E19D, where the multiconformer $\Delta\Delta G_{static}$ still displays considerable error with $\Delta\Delta G_{expt}$, although the trend is in the right direction from the single-conformer free-energy value of Table 3. One mutant where relaxation significantly helps is L49A. Nevertheless, calculations that are more consistent are obtained by including at the explicit level the $\Delta\Delta G_{conf}$ by the LRA method. An accurate modeling procedure cannot generally be developed from merely adding the $\Delta\Delta G_{conf}$ contribution to the single-conformer results. Again E19D provides an example, where $\Delta\Delta G_{conf} = 1.7$ kcal/mol and the addition of this term to the single-conformer $\Delta\Delta G_{bind}$ of 2.4 kcal/mol from Table 3 failed to improve the results. The second observation is that conformational averaging combined with the explicit determination of the reorganization component removes the inconsistencies in the application of ϵ_m. The NLPB model contribution that is calculated with the static limit of including only electronic polarizability ensures a proper treatment of the nonrelaxation free energies, whereas LRA accounts for the dipolar transitions without forcing homogeneity in ϵ_m.

Although the results of the NLPB–LRA calculations show excellent overall agreement with experiments, we conclude with a few cautionary comments regarding the statistical errors. Because of computational instabilities typically encountered in the relaxation of global structural changes, conformational sampling was limited to relaxation of local changes.

TABLE 4 Conformational reorganization free energies (kcal/mol) for U1A mutants

Mutant	Free Energy Term	LRA	Continuum model				
		$\epsilon_m = 1$	$\epsilon_m = 2$	$\epsilon_m = 4$	$\epsilon_m = 6$	$\epsilon_m = 8$	$\epsilon_m = 10$
Y13F	$\Delta\Delta G^{\mathcal{P}\to\mathcal{P}^*}_{s,reorg}$	−7.2	−26.0	−12.7	−8.2	−6.0	−4.6
	$\Delta\Delta G^{\mathcal{P}\to\mathcal{P}^*}_{intra,ele}$	8.4*	52.0	16.9	9.8	6.9	5.1
	$\Delta\Delta G^{\mathcal{P}\to\mathcal{P}^*}_{conf}$	1.2	26.0	4.2	1.6	0.9	0.5
N18A	$\Delta\Delta G^{\mathcal{P}\to\mathcal{P}^*}_{s,reorg}$	−5.6	−8.2	−4.0	−2.6	−1.9	−1.4
	$\Delta\Delta G^{\mathcal{P}\to\mathcal{P}^*}_{intra,ele}$	8.7*	16.4	5.3	3.1	2.2	1.6
	$\Delta\Delta G^{\mathcal{P}\to\mathcal{P}^*}_{conf}$	3.1	8.2	1.3	0.5	0.3	0.2
E19D	$\Delta\Delta G^{\mathcal{P}\to\mathcal{P}^*}_{s,reorg}$	−8.2	−18.0	−8.8	−5.8	−4.2	−3.3
	$\Delta\Delta G^{\mathcal{P}\to\mathcal{P}^*}_{intra,ele}$	10.9*	36.0	11.7	7.0	4.8	3.7
	$\Delta\Delta G^{\mathcal{P}\to\mathcal{P}^*}_{conf}$	2.7	18.0	2.9	1.2	0.6	0.4
L49A	$\Delta\Delta G^{\mathcal{P}\to\mathcal{P}^*}_{s,reorg}$	−8.2	−26.3	−12.6	−8.1	−5.8	−4.5
	$\Delta\Delta G^{\mathcal{P}\to\mathcal{P}^*}_{intra,ele}$	9.8*	52.6	16.8	9.7	6.6	5.0
	$\Delta\Delta G^{\mathcal{P}\to\mathcal{P}^*}_{conf}$	1.6	26.3	4.2	1.6	0.8	0.5
R52Q	$\Delta\Delta G^{\mathcal{P}\to\mathcal{P}^*}_{s,reorg}$	−1.1	−32.1	−15.5	−9.9	−7.2	−5.5
	$\Delta\Delta G^{\mathcal{P}\to\mathcal{P}^*}_{intra,ele}$	1.2*	64.2	20.7	11.9	8.2	6.1
	$\Delta\Delta G^{\mathcal{P}\to\mathcal{P}^*}_{conf}$	0.1	32.1	5.2	2.0	1.0	0.6

*Calculated using ϵ_m values of ~7 for Y13F, 3 for N18A, 4 for E19D, 6 for L49A, and 18 for R52Q. Statistical errors ~1–3 kcal/mol.

