



# NIH Public Access

## Author Manuscript

*Phys Chem Chem Phys.* Author manuscript; available in PMC 2009 August 28.

Published in final edited form as:

*Phys Chem Chem Phys.* 2008 August 28; 10(32): 4889–4902. doi:10.1039/b807384h.

## Assessing the performance of implicit solvation models at a nucleic acid surface

Feng Dong<sup>a</sup>, Jason A. Wagoner<sup>b</sup>, and Nathan A. Baker<sup>c</sup>

<sup>a</sup> Merck & Co., Inc., 770 Sumneytown Pike, P.O. Box 4, WP42-330, West Point, PA 19486, USA. E-mail: feng\_dong@merck.com

<sup>b</sup> Department of Chemistry, Stanford University, 333 Campus Drive #121, Mailbox 13, Stanford, CA 94305-5080, USA. E-mail: jwagoner@stanford.edu

### Abstract

Implicit solvation models are popular alternatives to explicit solvent methods due to their ability to “pre-average” solvent behavior and thus reduce the need for computationally-expensive sampling. Previously, we have demonstrated that Poisson-Boltzmann models for polar solvation and integral-based models for nonpolar solvation can reproduce explicit solvation forces in a low-charge density protein system. In the present work, we examine the ability of these continuum models to describe solvation forces at the surface of a RNA hairpin. While these models do not completely describe all of the details of solvent behavior at this highly-charged biomolecular interface, they do provide a reasonable description of average solvation forces and therefore show significant promise for developing more robust implicit descriptions of solvent around nucleic acid systems for use in biomolecular simulation and modeling. Additionally, we observe fairly good transferability in the nonpolar model parameters optimized for protein systems, suggesting its robustness for modeling general nonpolar solvation phenomena in biomolecular systems.

### 1 Introduction

Among the various components of molecular interactions, electrostatic interactions are of special importance [1–14] because of their long range and influence on polar or charged molecules, including water, aqueous ions, and amino or nucleic acids. In particular, robust models of electrostatic interactions are essential to understand the solvation properties of molecules and changes in solvation upon molecular folding, binding, and other transitions. Electrostatic interactions, therefore, are of central importance in visualizing and analyzing molecular structure and modeling the intramolecular and intermolecular interactions of biological molecules.

Solvation models can be roughly divided into two classes [2–4,6,8,12–16]: explicit solvent models that treat the solvent in molecular or atomic detail and implicit solvent models that generally replace the explicit solvent with a dielectric continuum. Each method has its strengths and weaknesses. While explicit solvent models offer some of the highest levels of detail, they generally require extensive sampling to converge thermodynamic or kinetic properties of interest. On the other hand, implicit solvent models trade detail and some accuracy to eliminate costly sampling of solvent degrees of freedom. Because of their fewer degrees of freedom,

<sup>c</sup> To whom correspondence should be addressed. Department of Biochemistry and Molecular Biophysics, Center for Computational Biology, Washington University in St. Louis, 700. S. Euclid Ave., St. Louis, MO 63110, USA. E-mail: baker@ccb.wustl.edu..

implicit solvent methods have become popular for many applications in molecular simulation [1–14].

Biomolecular interfaces provide a particularly interesting context for examining the nature of solvation and the ability of various models to accurately describe solvent-mediated interactions. Solvation is a crucial component of biomolecular simulation and structural analysis; accurate models of biomolecular solvation are essential for progress in modeling. In what follows, we review Poisson-Boltzmann implicit solvent models for polar solvation as well as recent developments in implicit solvent nonpolar models. We then apply these polar and nonpolar solvation models to an RNA hairpin system to examine the effect of this highly charged biomolecular interface on solvent properties and the performance of the implicit solvent approximation. In particular, model performance is assessed by comparing solvation mean forces with results from explicit solvent simulations. Unlike previous analyses of RNA solvation, which largely focus on global properties, this force-based analysis provides a unique atomic view of the performance of implicit solvent methods for this highly charged system.

### 1.1 Modeling molecular solvation

Solute-solvent interactions are generally described by solvation energies; that is, the free energy of transferring the solute from a vacuum to the solvent environment of interest (e.g., water at a certain ionic strength). This process is outlined in Fig. 1, adapted from Levy et al [17]. As Fig. 1 illustrates, solvation energies can be formally decomposed into polar and nonpolar parts related to electrostatic and other types of solute-solvent interactions. Roux and Simonson have provided a very useful framework [16] which we review here and use in our analyses, for understanding this decomposition in the context of implicit solvent models.

Consider a solute with  $N$  atoms in a conformation  $\underline{X} \in \mathbb{R}^{3N}$  surrounded by solvent molecules with a configuration  $\underline{Y} \in \mathbb{R}^M$ , where  $M$  is the number of solvent degrees of freedom. The solvation potential of mean force  $W(\underline{X})$  of the solute is defined by [16]

$$e^{-\beta W(\underline{X})} = \frac{\int d\underline{Y} e^{-\beta[U_{vv}(\underline{Y}) + U_{uv}(\underline{X}, \underline{Y})]}}{\int d\underline{Y} e^{-\beta U_{vv}(\underline{Y})}} \quad 1$$

where  $\beta = (k_B T)^{-1}$  is the inverse thermal energy,  $U_{vv}(\underline{Y})$  is the solvent-solvent potential, and  $U_{uv}(\underline{X}, \underline{Y})$  is the solute-solvent potential. This potential of mean force can be differentiated with respect to solute coordinates  $\underline{X}$  to give a solvation mean force:

$$\underline{F}(\underline{X}) = -\frac{\partial W(\underline{X})}{\partial \underline{X}} = -\frac{\int d\underline{Y} \frac{\partial U_{uv}(\underline{X}, \underline{Y})}{\partial \underline{X}} e^{-\beta[U_{vv}(\underline{Y}) + U_{uv}(\underline{X}, \underline{Y})]}}{\int d\underline{Y} e^{-\beta U_{vv}(\underline{Y})}} \quad 2$$

or

$$\underline{F}(\underline{X}) = -\left\langle \frac{\partial U_{uv}(\underline{X}, \underline{Y})}{\partial \underline{X}} \right\rangle_{\underline{Y}(p)}, \quad 3$$

where  $\langle \cdot \rangle_{\underline{Y}(p)}$  denotes the full (polar) ensemble average:

$$\langle f(\underline{X}, \underline{Y}) \rangle_{\underline{Y}(p)} = \frac{\int d\underline{Y} f(\underline{X}, \underline{Y}) e^{-\beta[U_{vv}(\underline{Y}) + U_{uv}(\underline{X}, \underline{Y})]}}{\int d\underline{Y} e^{-\beta U_{vv}(\underline{Y})}} \quad 4$$

It is often useful to decompose the solvent-solute interaction potential into polar  $U_{uv}^{(p)}$  and nonpolar  $U_{uv}^{(np)}$  components

$$U_{uv}(\underline{X}, \underline{Y}) = U_{uv}^{(p)}(\underline{X}, \underline{Y}) + U_{uv}^{(np)}(\underline{X}, \underline{Y}) \quad 5$$

where  $U_{uv}^{(p)}$  typically represents attractive and repulsive electrostatic interactions and  $U_{uv}^{(np)}$  includes steric repulsion and attractive solute-solvent dispersive interactions. Decomposition of the potential in this manner leads to the following polar  $W^{(p)}$  and nonpolar  $W^{(np)}$  solvation potentials of mean force [16]

$$e^{-\beta W^{(p)}(\underline{X})} = \frac{\int d\underline{Y} e^{-\beta[U_{vv}(\underline{Y}) + U_{uv}^{(p)}(\underline{X}, \underline{Y}) + U_{uv}^{(np)}(\underline{X}, \underline{Y})]}}{\int d\underline{Y} e^{-\beta[U_{vv}(\underline{Y}) + U_{uv}^{(np)}(\underline{X}, \underline{Y})]}} \quad 6$$

$$e^{-\beta W^{(np)}(\underline{X})} = \frac{\int d\underline{Y} e^{-\beta[U_{vv}(\underline{Y}) + U_{uv}^{(np)}(\underline{X}, \underline{Y})]}}{\int d\underline{Y} e^{-\beta[U_{vv}(\underline{Y})]}}. \quad 7$$

Differentiation of these potentials of mean force leads to the following polar  $F^{(p)}$  and nonpolar  $F^{(np)}$  solvation mean forces [18]:

$$\underline{F}^{(p)}(\underline{X}) = -\left\langle \frac{\partial U_{uv}^{(p)}(\underline{X}, \underline{Y})}{\partial \underline{X}} + \frac{\partial U_{uv}^{(np)}(\underline{X}, \underline{Y})}{\partial \underline{X}} \right\rangle_{Y,(p)} + \left\langle \frac{\partial U_{uv}^{(np)}(\underline{X}, \underline{Y})}{\partial \underline{X}} \right\rangle_{Y,(np)} \quad 8$$

$$\underline{F}^{(np)}(\underline{X}) = -\left\langle \frac{\partial U_{uv}^{(np)}(\underline{X}, \underline{Y})}{\partial \underline{X}} \right\rangle_{Y,(np)}. \quad 9$$

These expressions also introduce the nonpolar ensemble average defined as

$$\langle f(\underline{X}, \underline{Y}) \rangle_{Y,(np)} = \frac{\int d\underline{Y} f(\underline{X}, \underline{Y}) e^{-\beta[U_{vv}(\underline{Y}) + U_{uv}^{(np)}(\underline{X}, \underline{Y})]}}{\int d\underline{Y} e^{-\beta[U_{vv}(\underline{Y})]}}. \quad 10$$

Equations 8 and 9 are some of the primary metrics used in this study to evaluate the performance of implicit solvent methods by comparison with explicit solvent simulations.

## 1.2 Implicit solvent models

As mentioned above, implicit solvent methods have become popular alternatives to the computationally expensive explicit solvent approaches, although they have lower detail and accuracy [2–4,6,8,12–16]. In implicit solvent methods, the molecules of interest are treated explicitly while the solvent is represented by its average effect on the solute. Polar and nonpolar solvation influences are decomposed through the framework introduced above. In general biomolecular contexts, polar and nonpolar solvation terms act in opposing directions; nonpolar solvation favors compact structures with small areas and volumes, while polar solvation favors maximum solvent exposure for all polar groups in the solute.

### 1.2.1 Polar solvation models

As illustrated in Fig. 1, the polar solvation energy is generally associated with a difference in charging free energies in vacuum and in solvent. A variety of implicit solvent models are available to biomedical researchers to describe polar solvation [2–4,6,8,11–16,19–31]; however, the most widely-used methods are currently the Generalized Born [11,19–28] and Poisson-Boltzmann [2,3,6,8,13–15] models. This study focuses on the Poisson-Boltzmann method which can be formally derived from more detailed theories [32–34] and offers a somewhat slower, but often more accurate, method for evaluating polar solvation properties [20,35,36]. The Poisson-Boltzmann model is often used to evaluate thermodynamic properties [8,10,14,37–41] such as solvation energies and potentials of mean force for use in molecular calculations of binding constants,  $pK_a$ s, folding energies, and many other applications. Such energies can also be differentiated with respect to atom positions to provide mean solvation

forces (as in Sec. 1.1 and used throughout this study) for dynamics simulations [42–46] and related computational work [18,47].