Nevertheless, propagation of the statistical variances for the free-energy terms are quite significant (Table 5) and arise primarily from sampling electrostatic interactions at the complex configuration. The large fluctuations observed in the simulation trajectories of relaxing the interfacial perturbations are similarly problematic of other MD strategies applied to the U1A–RNA complex (Reyes and Kollman, 2000a). The LRA method of calculating free energies of solvation combined with the corresponding NLPB analysis are numerically more stable, resulting in errors of ~1–3 kcal/mol (Tables 2 and 4), which are comparable with reported protein–protein binding studies (see, e.g., Muegge et al., 1998). Without resorting to the significant computational demands of free-energy perturbation methods, the magnitudes of statistical errors render the calculations of $\Delta\Delta G_{bind}$ to be qualitative. Yet, the NLPB–LRA approach is sufficiently reliable and thus provides a powerful tool for gaining valuable information regarding the dielectric response underlying the variations in the free-energy values.

I thank C. Millard, M. Bryne, and F. Lebeda for many valuable discussions. Comments from the reviewers are also appreciated.

TABLE 5 Relative free-energy changes (kcal/mol) for U1A mutants calculated with full relaxation*

Mutant	$\Delta\Delta G_{ele}$	$\Delta\Delta G_{cav}$	$\Delta\Delta G_{static}$	$\Delta\Delta G_{conf}$	$\Delta\Delta G_{bind}$	$\Delta\Delta G_{expt}$
Y13F	1.8	−0.6	1.2	0.2	1.4	1.9
N18A	0.3	−1.1	−0.8	2.1	1.3	1.0
E19D	0.1	−1.3	−1.2	1.7	0.5	0.4
L49A	−2.0	2.7	0.7	0.6	1.3	0.4
R52Q	4.7	1.4	6.1	−0.9	5.2	>4.0

*Errors from averaging over the simulations for $\Delta\Delta G_{bind}$ range from ±5.3 kcal/mol to ±13.0 kcal/mol.

This work was supported by a generous grant of computer time from the Advanced Biomedical Computing Center of the Frederick Cancer Research and Development Center, National Cancer Institute.

REFERENCES

Åqvist, J., C. Medina, and J. E. Samuelsson. 1994. New method for predicting binding-affinity in computer-aided drug design. *Protein Eng.* 7:385–391.

Andersen, H. C. 1983. Rattle: a "velocity" version of the Shake algorithm for molecular dynamics calculations. *J. Chem. Phys.* 52:24–34.

Ben-Naim, A. 1990. Solvent effects on protein association and protein folding. *Biopolymers.* 29:567–596.

Connolly, M. L. 1981. Molecular surface program 429. The quantum chemistry program exchange, Indiana University, Bloomington, IN.

Del Buono, G., F. Figueirido, and R. Levy. 1994. Intrinsic pK_a's of ionizable residues in proteins: an explicit solvent calculation of lysozyme. *Proteins.* 20:85–97.

Ding, H. Q., N. Karasawa, and W. A. Goddard, III. 1992. Atomic level simulations on a million particles: the cell multipole method for Coulomb and London nonbonded interactions. *J. Chem. Phys.* 97: 4309–4315.

Freire, E. 1999. The propagation of binding interactions to remote sites in proteins: analysis of the binding of the monoclonal antibody D1.3 to lysozyme. *Proc. Natl. Acad. Sci. U.S.A.* 96:10118–10122.

Gilson, M. K., J. A. Given, B. L. Bush, and J. A. McCammon. 1997. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.* 72:1047–1069.

Gilson, M. K., and B. Honig. 1988. Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis. *Proteins.* 4:7–18.

Hermans, J., and L. Wang. 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.

Jackson, R. M., and M. J. E. Sternberg. 1994. Application of scaled particle theory to model the hydrophobic effect: implications for molecular association and protein stability. *Protein Eng.* 7:371–383.

Jackson, R. M., and M. J. E. Sternberg. 1995. A continuum model for protein–protein interactions: application to the docking problem. *J. Mol. Biol.* 250:258–275.