The nonlinear Poisson-Boltzmann equation [2,3,6,8,13–15] is an elliptic partial differential equation

$$-\nabla \cdot \epsilon(\underline{r}) \nabla \phi(\underline{r}) = \varrho(\underline{r}) + \sum_i^m c_i q_i e^{-\beta q_i [\phi(\underline{r}) + V_i(\underline{r})]}, \quad 11$$

for the potential  $\phi(\underline{r})$  at all points  $\underline{r}$  in a domain  $\Omega \subset \mathbb{R}^3$ . Eq. 11 is usually solved subject to Dirichlet boundary conditions which use an asymptotically-correct approximation to the potential on the domain boundary [48]. In this equation,  $\epsilon(\underline{r})$  is the dielectric coefficient,  $\varrho(\underline{r})$  is the solute charge distribution,  $c_i$  and  $q_i$  are the charges and concentrations of the  $m$  mobile ion species, and  $V_i(\underline{r})$  is the repulsive steric interaction function between the ions and the solute. The solution  $\phi$  to Eq. 11 minimizes the electrostatic free energy functional [8,14,33,49,50]

$$G[\phi; \underline{X}] = \int_{\Omega} \left\{ \varrho(\underline{r}) \phi(\underline{r}) - \frac{\epsilon(\underline{r})}{2} [\nabla \phi(\underline{r})]^2 + \sum_i^m c_i e^{-\beta V_i(\underline{r})} [e^{-\beta q_i \phi(\underline{r})} - 1] \right\} d\underline{r} \quad 12$$

which, in turn, can be used to calculate electrostatic energies for the system. Differentiation of this expression with respect to solute coordinates  $\underline{X}$ , which are implicitly part of the coefficients  $\epsilon$ ,  $\varrho$ , and  $V_i$ , leads to the electrostatic forces [51,52]

$$-\frac{\partial G[\phi; \underline{X}]}{\partial \underline{X}} = - \int_{\Omega} \left\{ \frac{\partial \varrho(\underline{r})}{\partial \underline{X}} \phi(\underline{r}) - \frac{1}{2} \frac{\partial \epsilon(\underline{r})}{\partial \underline{X}} [\nabla \phi(\underline{r})]^2 - \beta \sum_i^m c_i \frac{\partial V_i(\underline{r})}{\partial \underline{X}} e^{-\beta V_i(\underline{r})} [e^{-\beta q_i \phi(\underline{r})} - 1] \right\} d\underline{r} \quad 13$$

Note that the free energy in Eq. 12 is not equivalent to  $W^{(p)}$  from Sec. 1.1 and the force in Eq. 13 is not equivalent to  $F^{(p)}$  because these quantities include both solute-solute and solute-solvent/ion electrostatic contributions. However, the solute-solute electrostatic interaction is straightforward to remove. In the absence of solvent and ions, the electrostatic potential  $\psi$  due to the solute is governed by

$$-\epsilon_s \nabla^2 \psi(\underline{r}) = \varrho(\underline{r}) \quad 14$$

that is solved subject to Coulomb potential Dirichlet boundary conditions for a homogeneous dielectric of  $\epsilon_s$ , a constant which is set to the internal dielectric constant of the solute. The potential  $\psi$  can be used to define a free energy

$$G_\psi[\psi; \underline{X}] = \int_{\Omega} \left\{ \varrho(\underline{r}) \psi(\underline{r}) - \frac{\epsilon_s}{2} [\nabla \psi(\underline{r})]^2 \right\} d\underline{r} \quad 15$$

that can be subtracted from the free energy for the inhomogeneous system (Eq. 12) so that the Poisson-Boltzmann model for the polar solvation energy  $W^{(p)}$  can be written as

$$W^{(p)}(\underline{X}) = G[\phi; \underline{X}] - G_\psi[\psi; \underline{X}] \quad 16$$

Likewise, we can write a Poisson-Boltzmann solvent model for the polar solvation force  $F^{(p)}$  as

$$F^{(p)}(\underline{X}) = - \frac{\partial G[\phi; \underline{X}]}{\partial \underline{X}} + \frac{\partial G_\psi[\phi; \underline{X}]}{\partial \underline{X}} = - \int_{\Omega} \left\{ \frac{\partial \varrho(\underline{r})}{\partial \underline{X}} [\phi(\underline{r}) - \psi(\underline{r})] - \frac{1}{2} \frac{\partial \epsilon(\underline{r})}{\partial \underline{X}} [\nabla \phi(\underline{r})]^2 - \beta \sum_i^m c_i \frac{\partial V_i(\underline{r})}{\partial \underline{X}} e^{-\beta V_i(\underline{r})} [e^{-\beta q_i \phi(\underline{r})} - 1] \right\} d\underline{r}. \quad 17$$

Equation 17 will be used for the implicit polar solvation forces used in this study.

### 1.2.2 Nonpolar solvation models

Poisson-Boltzmann methods provide *polar* solvation energies and therefore must be complemented by *nonpolar* solvation models to provide a complete view of biomolecular solvent-solute interactions. As illustrated in Fig. 1, nonpolar solvation is generally associated with the insertion of the uncharged solute into solvent. There are many such nonpolar solvation models available; however, recent work by Levy, Gallicchio, and others [17,23,53,54] as well as our own research [47] has demonstrated the importance of nonpolar implicit solvent models which include treatment of attractive solute-solvent dispersion terms (Step 5 in Fig. 1) as well as full models of solvent-solvent repulsive interactions (Step 4 in Fig. 1) that include both area and volume contributions [47].

Previously, we demonstrated that nonpolar solvation forces on a protein surface could be accurately reproduced by an implicit solvent model of the form [47]

$$W^{(np)}(\underline{X}) = \gamma A(\underline{X}) + pV(\underline{X}) + \bar{\rho} \int_{\Omega} U_{uv}^{(np.att)}(\underline{X}, \underline{r}) g_0(\underline{r}, \underline{X}) d\underline{r}, \quad 18$$

where  $\gamma$  is a solvent surface tension,  $A(\underline{X})$  is the area of the solute in conformation  $\underline{X}$ ,  $p$  is a solvent pressure,  $V(\underline{X})$  is the volume of the solute in conformation  $\underline{X}$ ,  $\bar{\rho}$  is the bulk density of the solvent,  $\Omega$  is the system domain of interest,  $U_{uv}^{(np.att)}(\underline{X}, \underline{r})$  is the attractive dispersive component of the interaction between the solute in conformation  $\underline{X}$  and the volume of solvent at point  $\underline{r}$ , and  $g_0(\underline{r}, \underline{X})$  is a model distribution function for solvent around a solute cavity in conformation  $\underline{X}$ . The first two terms of this expression represent a scaled particle theory [55–57] model of cavity creation due to the repulsive interactions between solute and solvent. The last term represents a perturbative model of the attractive nonpolar solute-solvent interactions [58]. Equation 18 can be differentiated with respect to the solute coordinates  $\underline{X}$  to provide our implicit model of nonpolar solvation mean forces:

$$\underline{F}^{(np)}(\underline{X}) = -\frac{\partial W^{(np)}(\underline{X})}{\partial \underline{X}} = -\gamma \frac{\partial A(\underline{X})}{\partial \underline{X}} - p \frac{\partial V(\underline{X})}{\partial \underline{X}} - \bar{\rho} \int_{\Omega} \frac{\partial U_{uv}^{(np.att)}(\underline{X}, \underline{r})}{\partial \underline{X}} g_0(\underline{r}, \underline{X}) d\underline{r} - \underbrace{\bar{\rho} \int_{\Omega} U_{uv}^{(np.att)}(\underline{X}, \underline{r}) \frac{\partial g_0(\underline{r}, \underline{X})}{\partial \underline{X}} d\underline{r}}_{\approx 0} \quad 19$$

The second integral term in the force expression is typically removed based on physical arguments (e.g., its contribution is negligible in the absence of phase transitions) or based on the observation that it results in a surface-area based force that would be accounted for in parameterization of the repulsive components [47]. Equation 19 will be used for the implicit nonpolar solvation forces used in this study. The implementation details for Eqs. 18 and 19, including the functional forms for  $A$ ,  $V$ , and  $g_0$ , have been described in detail previously [47]. In our implementation, a probe radius  $\sigma_s$  is used to define a very simple step-function behavior for  $g_0$ :  $g_0$  is zero within the solvent-inaccessible (solute) regions of the problem domain  $\Omega$  and 1 everywhere else. This region of solvent-inaccessibility is also used to define the solvent-accessible surface area  $A(\underline{X})$  and volume  $V(\underline{X})$  used in this model. Therefore, the probe radius  $\sigma_s$  plays a central role in the implementation and parameterization of this implicit nonpolar solvation model.

### 1.3 Testing implicit solvent models

**1.3.1 Implicit solvent model shortcomings**—Poisson-Boltzmann theory is approximate and, as a result, has several well-known limitations which can affect its accuracy [13,14,33,34]. These limitations have been reviewed in previous articles and will only be briefly summarized here. First, most continuum models assume linear and local solvent response [4,31,32,59]. However, nonlinear solvent response (usually through dielectric saturation or electrostriction), can be important in regions of strong electric field [4,31,59]. Biologically-relevant examples of nonlinear solvent response have been found near highly charged ions,

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

biomolecules, and other interfaces. Nonlocal solvent response generally involves the finite nonzero size of water and its unique hydrogen bonding with solute and other solvent molecules. Such nonlocal response can be important in describing the orientation of water at biomolecular interfaces [60], differing solvation of cations and anions [61], and the solvation of asymmetric charge distributions [62]. The second major limitation is the mean-field treatment of ions in Poisson-Boltzmann theory [14,33,34]. Mean field models assume that each ion experiences only the average influence of the other ions in solution. Such averaging precludes detailed ion-ion interactions involving steric repulsion of ions (or their solvation shells) and Coulombic interaction of ions, including repulsion and attractive pairing. The mean field assumption thereby eliminates correlations and fluctuations which can have important energetic and structural consequences for solutions of divalent and multivalent ions surrounding highly-charged molecules such as nucleic acids [63–67]. As suggested by the limitations above, Poisson-Boltzmann models also neglect detailed ion-solvent interactions which eliminate differences between ion species in solution and thereby preclude analysis of specific ion species effects [68–71]. However, despite these limitations, Poisson-Boltzmann methods are still very important for biomolecular structural analysis, modeling, and simulation. Furthermore, these limitations are currently being addressed through new implicit solvent models [33,63,64,72,73] and hybrid treatments [74–78] that extend the applicability of Poisson-Boltzmann theory while preserving some of its computational efficiency through pre-averaging solvent and ion response.

Implicit models for nonpolar solvation also have limitations. In general, nonpolar solvation effects can be decomposed into repulsive and attractive interactions. The repulsive contribution, or the entropic penalty associated with cavity formation, is often approximated using scaled particle theory. Scaled particle theory uses macroscopic arguments to model this cavity formation component based on global properties of surface area, volume, and curvature for simple geometries [55,57]. Unfortunately, this theory does not include the structural effects that must be present for macromolecules of complex geometries. Though the crossover between volume and surface-area dependence for cavity creation has been studied for spherical cavities [79–81], there is still no robust model that clearly predicts how such terms can be used in tandem to describe structures containing features at a variety of length scales.

The attractive component of nonpolar solvation, calculated via the integral component in Eq. 18, has been shown to produce very accurate results with respect to explicit solvent calculations [17,23,47,53,54]. However, the distribution function  $g_0$  is generally implemented as a step function between the solute interior and the solvent exterior [17,47]. This step-like model is a very rough approximation to the true distribution function that should be used to account for these interactions. The use of this approximate distribution function necessitates the parameterization of the solvent probe radius to determine molecular surfaces and volumes defined by  $g_0$  [47,54]. Furthermore, this approximate form of  $g_0$  can lead to significant error, particularly when dramatic deviations from bulk solvent density are encountered near the molecular surface.

Finally, recent work by Dzubiella and others [82–84] has suggested the need for theories which couple the polar and nonpolar components of implicit solvent models. Such theories are particularly interesting for highly charged solutes such as the RNA under consideration in this work. However, the current study is limited to the more widely-used implicit solvent models described above; future work will investigate the coupled models of Dzubiella et al. in the context of nucleic acid solvation.

**1.3.2 Past work**—There have been a variety of efforts to assess the accuracy of implicit solvent models in biomolecular systems. Such work often analyzes solvation free energies and related quantities [37–41,44,85], conformational and structural properties [35,42,86], or other

global metrics [60] to assess the performance and accuracy of the implicit solvent model. Our recent work [18,47] has focused on the examination of local solvation properties to help identify the sources of error in implicit solvent models. In particular, we have compared mean polar and nonpolar solvation forces ( $F^{(p)}$  and  $F^{(np)}$  defined above) between implicit and explicit simulations of solvated proteins [18,47]. Such comparisons provide a detailed view of solvation model performance at specific regions of the biomolecule-solvent interface. These types of analyses rely on the comparison of the model with simulations (rather than experiment). However, they are nonetheless useful tests of the importance of discreteness in solvation phenomena. In particular, these analyses allow us to assess implicit solvent performance in the context of the limitations discussed above (see Sec. 1.3.1).

**1.3.3 RNA solvation**—In the current work, we move from our past studies of protein solvation to a much more challenging biomolecule-solvent interface: nucleic acids. In particular, we will examine the solvation properties of the Iron Responsive Element (IRE) RNA hairpin, previously studied by Showalter et al. [87] for the analysis of its dynamic properties in the context of NMR measurements. Computational investigation of nucleic acid solvation has been performed by several other groups [88] who have examined structure and dynamics of solvent and ions around DNA [66,89–101] and RNA [93,94,102–106] as well as the role of water in nucleic acid-protein interactions [86,107–109]. While comparisons between solvation models have been made previously, most of these comparisons have been global in nature and have focused on ion-nucleic acid interactions. The current work is the first investigation of the ability of the Poisson-Boltzmann polar solvation model and a new nonpolar solvation model to accurately reproduce mean solvation *forces* at RNA-water interfaces. In the spirit of this themed issue of water at interfaces, this work focuses exclusively on the well-converged behavior of water around the IRE RNA hairpin and defers a similar study of ionic forces to future work.

## 2 Results

### 2.1 Nonpolar solvation forces

We first examined the ability of the implicit solvent model (Eq. 19) to reproduce nonpolar solvation forces obtained from the explicit solvent simulations (Eq. 9). In order to illustrate the importance of each component of the nonpolar implicit solvent model, we present results for three model variants that each include attractive dispersive forces (integral in term Eqs. 18 and 19) together with versions of the repulsive component that include: only the  $\gamma A(x)$  term (“SASA only”), only the  $pV(x)$  term (“SAV only”), or both terms (“SASA and SAV”). Table 2 illustrates the accuracy of these three variants of the nonpolar implicit solvent models using values of  $\gamma$ ,  $p$ , and  $\sigma_s$  that are optimized against the explicit solvent forces, as described in the Methods section. Confidence intervals for these parameters were assigned using a standard F-test [110] at the 99% confidence interval. Figure 3 highlights the mean-squared errors obtained for the full model (“SASA and SAV”) over a range of probe radii, thus illustrating the fitness landscape for our parameter optimization process. To contrast optimized parameters with chosen values, Table 3 presents results for the three variants of the nonpolar implicit solvent model using the popular choice of  $\sigma_s = 1.4 \text{ \AA}$ , as well as optimal parameters determined from a previous study [47] for a low charge density protein (intestinal fatty acid binding protein or IFABP). Finally, Fig. 4 illustrates the per-component comparison of implicit and explicit model nonpolar solvation forces summarized in Tables 2 and 3. Figure 5 gives structural context to this comparison by mapping onto the IRE structure the distribution of differences in nonpolar solvation force magnitudes between the implicit and explicit solvent models.

## 2.2 Polar solvation forces

We next compared polar solvation forces from the Poisson-Boltzmann model (Eq. 17) to those obtained from the explicit solvent simulation (Eq. 8). Figure 6 summarizes the per-component comparison of implicit and explicit model polar solvation forces on the IRE structure. In particular, Fig. 6 correlates 6132 explicit and implicit model results: data for each  $x$ ,  $y$ , and  $z$  polar solvation force component on each of the 511 atoms for each of the 4 structural representatives. Implicit and explicit model polar solvation forces were correlated with a Pearson coefficient of  $R = 0.800$ , a regression coefficient (slope) of  $r = 0.997$ , and a total mean squared error of  $10.76 \text{ kcal}^2 \text{ mol}^{-2} \text{ \AA}^{-2}$ . Figures 7A and B give structural context to this comparison by illustrating the distribution of polar solvation forces on a sample IRE conformation. Figure 7C highlights the structural regions of largest difference between the implicit and explicit polar solvation forces.

## 2.3 Solvent polarization

In order to better understand the nature of polar solvation at the nucleic acid-water interface, we also analyzed solvent polarization around the RNA hairpin. Figure 8 provides a comparison of solvent polarization between the implicit and explicit models. This figure correlates 109038 data points: polarization  $x$ ,  $y$ , and  $z$  components for regions of the simulation domain with non-zero solvent density (see Methods) for all 4 IRE structural representations. Implicit and explicit model polarizations were correlated with a Pearson coefficient of  $R = 0.840$ , a regression coefficient of  $r = 0.891$ , and a total mean squared error of  $1.78 \times 10^{-7} \text{ e}^2 \text{ \AA}^{-4}$ . Figure 9 gives structural context to this comparison by illustrating the distribution of polarization and polarization error around a sample IRE conformation. Note that this figure only shows the results for a single conformation (II) for illustration purposes; conformational fluctuations make averaging over several conformations difficult to represent on a single structure.

# 3 Discussion

## 3.1 Nonpolar solvation model performance

As is clearly demonstrated in Tables 2 and 3, the inclusion of an appropriate cavity model is essential for the accurate representation of nonpolar solvation effects. Though commonly neglected, the inclusion of the  $pV(X)$  term provides more accurate forces than are obtained with the  $\gamma A(X)$  term. This result is in accordance with studies that have demonstrated the volume dependence of cavity formation for sub-nanometer length scales [79–81], likely the relevant length scale for *atomic* solvation forces. Though the inclusion of the  $\gamma A(x)$  term into the full model (“SASA and SAV”) has little impact on these results, it is expected that a full model of nonpolar solvation will incorporate both area and volume terms in order to capture the appropriate contributions to cavity formation at a variety of length scales.

With the inclusion of the  $pV(X)$  term, nonpolar forces obtained from the implicit solvent model with optimized parameters compare very favorably to those from the explicit solvent simulations (see Fig. 4). However, the optimal probe radius for the full model (0.93 Å) is significantly smaller than a physically-realistic water probe radius of approximately 1.4 Å. As discussed previously, such an unrealistic value for the probe radius originates from the approximate form of the distribution function,  $g_0$ , used in equations 18 and 19. Nonetheless, the results presented in Table 3 demonstrate that the nonpolar implicit solvent model presented in this work is also capable of providing accurate forces using this more popular (and meaningful) choice of  $\sigma_s = 1.4 \text{ \AA}$ .

Figure 5 reveals interesting structural origins for the slight deviations of this nonpolar model with respect to the explicit solvent results. For the small-but-optimal 0.93 Å probe radius (Fig. 5A), the errors are very strongly localized to hydrogen H5 of the pyrimidine bases in the

molecule, with a small amount of error also present on highly exposed sugar carbons (C5'). As the probe radius increases to the more realistic 1.40 Å value (Fig. 5B), the error becomes more broadly distributed across the molecule with somewhat less prominent error on the pyrimidine H5 atoms and more error distributed over the sugar backbone. Similar results are seen with the nonpolar protein-derived IFABP parameters (Fig. 5C). The localization of error to specific atom types can be explained as a result of approximating the *true* form of  $g_0$  using a globally parameterized probe radius, as discussed in Sec. 1.3.1. We are currently exploring ways in which to improve  $g_0$  definitions to relax this global parameterization constraint by incorporating the heterogeneity of surface atom types and thereby improving these types of errors.

The optimal parameters obtained for the nonpolar implicit solvent model for solvation forces on IRE deviate from similar calculations on protein conformations in a previous study [47]. For the purposes of assessing parameter transferability, Table 3 provides the results of calculating nonpolar forces on IRE using these previously obtained parameters. With these transferred parameters, the results are still very accurate when compared to models that contain only SASA terms for the repulsive contribution. Though the values of  $p$  and  $\sigma_s$  obtained in this study do not agree within statistical error to those obtained from protein structures [47], it should be noted that the value for  $\gamma$  is the same for both parameter sets and that the relative magnitude of the  $pV(x)$  term is very similar. It can be expected that appropriate parameterization of the  $\sigma_s$  and  $p$  parameters over a range of structures in a variety of chemical environments will yield a set of transferable parameters for this implicit solvent model. We leave this, as well as the appropriate weighting of volume and area terms for different length scales, to a future study.

### 3.2 Polar solvation model performance

The agreement between implicit and explicit polar solvation forces, as illustrated in Fig. 6, is reasonable. The two datasets correlate with an excellent regression coefficient of  $r = 0.997$  but with significant scatter (Pearson coefficient of  $R = 0.800$ ), contributing to the observed root mean squared error of 3.28 kcal mol<sup>-1</sup> Å<sup>-1</sup>. Given the limitations of polar implicit solvent models discussed above (Sec. 1.3.1), particularly at highly charged interfaces, this level of error is not particularly surprising. In fact, it is encouraging that this correlation is comparable to that of previous studies of the low-charge-density IFABP protein [18]. Furthermore, the error between implicit and explicit models is only moderately larger than the IFABP polar solvation force errors [111] (around 1 kcal mol<sup>-1</sup> Å<sup>-1</sup>). It is important to note that both the previous study of IFABP [18] and the current study of IRE were performed using *unoptimized* radii for the implicit solvent calculations. Therefore, it is likely that optimization of radii to reproduce solvation forces and energies will help further reduce the differences between the explicit and implicit models [111].

The maps of solvation forces and solvation force errors (Fig. 7) on the IRE structure provide additional information about the ability of Poisson-Boltzmann models to describe polar solvation forces at nucleic acid-water interfaces. Qualitative comparison of the explicit (Fig. 7A) and implicit (Fig. 7B) force distributions on the IRE surface reveals some common themes. First, as expected, the highly charged phosphate backbone is among the most strongly solvated regions of the molecule for both implicit and explicit models. Additionally, amines of exposed bases are also strongly solvated in both models. However, the specific distribution of polar solvation forces differs somewhat between the explicit and implicit models. In particular, relative polar solvation forces on phosphate groups are significantly weaker in the implicit model. Given their high charge density, this error is not surprising: it is expected that phosphates would provide the most problematic region of the molecule for an implicit solvent description. Conversely, as illustrated in Figs. 7B and C, the implicit model places a greater emphasis on

solvation of polar atoms in the bases, particularly the highly exposed C11 at the “top” of the molecule and the base-paired G4 at the “bottom” of the hairpin.

The grouping of differences between the implicit and explicit models onto specific atom types emphasizes the importance of our current work in the development of parameter sets for implicit solvent simulations of nucleic acids. In particular, these results suggest that better agreement between implicit and explicit solvation forces might be obtained by decreasing the implicit solvent phosphate group radii (thus increasing the solvation force) and increasing the radii of base amines (thus decreasing the solvation force). However, parameter changes are generally very strongly correlated and such parameterization must be carried out in a systematic way to avoid the introduction of bias and additional sources of error.

The agreement between implicit and explicit model solvent polarization, as illustrated in Fig. 8, is strong with a good regression coefficient ( $r = 0.910$ ) and correlation (Pearson  $R = 0.846$ ), contributing to the observed small root mean squared error of  $4.22 \times 10^{-4} \text{ e } \text{\AA}^{-2}$ . This favorable error arises from the good agreement between the explicit and implicit models for the majority of small polarization values in the correlation. In fact, most of the apparent scatter originates in larger polarization values with absolute values above  $5 \times 10^{-3} \text{ e } \text{\AA}^{-3}$ . Given the limitations of polar implicit solvent models (Sec. 1.3.1), this relatively small level of error is surprising and encouraging. Furthermore, it is expected that the regions of largest polarization would be the most poorly described by the linear response implicit model framework.

Figure 9 helps to identify the sources of polarization differences between the models by showing implicit/explicit solvent polarization and polarization differences in the context of the IRE structure. Note that all of these figures illustrate *relative* polarization or polarization differences to help readers identify important features in these quantities that would be obscured if an absolute polarization scale was used. In all models, polarization density is very strongly peaked near charged groups. Therefore, Fig. 9 uses relatively low isocontour values (30% and 40% of the maximum polarization magnitude) to illustrate large-scale behavior in the system rather than simply identifying strong polarization peaks in the solvent. From Fig. 9A, it is immediately evident that strong polarization is distributed across large portions of the molecular surface for the explicit solvent model. Not surprisingly, the largest regions of polarization density in the explicit solvent model are along the phosphate backbone and in the groove at the “back” face of the IRE molecule (right image in Fig. 9A). On the other hand, Fig. 9B indicates that the implicit solvent model tends to localize strong polarization in specific regions of high charge density. In particular, implicit solvent model polarization appears to be most sensitive to the local geometry and packing of polar groups: density is highest in the concave “front” of the IRE molecule (left image in Fig. 9B), along localized regions of the phosphate backbone, and near the “top” of the IRE where there is a locally-higher region of charge density due to the proximity of polar basic groups and phosphate backbone.