Jensen, T. H., C. Oubridge, C. H. Teo, C. Pritchard, and K. Nagai. 1991. Identification of molecular contacts between the U1A small nuclear ribonucleoprotein and U1 RNA. *EMBO J.* 10:3447–3456.

Jones-Hertzog, D. K., and W. L. Jorgensen. 1997. Binding affinities for sulfonamide inhibitors with human thrombin using Monte Carlo simulations with a linear response method. *J. Med. Chem.* 40:1539–1549.

Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulation of liquid water. *J. Chem. Phys.* 79:926–935.

Karplus, M., and J. Janin. 1999. Comment on: the entropy cost of protein association. *Protein Eng.* 12:185–186.

Kranz, J. K., and K. B. Hall. 1998. RNA binding mediates the local cooperativity between the beta-sheet and the C-terminal tail of the human U1A RBD1 protein. *J. Mol. Biol.* 275:465–481.

Kranz, J. K., and K. B. Hall. 1999. RNA recognition by the human U1A protein is mediated by a network of local cooperative interactions that create the optimal binding surface. *J. Mol. Biol.* 285:215–231.

Kranz, J. K., J. R. Lu, and K. B. Hall. 1996. Contribution of the tyrosines to the structure and function of human U1A N-terminal RNA binding domain. *Protein Sci.* 5:1567–1583.

King, G., F. Lee, and A. Warshel. 1991. Microscopic simulations of macroscopic dielectric constants of solvated proteins. *J. Chem. Phys.* 95:4366–4377.

Lee, F. S., Z. T. Chu, M. B. Bolger, and A. Warshel. 1992. Calculations of antibody–antigen interactions: microscopic and semi-microscopic evaluation of the free energies of binding of phosphorylcholine analogs to McPC603. *Protein Eng.* 5:215–228.

Lee, F. S., Z.-T. Chu, and A. Warshel. 1993. A microscopic and semimicroscopic calculations of electrostatic energies in proteins by the POLARIS and ENZYMIX programs. *J. Comput. Chem.* 14:161–185.

Lu, J., and K. B. Hall. 1997. Thermal unfolding of the N-terminal RNA binding domain of the human U1A protein studies by differential scanning calorimetry. *Biophys. Chem.* 64:111–119.

Misra, V. K., and B. Honig. 1995. On the magnitude of the electrostatic contribution to ligand–DNA interactions. *Proc. Natl. Acad. Sci. U.S.A.* 92:4691–4695.

Misra, V. K., J. L. Hecht, A.-S. Yang, and B. Honig. 1998. Electrostatic contributions to the binding free energy of the λcI repressor to DNA. *Biophys. J.* 75:2262–2273.

Muegge, I., H. Tao, and A. Warshel. 1997. A fast estimate of electrostatic group contributions to the free energy of protein-inhibitor binding. *Protein Eng.* 10:1363–1372.

Muegge, I., T. Schweins, and A. Warshel. 1998. Electrostatic contributions to protein–protein binding affinities: application to Rap/Raf interaction. *Proteins.* 30:407–423.

Nicholls, A., K. A. Sharp, and B. Honig. 1991. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins.* 11:281–296.

Novotny, J., R. E. Bruccoleri, M. Davis, and K. A. Sharp. 1997. Empirical free energy calculations—a blind test and further improvements to the method. *J. Mol. Biol.* 268:401–411.

Pellequer, J. L., and S. W. Chen. 1997. Does conformational free energy distinguish loop conformations in proteins? *Biophys. J.* 73:2359–2375.

Olson, M. A. 1998. Mean-field analysis of protein–protein interactions. *Biophys. Chem.* 75:115–128.

Olson, M. A. 1999. Continuum models of macromolecular association in aqueous solution. *In* Computational Chemistry: Reviews of Current Trends. Vol. 4. Jerzy Leszczynski, editor. World Scientific Publishing Co. Pte. Ltd., Singapore. 153–190.

Olson, M. A., and L. Cuff. 1999. Free energy determinants of binding the rRNA substrate and small ligands to ricin A-chain. *Biophys. J.* 76:28–39.

Olson, M. A., and L. T. Reinke. 2000. Modeling implicit reorganization in continuum descriptions of protein–protein interactions. *Proteins.* 38: 115–119.

Oubridge, C., H. Ito, P. R. Evans, C. H. Teo, and K. Nagai. 1994. Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature.* 372: 432–438.