Differences between the implicit and explicit model polarization (Fig. 9C) clearly reiterate the greater locality of polarization in the implicit solvent results. However, as discussed in Sec. 1.3.1, this locality is a fundamental feature of most continuum models which assume a local relationship between electrostatic field and induced polarization. The explicit solvent model provides a much more delocalized response due to the finite size of solvent and the ability of water to communicate changes in polarization to other nearby water molecules through hydrogen bonds. Traditional continuum models of electrostatic interactions and polar solvation are fundamentally unable to account for such nonlocal response. Instead, applications which require an accurate description of such behavior will likely need to use explicit solvent simulations, hybrid models [75–78,112], integral equation formulations [32,112–114], or new nonlocal models of continuum electrostatics [115,116]. Unfortunately, it is currently unclear

how strongly the polarization nonlocality affects solvation force error. While the errors in solvation forces (Fig. 7C) and polarization (Fig. 9C) are clearly localized to similar regions of the IRE structure, these differences between the implicit and explicit models are likely due to *both* nonlocal response and the poor implicit solvent parameters discussed above. Therefore, a more thorough investigation of the importance of nonlocal solvent response on solvation force must wait until optimal implicit solvent parameters have been developed.

### 3.3 Caveats of comparisons

The current study is subject to a few caveats which warrant additional discussion.

First, we are comparing implicit solvent *models* to nonpolarizable force field *simulations*. The justification for this comparison has been discussed earlier in this manuscript; this approach allows us to address the impact of model granularity on observables such a solvation force. However, current explicit solvent simulations have many well-known artifacts, some of which are exacerbated in nonpolarizable force fields [38,117–121]. In the context of the current work, we view explicit solvent force field polarizability as an additional variable when assessing the performance of implicit solvent models. While the current study has focused on the importance of discreteness with nonpolarizable force fields, future work will examine the influence of polarizability on our results using recent Poisson-Boltzmann models which include atomic polarizability [38,122].

Second, the current analysis focused exclusively on water contributions to nonpolar solvation, total polar solvation forces, and solvent polarization without a detailed analysis of ionic forces or behavior. However, ions clearly play an important role in nucleic acid electrostatic properties and warrant extensive study. Our current neglect of detailed ion analysis was motivated, in part, by the thematic nature of this special issue. Additionally, the proper sampling of ion distributions suitable for comparison with Poisson-Boltzmann ion densities is extremely computationally-demanding: adequate sampling is complicated by long ion relaxation times [123–125] and severe finite size effects related to simulation domain sizes (Rohit Pappu and Alan Chen, personal communication). Therefore, a proper comparison of ion effects in implicit and explicit solvation models will require a study of equal or greater length than the current manuscript and is best left to future work.

## 4 Methods

### 4.1 Solute conformations

The solvation force analyses described in this paper were carried out for four conformations of the Iron Responsive Element (IRE) RNA hairpin (Fig. 2). These four conformations were obtained from a 12 ns molecular dynamics (MD) trajectory of the IRE hairpin, described previously [87]. The last 10 ns of the 12 ns IRE trajectory was classified into six conformational clusters using the “GROMOS clustering method” as implemented in the g\_cluster tool of GROMACS 3.2.1 [126]. A 1.9 Å all-atom cutoff was used to define the clusters. The first four clusters cover 98.4% of the IRE conformations in the studied trajectory and have pairwise all-atom root-mean-squared conformational distances ranging from 1.20 to 5.27 Å, as summarized in Table 1. The median conformation of each cluster was used in all subsequent analyses.

### 4.2 Calculation of explicit solvation forces

As suggested by the theory (see Sec. 1.1), we used two different types of explicit solvent molecular dynamics simulations to generate  $\underline{F}^{(p)}(\underline{X})$  and  $\underline{F}^{(np)}(\underline{X})$  for our four IRE conformations ( $\underline{X}$ ): simulations with a full (polar) ensemble which contains the full set of molecular mechanics interactions and simulations with a nonpolar ensemble which eliminates all charge-charge interactions between the solute and its environment. For both types of

simulation, each of the four static IRE conformations was solvated in a cubic box (67 Å sides) with 9050 TIP3P [127] water molecules as well as 45 Na<sup>+</sup>, 32 Cl<sup>-</sup>, and 1 Mg<sup>++</sup> ions. As discussed in Showalter et al [87], this simulation setup was determined by grand canonical Monte Carlo simulations [74] and corresponds to bulk concentrations of 50 mM NaCl and 1 mM MgCl<sub>2</sub> and provides an electroneutral system with effective local concentrations of 250 mM Na<sup>+</sup>, 180 mM Cl<sup>-</sup>, and 5 mM Mg<sup>++</sup>. Both types of simulations were performed with a “belly-restrained” IRE solute as implemented in AMBER 8 [128] and used previously in similar protein solvation force calculations [18,47] and electrostatic potential analyses [60]. Both the full-charge and uncharged solute simulations used the AMBER 94 force field [129]; the full simulations included all solute-solvent interactions while the nonpolar simulation set all solute charges to zero. The simulations were run for 55 ns with 2 fs time steps, SHAKE constraints on hydrogen-heavy atom bonds, [130] and isobaric-isothermal conditions (NPT) at 300K temperature and 1 atm pressure using the Berendsen weak coupling scheme [131].

To generate full-charge solute (polar) ensemble averages  $\langle \cdot \rangle_{Y,(p)}$ , we performed the molecular dynamics simulations with all AMBER94 force field solute-solvent interactions present. To generate the uncharged solute (nonpolar) ensemble, we performed the simulations with only nonpolar AMBER94 solute-solvent interactions. This was accomplished by setting all charges on the restrained RNA solute to zero.

Based on convergence of the solvation forces, the last 50 ns of each trajectory was used to calculate explicit solvent forces. Snapshots were taken every 4 ps, resulting in 12500 conformations for each trajectory used in the analyses. Explicit solvent forces were evaluated using the TINKER software [132] with the AMBER94 force field [129]. For every snapshot, forces were calculated on each IRE atom in the presence and absence of solvent and ions. Mean solvation forces were then calculated by subtracting the “unsolvated” forces from the “solvated” forces and averaging over the trajectory. As described above (Eqs. 8 and 9), the nonpolar mean solvation forces  $F^{(np)}$  are obtained directly from the uncharged solute trajectory with only nonpolar solute-solvent interactions; e.g., zero solute charge. The polar mean solvation forces  $F^{(p)}$  are calculated from the difference of the mean solvation forces obtained from the full-charge and uncharged solute trajectories.

#### 4.3 Calculation of implicit solvent forces

APBS 0.5.1 implements the polar solvation force (Eq. 17) using a generalized version of the scheme suggested by Im et al. [52] but with the straightforward inclusion of nonlinear Poisson-Boltzmann equation ionic forces. We used continuous cubic B-spline boundaries for all dielectric and ion-accessibility interfaces and cubic B-spline discretization of the charge distributions. Poisson-Boltzmann calculations were performed at 300 K with a dielectric constant of 78.54 for water<sup>\*</sup> and a dielectric constant of 1.00 for the protein interior<sup>†</sup>. The ion concentrations were set to the model concentrations used for the explicit solvent simulations: 250 mM Na<sup>+</sup>, 180 mM Cl<sup>-</sup>, and 5 mM Mg<sup>++</sup>. All ions were modeled with the “typical” radius of 2.0 Å<sup>‡</sup>. Multiple-charge Debye-Hückel Dirichlet boundary conditions were used to solve the nonlinear Poisson-Boltzmann equation (Eq. 11) on a grid with the same dimensions as the explicit solvent simulation (a cubic box of 67 Å length). The Poisson-Boltzmann equation was solved using a multigrid finite difference approach on a grid with 225 × 225 × 225 unknowns at a resolution of 0.3 Å. The Supplementary Data provided with this paper explores the sensitivity of solvent polarization and polar solvation force to grid spacing. Both quantities are

<sup>\*</sup>A dielectric constant of 78.54 was chosen because it is a standard setting for continuum solvent calculations. This choice was made in spite of the known differences of TIP3P dielectric constants from this value [133–135].

<sup>†</sup>An interior dielectric constant of 1.0 was chosen to mimic the nonpolarizable interior of the protein in the fixed charge AMBER94 force field.

<sup>‡</sup>Sensitivity analysis showed very little dependence of the reported results to this choice of radius.

converged at the 0.3 Å spacing used in this study. IRE atom charges were assigned from AMBER94 force field [129]. Three radii sets were initially tested for these polar solvation calculations: AMBER94 radii [129], PARSE radii [37,40], and a custom set of radii adapted from OPLS [136] and Bondi [137]<sup>§</sup>. The Bondi radii gave the best correlation between implicit and explicit polar solvation forces and are used for all results reported here.

Nonpolar forces were also calculated using APBS. The dispersive component of these forces were integrated numerically on a grid with 0.5 Å spacing, assuming the bulk density of water  $\bar{\rho} = 0.033428 \text{ Å}^{-3}$ . Solvent accessible areas and volumes were calculated using the Shrike-Rupley method [138]; volume derivatives were calculated directly using the surface quadrature, while area derivatives were obtained by central differences of the area with respect to atomic displacements. The volume components were calculated with a grid density of 100 points Å<sup>-3</sup>, while the area components were calculated with a sphere density of 1500 points Å<sup>-2</sup>. The fine surface resolution for the area components is necessary to obtain convergence in the central-differencing used for these nonpolar force calculations. We used the AMBER94 force field [129] for nonpolar radii and the integral term Lennard-Jones well depth.

The optimal values of  $\gamma$ ,  $p$ , and  $\sigma_s$  were obtained by simultaneously optimizing all three parameters to minimize the mean squared error, defined as

$$\chi^2 = \frac{1}{12N} \sum_{c=1}^4 \sum_{i=1}^N \| \underline{f}_{i,c}^{(\text{np})} - \underline{F}_{i,c}^{(\text{np})} \|^2 \quad 20$$

where the  $c$  index denotes each of the 4 conformers, the  $i$  index denotes each of the  $N$  atoms,  $\underline{f}_{i,c}^{(\text{np})}$  is the nonpolar implicit solvent force for atom  $i$  of conformer  $c$  and  $\underline{F}_{i,c}^{(\text{np})}$  is the nonpolar explicit solvent force for atom  $i$  of conformer  $c$ .

#### 4.4 Water polarization

Solvent polarization was compared between the implicit and explicit solvent models. For the explicit model, the simulation (cubic with 67 Å sides) was discretized into a set of 21×21×21 cubic cells of 3.2 Å length. For each snapshot of the simulation, water molecules were assigned to these cells based on the location of their oxygen atom. The total dipole moment of each cell was calculated from the assigned water molecules and averaged over the last 50 ns of the explicit solvent trajectory to obtain the dipole density or polarization  $P(r)$  of each cell (in units of  $e \text{ Å}^{-2}$ ). To correct for solvent polarization around the uncharged biomolecule [60], the resulting polarization densities for the charged and uncharged solute were subtracted to obtain the final explicit solvent polarization used in these analyses.

Implicit model polarization was also calculated for each cell using the constitutive relation for continuum electrostatics:

$$\underline{P}(r) = -[\epsilon(r) - 1] \nabla \phi(r) \quad 21$$

where  $\phi$  is the solution to the nonlinear Poisson-Boltzmann equation obtained from APBS.

### 5 Conclusion

We have examined the behavior of implicit and explicit solvation models at highly charged nucleic acid-water interfaces. This study, in accordance with previous work [47], demonstrates that the inclusion of both dispersive and solvent-accessible components provides a robust model for modeling RNA hairpin nonpolar solvation forces. Analysis of polar solvation forces

---

<sup>§</sup>In particular, the following radii were assigned: carbon (1.9 Å), hydrogen (1.2 Å), nitrogen (1.625 Å), oxygen (1.775 Å).

showed that Poisson-Boltzmann models could reproduce trends in polar solvation across the molecular surface but with a significant amount of scatter. This scatter was largely due to a few groups of atom types in the IRE hairpin structure which likely require reparameterization to improve agreement between the models. Additionally, analysis of polarization showed overall agreement between the implicit and explicit solvent models. However, explicit solvent polarization demonstrated a substantially greater degree of nonlocality than the Poisson-Boltzmann based polarization.

As discussed above, the current study lays the groundwork for several new and ongoing investigations. Nonpolar model performance and transferability can likely be improved by modifications to better describe the geometry of the solute-solvent interface. Polar model performance for nucleic acid systems could be improved in a number of ways, including optimized implicit solvent parameters, better dielectric surface descriptions, and continuum models which provide more realistic descriptions of nonlocal solvent response. However, the results of this initial study are encouraging and suggest an important role for continuum solvent models in future simulations of nucleic acids.

## Acknowledgments

The authors would like to thank Rohit Pappu, Ron Levy, and Emilio Gallicchio for stimulating discussions on solvation forces and implicit solvent theory, Hui Lu and Nitin Bhardwaj for discussions on parameter optimization and their reading of the manuscript, Kathleen Hall for discussions on the IRE system, and Rachel Rice her help with manuscript editing. This work was supported in part by R01 GM069702 and an Alfred P. Sloan Research Fellowship (NAB).

## References

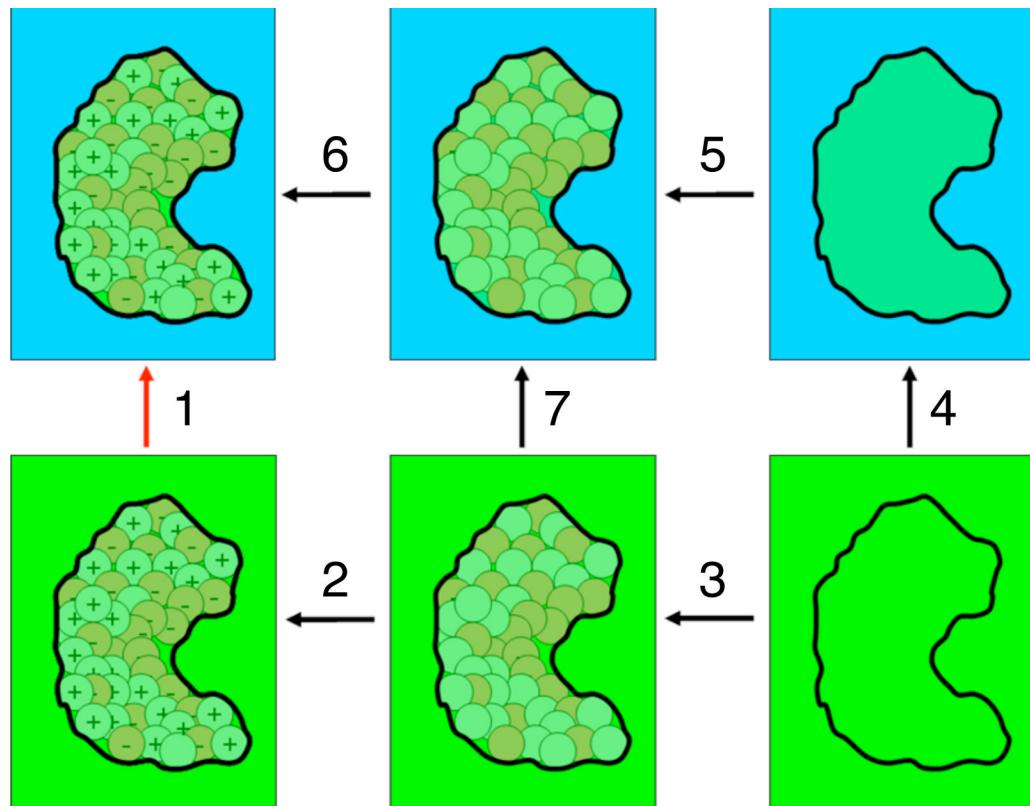
1. Honig BH, Hubbell WL, Flewelling RF. Annual Review of Biophysics and Biophysical Chemistry 1986;15:163–93.
2. Davis ME, McCammon JA. Chemical Reviews 1990;94:7684–7692.
3. Sharp KA, Honig B. Annual Review of Biophysics and Biophysical Chemistry 1990;19:301–332.
4. Warshel A, Papazyan A. Current Opinion in Structural Biology 1998;8(2):211–7. [PubMed: 9631295]
5. Simonson T. Current Opinion in Structural Biology 2001;11(2):243–252. [PubMed: 11297935]
6. Fogolari F, Brigo A, Molinari H. Journal of Molecular Recognition 2002;15(6):377–92. [PubMed: 12501158]
7. Simonson T. Reports on Progress in Physics 2003;66(5):737–87.
8. Baker NA. Methods in Enzymology 2004;383:94–118. [PubMed: 15063648]
9. Feig M, Brooks I, L. C. Current Opinion in Structural Biology 2004;14(2):217–224. [PubMed: 15093837]
10. Baker, NA. Reviews in Computational Chemistry. Lipkowitz, KB.; Larter, R.; Cundari, TR., editors. 21. John Wiley and Sons; Hoboken, NJ: 2005.
11. Koehl P. Current Opinion in Structural Biology 2006;16(2):142–51. [PubMed: 16540310]
12. Warshel A, Sharma PK, Kato M, Parson WW. Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics 2006;1764(11):1647–76.
13. Dong F, Olsen B, Baker NA. Methods in Cell Biology 2008;84:843–70. [PubMed: 17964951]
14. Grochowski P, Trylska J. Biopolymers 2007;89(2):93–113. [PubMed: 17969016]
15. Lamm, G. Reviews in Computational Chemistry. Lipkowitz, KB.; Larter, R.; Cundari, TR., editors. John Wiley and Sons, Inc.; Hoboken, N.J.: 2003. p. 147–366.
16. Roux B, Simonson T. Biophysical Chemistry 1999;78(12):1–20. [PubMed: 17030302]
17. Levy RM, Zhang LY, Gallicchio E, Felts AK. Journal of the American Chemical Society 2003;125 (31):9523–9530. [PubMed: 12889983]
18. Wagoner J, Baker NA. Journal of Computational Chemistry 2004;25(13):1623–9. [PubMed: 15264256]
19. Dominy BN, Brooks CL III. Journal of Physical Chemistry B 1999;103(18):3765–3773.

20. Bashford D, Case DA. Annual Review of Physical Chemistry 2000;51:129–152.
21. Tsui V, Case DA. Biopolymers 2000;56(4):275–291. [PubMed: 11754341]
22. Onufriev A, Case DA, Bashford D. Journal of Computational Chemistry 2002;23(14):1297–304. [PubMed: 12214312]
23. Gallicchio E, Zhang LY, Levy RM. Journal of Computational Chemistry 2002;23(5):517–29. [PubMed: 11948578]
24. Zhu J, Alexov E, Honig B. Journal of Physical Chemistry B 2005;109(7):3008–22.
25. Tjong H, Zhou HX. Journal of Chemical Physics 2007;126:195102. [PubMed: 17523838]
26. Mongan J, Simmerling C, McCammon JA, Case DA, Onufriev A. Journal of Chemical Theory and Computation 2007;3(1):159–69.
27. Chen J, Brooks, Khandogin J. Current Opinion in Structural Biology 2008;18(2):140–8. [PubMed: 18304802]
28. Grant JA, Pickup BT, Sykes MT, Kitchen CA, Nicholls A. Physical Chemistry Chemical Physics 2007;9:4913–22. [PubMed: 17912422]
29. Schaefer M, Karplus M. Journal of Physical Chemistry 1996;100(5):1578–1599.
30. Vitalis A, Pappu RV. Journal of Computational Chemistry. accepted
31. Sham YY, Muegge I, Warshel A. Biophysical Journal 1998;74(4):1744–1753. [PubMed: 9545037]
32. Beglov D, Roux B. Journal of Chemical Physics 1996;104(21):8678–8689.
33. Netz RR, Orland H. European Physical Journal E 2000;1(23):203–14.
34. Holm, C.; Kekicheff, P.; Podgornik, R. Electrostatic effects in soft matter and biophysics; NATO Science Series. Kluwer Academic Publishers; Boston: 2001.
35. David L, Luo R, Gilson MK. Journal of Computational Chemistry 2000;21(4):295–309.
36. Onufriev A, Bashford D, Case DA. Journal of Physical Chemistry B 2000;104(15):3712–3720.
37. Tang CL, Alexov E, Pyle AM, Honig B. Journal of Molecular Biology 2007;366(5):1475–96. [PubMed: 17223134]
38. MacDermaid CM, Kaminski GA. Journal of Physical Chemistry B 2007;111(30):9036–44.
39. Tsui V, Case DA. Journal of Physical Chemistry B 2001;105(45):11314–11325.
40. Sitkoff D, Sharp KA, Honig B. Journal of Physical Chemistry 1994;98(7):1978–1988.
41. Massova I, Kollman PA. Journal of the American Chemical Society 1999;121(36):8133–43.
42. Tan C, Yang L, Luo R. Journal of Physical Chemistry B 2006;110(37):18680–18687.
43. Prabhu NV, Panda M, Yang Q, Sharp KA. Journal of Computational Chemistry 2008;29(7):1113–30. [PubMed: 18074338]
44. Prabhu NV, Zhu P, Sharp KA. Journal of Computational Chemistry 2004;25(16):2049–2064. [PubMed: 15481091]
45. Luo R, David L, Gilson MK. Journal of Computational Chemistry 2002;23(13):1244–53. [PubMed: 12210150]
46. Lu Q, Luo R. Journal of Chemical Physics 2003;119(21):11035–11047.
47. Wagoner JA, Baker NA. Proceedings of the National Academy of Sciences of the United States of America 2006;103(22):8331–6. [PubMed: 16709675]
48. Baker N, Holst M, Wang F. Journal of Computational Chemistry 2000;21(15):1343–1352.
49. Sharp KA, Honig B. Journal of Physical Chemistry 1990;94(19):7684–7692.
50. Micu AM, Bagheri B, Iljin AV, Scott LR, Pettitt BM. Journal of Computational Physics 1997;136(2):263–271.
51. Gilson MK, Davis ME, Luty BA, McCammon JA. Journal of Physical Chemistry 1993;97(14):3591–3600.
52. Im W, Beglov D, Roux B. Computer Physics Communications 1998;111(13):59–75.
53. Gallicchio E, Kubo MM, Levy RM. Journal of Physical Chemistry B 2000;104(26):6271–85.
54. Gallicchio E, Levy RM. Journal of Computational Chemistry 2004;25(4):479–499. [PubMed: 14735568]
55. Pierotti RA. Chemical Reviews 1976;76(6):717–726.