Pratt, L. R., and D. Chandler. 1977. Theory of the hydrophobic effect. *J. Chem. Phys.* 67:3683–3704.

Privalov, P. L., and A. Tamura. 1999. Comments on the comments. *Protein Eng.* 12:187.

Reiner, E. S., and C. J. Radke. 1990. Variational approach to the electrostatic free energy in charged colloidal suspensions: general theory for open systems. *J. Chem. Soc. Faraday Trans.* 86:3901–3912.

Reyes, C. M., and P. A. Kollman. 2000a. Investigating the binding specificity of U1A–RNA by computational mutagenesis. *J. Mol. Biol.* 295: 1–6.

Reyes, C. M., and P. A. Kollman. 2000b. Structure and thermodynamics of RNA–protein binding: using molecular dynamics and free energy analyses to calculate the free energy of binding and conformational changes. *J. Mol. Biol.* 297:1145–1158.

Roux, B., and T. Simonson. 1999. Implicit solvent models. *Biophys. Chem.* 78:1–20.

Sham, Y., Z. Chu, and A. Warshel. 1997. Consistent calculations of pK_a's of ionizable residues in proteins: semi-microscopic and microscopic approaches. *J. Phys. Chem. B.* 101:4458–4472.

Sham, Y. Y., Z. T. Chu, H. Tao, and A. Warshel. 2000. Examining methods for calculations of binding free energies: LRA, LIE, PDLD-LRA, and PDLD/S-LRA calculations of ligands binding to an HIV protease. *Proteins.* 39:393–407.

Sham, Y., I. Muegge, and A. Warshel. 1998. The effect of protein relaxation on charge–charge interactions and dielectric constants of proteins. *Biophys. J.* 74:1744–1753.

Sharp, K. A. 1998. Calculations of HyHEL10-lysozyme binding free energy changes: effect of ten point mutations. *Proteins.* 33:39–48.

Sharp, K. A., and B. Honig. 1990. Calculating total electrostatic energies with the nonlinear Poisson–Boltzmann equation. *J. Phys. Chem.* 94: 7684–7692.

Simonson, T., G. Archontis, and M. Karplus. 1999. A Poisson–Boltzmann study of charge insertion in an enzyme active site: the effect of dielectric relaxation. *J. Phys. Chem. B.* 103:6142–6156.

Simonson, T., and C. L. Brooks. 1996. Charge screening and the dielectric constant of proteins: insights from molecular dynamics. *J. Am. Chem. Soc.* 118:8452–8458.

Simonson, T., and D. Perahia. 1995. Microscopic dielectric properties of cytochrome *c* from molecular dynamics simulations in aqueous solution. *J. Am. Chem. Soc.* 117:7987–8000.

Tang, Y., and L. Nilsson. 1999. Molecular dynamics simulations of the complex between human U1A protein and hairpin II of U1 small nuclear RNA and of free RNA in solution. *Biophys. J.* 77:1284–1305.

Tsai, C.-J., B. Ma, and R. Nussinov. 1999. Folding and binding cascades: shifts in energy landscapes. *Proc. Natl. Acad. Sci. U.S.A.* 96:9970–9972.

Warshel, A., and J. Åqvist. 1991. Electrostatic energy and macromolecular function. *Ann. Rev. Biophys. Biophys. Chem.* 20:267–298.

Warshel, A., and A. Papazyan. 1998. Electrostatic effects in macromolecules: fundamental concepts and practical modeling. *Curr. Opin. Struct. Biol.* 8:211–217.

Warshel, A., A. Papazyan, and I. Muegge. 1997. Microscopic and semimacroscopic redox calculations: what can and cannot be learned from continuum models. *J. Biol. Inorg. Chem.* 2:143–152.

Weiner, S. J., P. A. Kollman, D. T. Nguyen, and D. A. Case. 1986. An all atom forcefield for simulations of proteins and nucleic acid. *J. Comput. Chem.* 7:230–252.

Williams, D. J., and K. B. Hall. 1996. RNA hairpins with non-nucleotide spacers bind efficiently to the human U1A protein. *J. Mol. Biol.* 257: 265–275.

Zwanzig, R. W. 1954. High temperature equation of state by a perturbation method. *J. Chem. Phys.* 22:1420–1426.