56. Stillinger FH, Debenedetti PG, Chatterjee S. *Journal of Chemical Physics* 2006;125:204504. [PubMed: 17144712]
57. Stillinger F. *Journal of Solution Chemistry* 1973;2:141–158.
58. Weeks JD, Chandler D, Andersen HC. *Journal of Chemical Physics* 1971;54(12):5237–47.
59. Fedorov MV, Kornyshev AA. *Molecular Physics* 2007;105(1):1–16.
60. Cerutti DS, Baker NA, McCammon JA. *The Journal of Chemical Physics* 2007;127(15):155101+. [PubMed: 17949217]
61. Grossfield A. *Journal of Chemical Physics* 2005;122:024506. [PubMed: 15638597]
62. Mobley DL, Barber AE, Fennell CJ, Dill KA. *Journal of Physical Chemistry B* 2008;112(8):2405–14.
63. Chu VB, Bai Y, Lipfert J, Herschlag D, Doniach S. *Biophysical Journal* 2007;93(9):3202–9. [PubMed: 17604318]
64. Tan ZJ, Chen SJ. *Journal of Chemical Physics* 2005;122:044903.
65. Tanaka M, Grosberg AY. *Journal of Chemical Physics* 2001;115(1):567–574.
66. Savelyev A, Papoian GA. *Journal of the American Chemical Society* 2007;129(19):6060–1. [PubMed: 17455935]
67. Chen YG, Weeks JD. *Proceedings of the National Academy of Sciences of the United States of America* 2006;103(20):7560–5. [PubMed: 16670200]
68. Chen AA, Pappu RV. *Journal of Physical Chemistry B* 2007;111(23):6469–78.
69. Tobias DJ, Hemminger JC. *Science* 2008;319(5867):1197–1198. [PubMed: 18309069]
70. Horinek D, Netz RR. *Physical Review Letters* 2007;99(22):226104. [PubMed: 18233302]
71. Horinek D, Serr A, Bonthuis DJ, Bostrom M, Kunz W, Netz RR. *Langmuir* 2008;24(4):1271–1283. [PubMed: 18220430]
72. Azuara C, Lindahl E, Koehl P, Orland H, Delarue M. *Nucleic Acids Research* 2006;34:W38–W42. [PubMed: 16845031]
73. Penfold R, Nordholm S, Jönsson B, Woodward CE. *Journal of Chemical Physics* 1990;92(3):1915–1922.
74. Vitalis A, Baker NA, McCammon JA. *Molecular Simulation* 2004;30(1):45–61.
75. Banavali NK, Im W, Roux B. *Journal of Chemical Physics* 2002;117(15):7381–8.
76. Lee MS, Salsbury J, R. F, Olson MA. *Journal of Computational Chemistry* 2004;25(16):1967–78. [PubMed: 15470756]
77. Okur A, Wickstrom L, Layten M, Geney R, Song K, Hornak V, Simmerling C. *Journal of Chemical Theory and Computation* 2006;2(2):420–433.
78. Mu Y, Yang Y, Xu W. *Journal of Chemical Physics* 2007;127(8)
79. Lum K, Chandler D, Weeks J. *Journal of Physical Chemistry B* 1999;103:4570–4577.
80. Chandler D. *Nature* 2002;417:491. [PubMed: 12037545]
81. Graziano G. *Journal of Physical Chemistry B* 2006;110:11421–11426.
82. Dzubiella J, Swanson JMJ, McCammon JA. *Physical Review Letters* 2006;96:087802. [PubMed: 16606226]
83. Dzubiella J, Swanson JMJ, McCammon JA. *Journal of Chemical Physics* 2006;124:084905. [PubMed: 16512740]
84. Che J, Dzubiella J, Li B, Mccammon JA. *Journal of Physical Chemistry B*. in press
85. Mobley DL, Dill KA, Chodera JD. *Journal of Physical Chemistry B* 2008;112(3):938–946.
86. Tsui V, Radhakrishnan I, Wright PE, Case DA. *Journal of Molecular Biology* 2000;302(5):1101–1117. [PubMed: 11183777]
87. Showalter SA, Baker NA, Tang C, Hall KB. *Journal of Biomolecular NMR* 2004;32(3):179–93. [PubMed: 16132819]
88. Auffinger P, Hashem Y. *Current Opinion in Structural Biology* 2007;17(3):325–33. [PubMed: 17574833]
89. Feig M, Pettitt BM. *Biophysical Journal* 1999;77(4):1769–1781. [PubMed: 10512802]

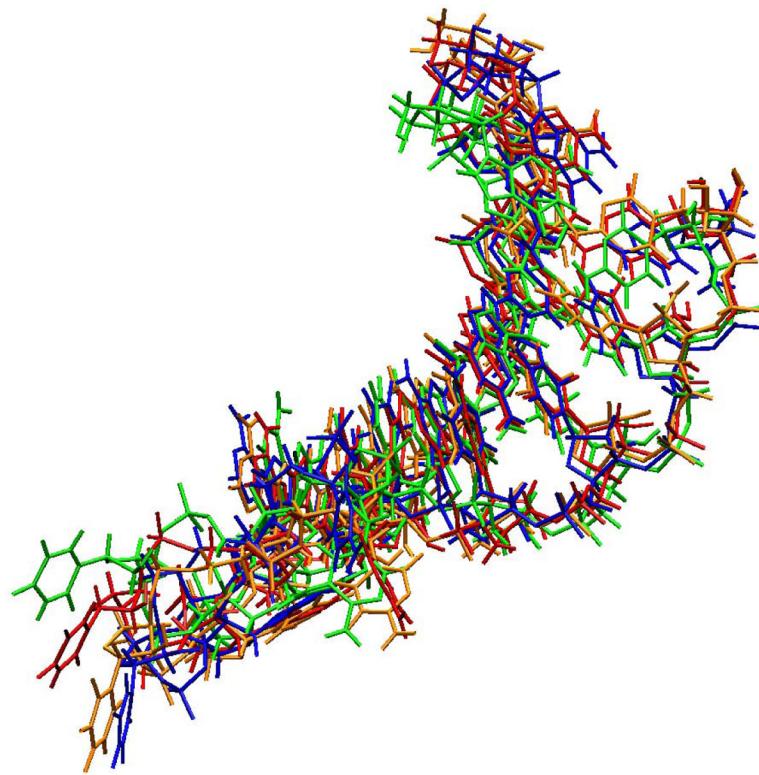
90. Makarov VA, Feig M, Andrews BK, Pettitt BM. *Biophysical Journal* 1998;75(1):150–8. [PubMed: 9649375]
91. Feig M, Pettitt MB. *Journal of Molecular Biology* 1999;286(4):1075–1095. [PubMed: 10047483]
92. Feig M, Pettitt MB. *Biopolymers* 1998;48(4):199–209. [PubMed: 10699840]
93. Auffinger P, Westhof E. *Journal of Molecular Biology* 2000;300(5):1113–1131. [PubMed: 10903858]
94. Auffinger P, Westhof E. *Journal of Molecular Biology* 2001;305(5):1057–1072. [PubMed: 11162114]
95. Young MA, Jayaram B, Beveridge DL. *Journal of the American Chemical Society* 1997;119(1):59–69.
96. Ponomarev SY, Thayer KM, Beveridge DL. *Proceedings of the National Academy of Sciences of the United States of America* 2004;101(41):14771–5. [PubMed: 15465909]
97. Young MA, Ravishanker G, Beveridge DL. *Biophysical Journal* 1997;73(5):2313–2336. [PubMed: 9370428]
98. Mocci F, Laaksonen A, Lyubartsev A, Saba G. *Journal of Physical Chemistry B* 2004;108(41):16295–302.
99. Rueda M, Cubero E, Laughton CA, Orozco M. *Biophysical Journal* 2004;87(2):800–11. [PubMed: 15298889]
100. Perez A, Luque FJ, Orozco M. *Journal of the American Chemical Society* 2007;129(47):14739–14745. [PubMed: 17985896]
101. Savelyev A, Papoian GA. *Journal of the American Chemical Society* 2006;128(45):14506–18. [PubMed: 17090034]
102. Auffinger P, Westhof E. *Biopolymers* 2000;56(4):266–274. [PubMed: 11754340]
103. Csaszar K, Spacková N, Stefl R, Sponer J, Leontis NB. *Journal of Molecular Biology* 2001;313(5):1073–1091. [PubMed: 11700064]
104. Tsui V, Case DA. *Journal of the American Chemical Society* 2000;122(11):2489–2498.
105. Sykes MT, Levitt M. *Proceedings of the National Academy of Sciences* 2007;104(30):12336–12340.
106. Reblova K, Spackova N, Stefl R, Csaszar K, Koca J, Leontis NB, Sponer J. *Biophysical Journal* 2003;84(6):3564–3582. [PubMed: 12770867]
107. Bonvin AM, Sunnerhagen M, Otting G, van Gunsteren WF. *Journal of Molecular Biology* 1998;282(4):859–873. [PubMed: 9743632]
108. Billeter M, Güntert P, Luginbühl P, Wüthrich K. *Cell* 1996;85(7):1057–1065. [PubMed: 8674112]
109. Reddy CK, Das A, Jayaram B. *Journal of Molecular Biology* 2001;314(3):619–632. [PubMed: 11846571]
110. Press, WH.; Teukolsky, SA.; Vetterling, WT.; Flannery, BP. *Numerical Recipes in C*. Cambridge University Press; New York: 1992.
111. Swanson JMJ, Wagoner JA, Baker NA, McCammon JA. *Journal of Chemical Theory and Computation* 2007;3(1):170–83.
112. Kovalenko A, Hirata F. *Journal of Chemical Physics* 2000;113(7):2793–805.
113. Marucho M, Pettitt BM. *Journal of Chemical Physics* 2007;126:124107. [PubMed: 17411108]
114. Yoshida N, Phongphanphanee S, Maruyama Y, Imai T, Hirata F. *Journal of the American Chemical Society* 2006;128(37):12042–3. [PubMed: 16967934]
115. Hildebrandt A, Blossey R, Rjasanow S, Kohlbacher O, Lenhof HP. *Physical Review Letters* 2004;93(10):108104. [PubMed: 15447456]
116. Hildebrandt A, Blossey R, Rjasanow S, Kohlbacher O, Lenhof HP. *Bioinformatics* 2007;23(2):e99–e103. [PubMed: 17237112]
117. Warshel A, Kato M, Pisliakov AV. *Journal of Chemical Theory and Computation* 2007;3(6):2034–45.
118. Lamoureux G, Roux B. *Journal of Physical Chemistry B* 2006;110(7):3308–22.
119. Jiao D, King C, Grossfield A, Darden TA, Ren P. *Journal of Physical Chemistry B* 2006;110(37):18553–9.
120. Ren P, Ponder JW. *Journal of Physical Chemistry B* 2004;108(35):13427–37.
121. Ponder JW, Case DA. *Advances in Protein Chemistry* 2003;66:27–85. [PubMed: 14631816]

122. Schnieders MJ, Baker NA, Ren P, Ponder JW. *Journal of Chemical Physics* 2007;126:124114. [PubMed: 17411115]
123. Marincola FC, Denisov VP, Halle B. *Journal of the American Chemical Society* 2004;126(21):6739–50. [PubMed: 15161302]
124. Dixit SB, Beveridge DL, Case DA, Cheatham r. E. T, Giudice E, Lankas F, Lavery R, Maddocks JH, Osman R, Sklenar H, Thayer KM, Varnai P. *Biophysical Journal* 2005;89(6):3721–40. [PubMed: 16169978]
125. Cheatham TE III, Young MA. *Biopolymers* 2001;56(4):232–256. [PubMed: 11754338]
126. Lindahl E, Hess B, van der Spoel D. *Journal of Molecular Modeling* 2001;7(8):306–317.
127. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. *Journal of Chemical Physics* 1983;79(2):926–935.
128. Pearlmann DA, Case DA, Caldwell JW, Ross WS, Cheatham I, E. T, DeBolt S, Ferguson D, Seibel G, Kollman P. *Computer Physics Communications* 1995;91(13):1–41.
129. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. *Journal of the American Chemical Society* 1995;118(9):2309–2309.
130. Ryckaert JP, Ciccotti G, Berendsen HJC. *Journal of Computational Physics* 1977;23(3):327–341.
131. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. *Journal of Chemical Physics* 1984;81(8):3684–90.
132. Tinker, JW. Ponder. St. Louis; MO: 2008.
133. Price DJ, Brooks I, L. C. *Journal of Chemical Physics* 2004;121(20):10096–103. [PubMed: 15549884]
134. Hess B, van der Vegt NFA. *Journal of Physical Chemistry B* 2006;110(35):17616–26.
135. Höchtl P, Boresch S, Bitomsky W, Steinhauser O. *Journal of Chemical Physics* 1998;109(12):4927–4937.
136. Jorgensen WL, Maxwell DS, Tirado-Rives J. *Journal of the American Chemical Society* 1996;118(45):11225–36.
137. Bondi A. *Journal of Physical Chemistry* 1964;68(3):441–451.
138. Shrake A, Rupley JA. *Journal of Molecular Biology* 1973;79(2):351–64. [PubMed: 4760134]



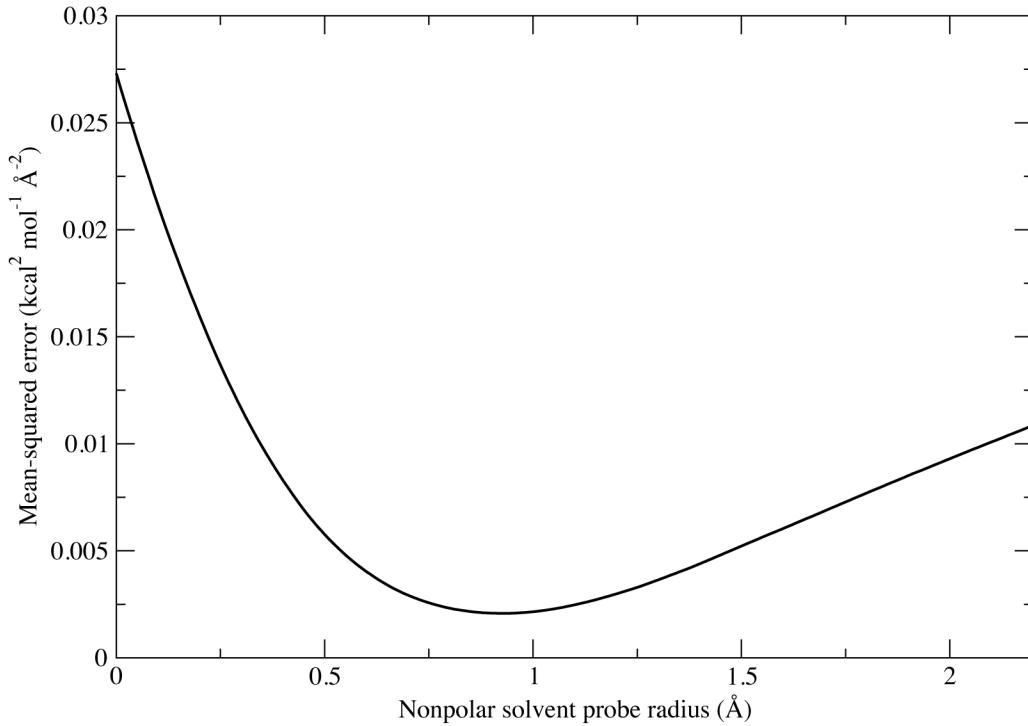
**Figure 1.**

A solvation free energy cycle adapted from Levy et al [17]. The total solvation energy (1) is decomposed into several steps: “charging” the solute in solvent (6) and vacuum (2), including attractive dispersive solute-solvent interactions in solvent (5) and vacuum (3), and cavity formation associated with repulsive solute-solvent interactions (4). The energy associated with Step (7) is generally termed a “nonpolar solvation energy” while the difference in energies associated with Steps (1) and (7) is generally considered as “polar solvation energy”.

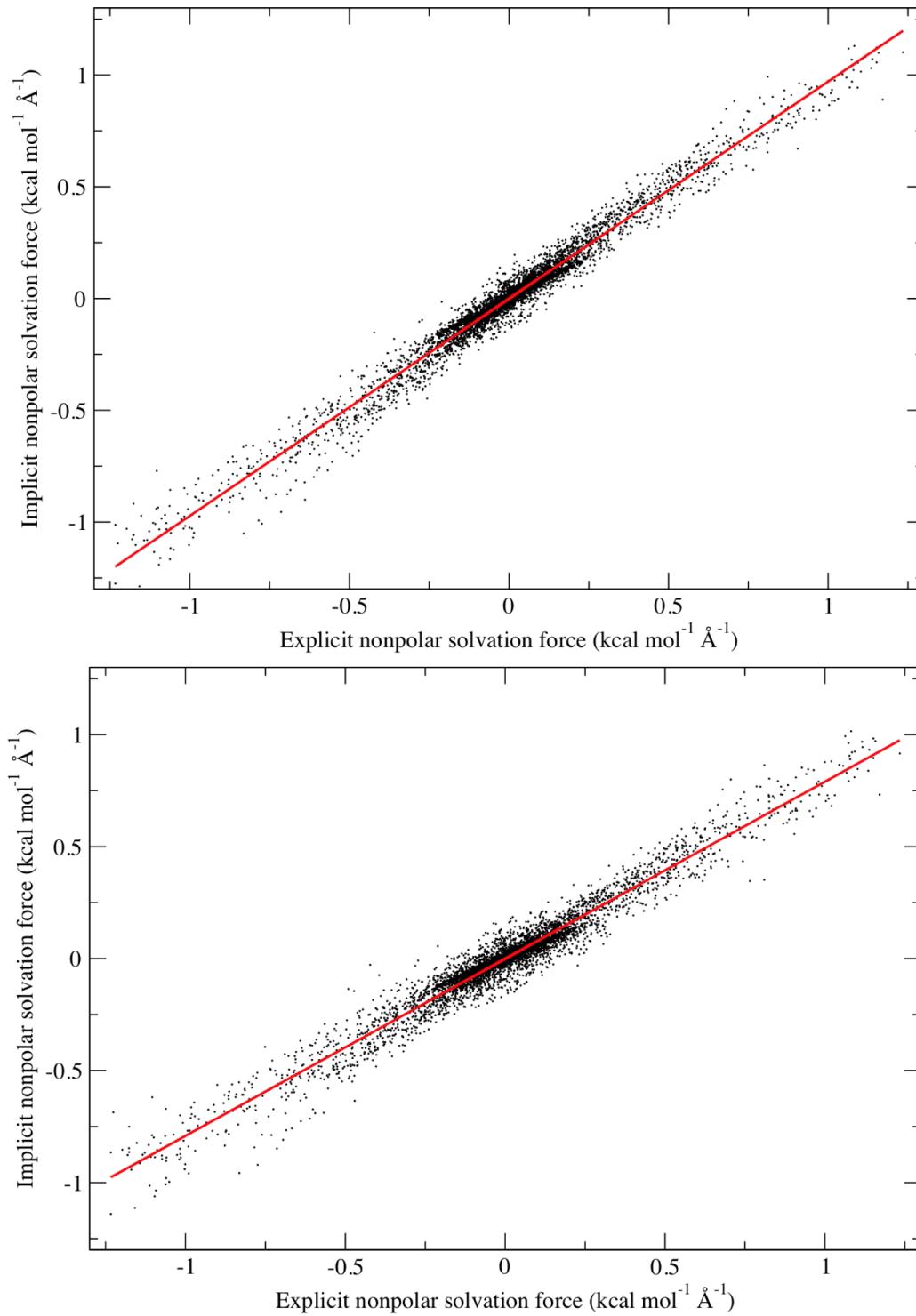


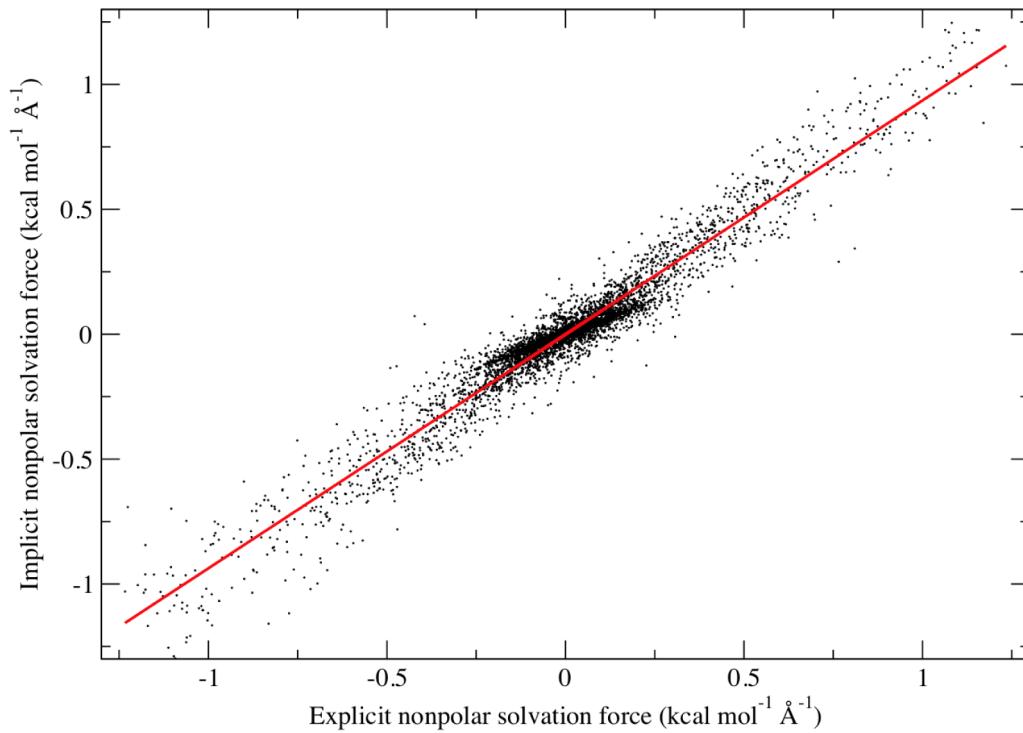
**Figure 2.**

The four representative structures of the RNA hairpin conformations used in for the solvation studies.

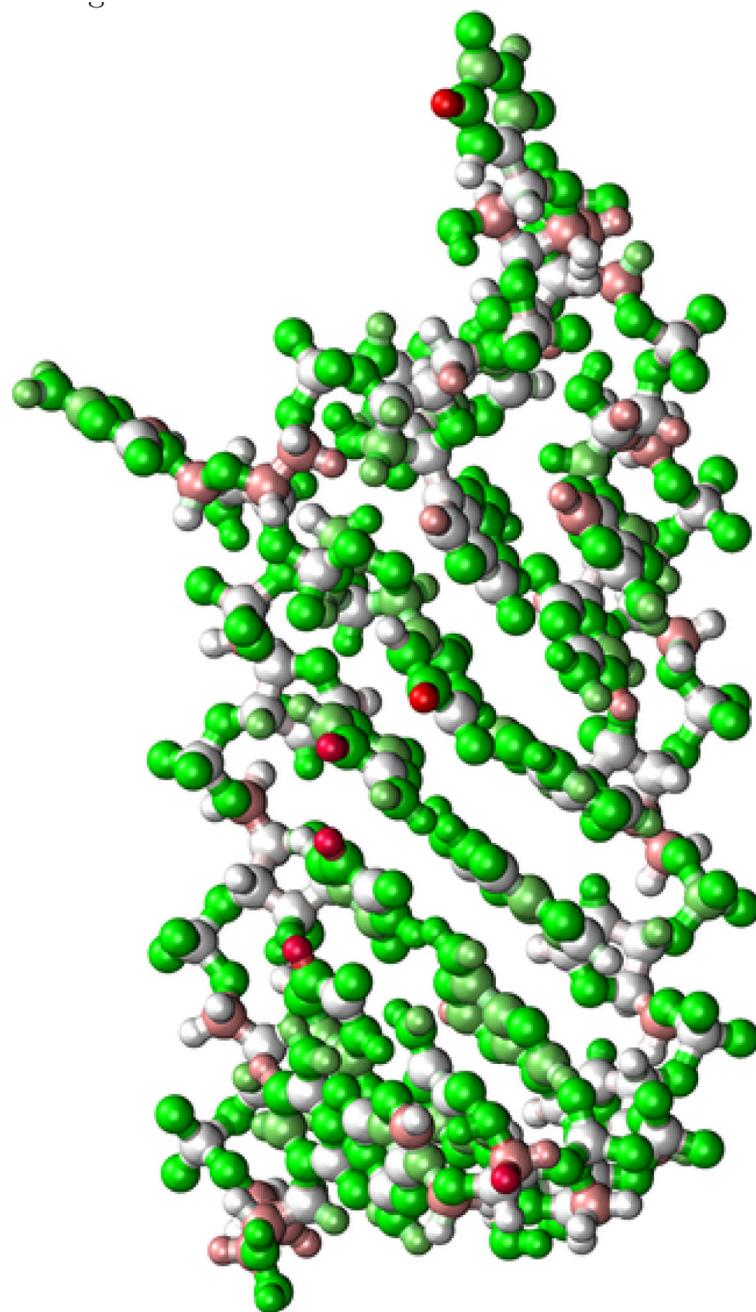
**Figure 3.**

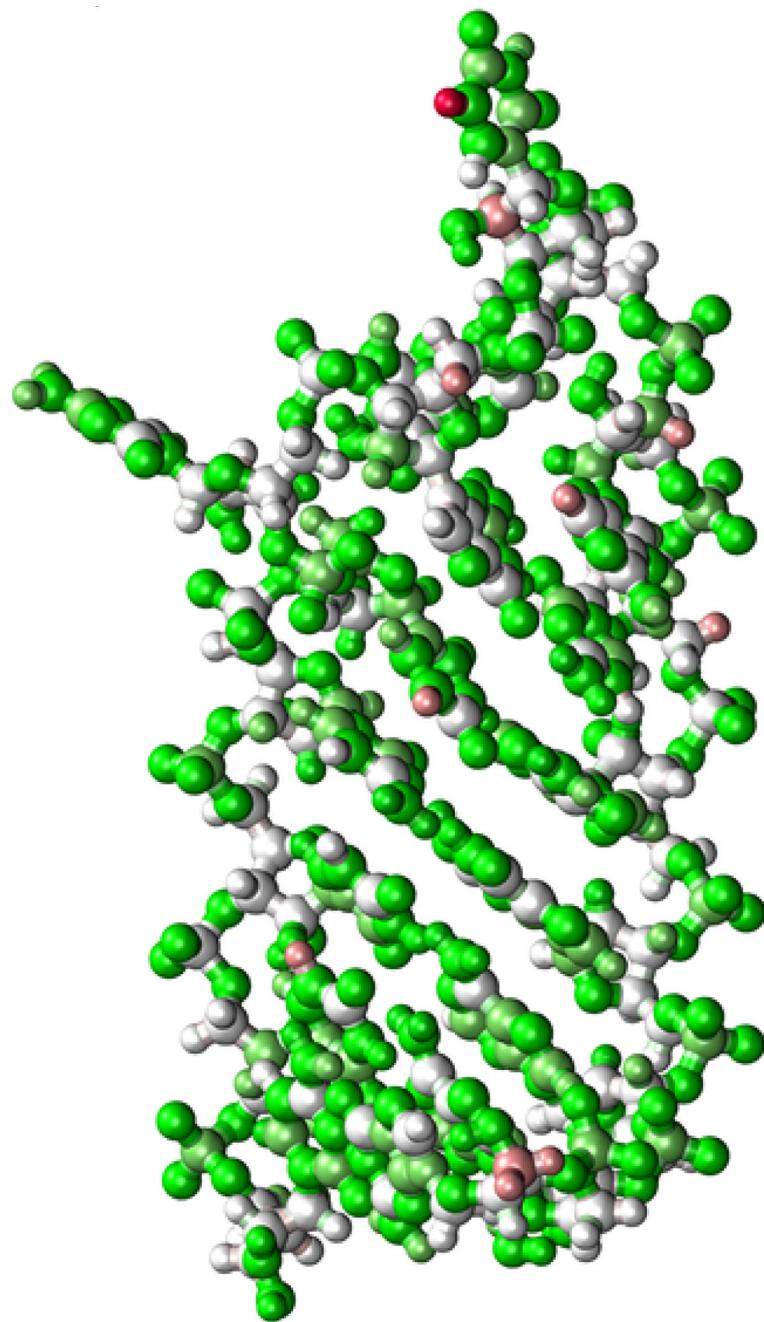
Optimization landscape for the nonpolar implicit solvation model (Eqs. 9 and 19) probe radius. Mean-squared error in the nonpolar solvation force between the explicit and implicit models was used as the fitness measure.

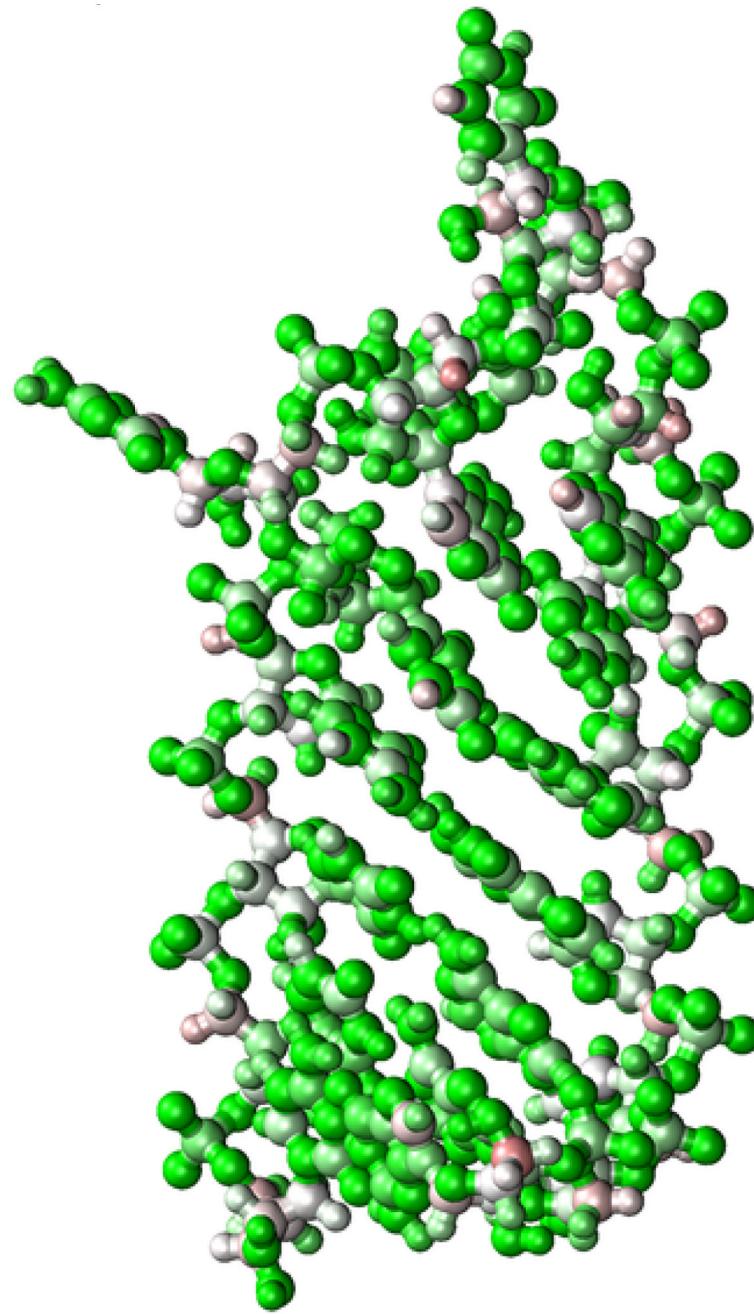


**Figure 4.**

Comparisons of nonpolar solvation forces between implicit and explicit solvent models. The red line represents the best linear regression fit to the data. A. Nonpolar solvation force comparison with the optimal probe radius  $\sigma_s = 0.93 \text{ \AA}$ . See Table 2 for parameter values and associated errors for this plot. B. Nonpolar solvation force comparison with the protein-optimized solvent probe radius  $\sigma_s = 1.25 \text{ \AA}$ . C. Nonpolar solvation force comparison with the traditional solvent probe radius  $\sigma_s = 1.40 \text{ \AA}$ . See Table 3 for parameter values and associated errors for plots B and C.

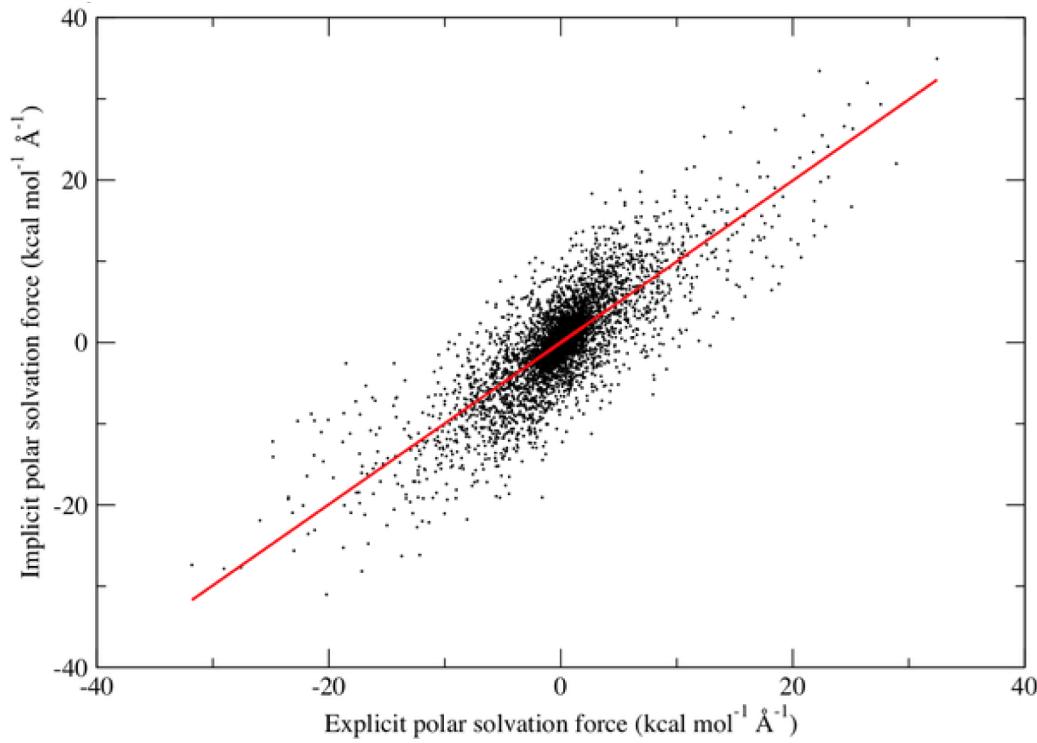




**Figure 5.**

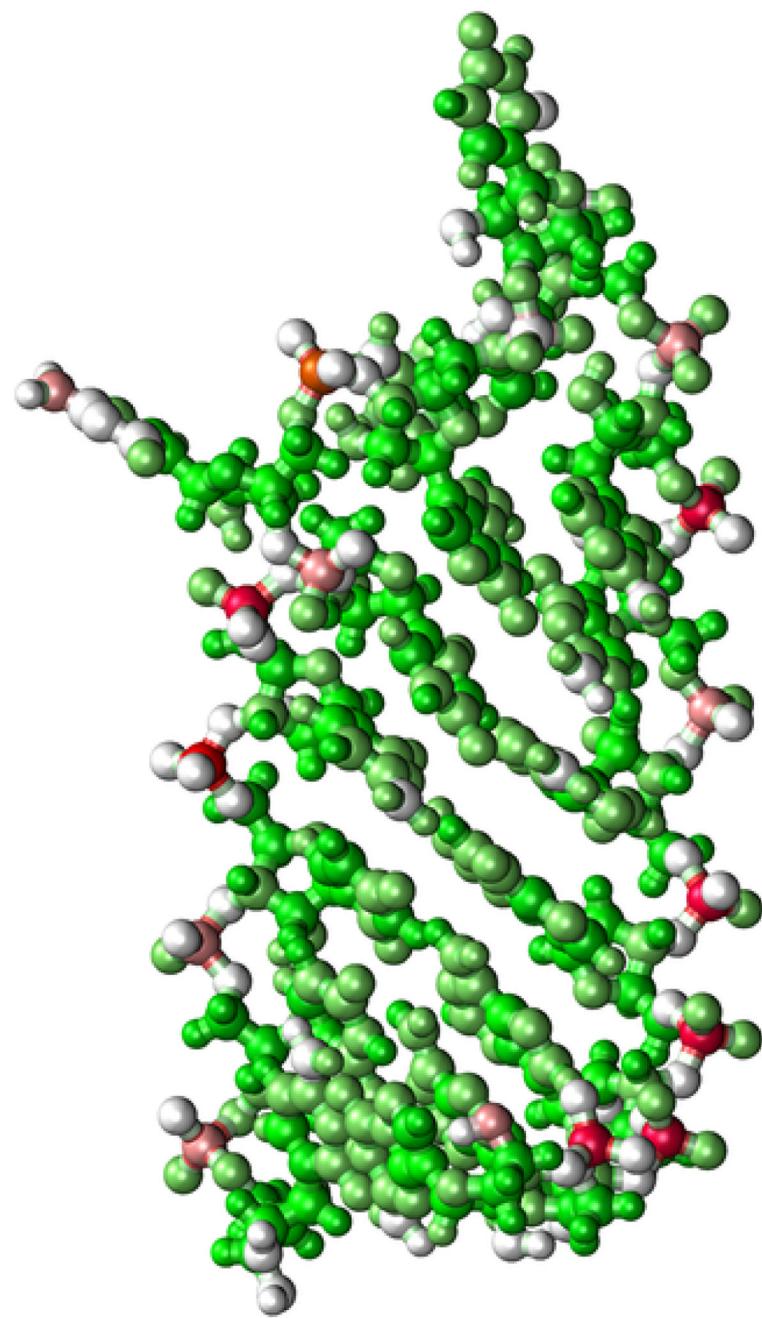
The distribution of *differences* in nonpolar solvation forces on the IRE structure. This figure depicts IRE conformation II colored by the magnitude of the difference in nonpolar solvation force between the explicit and implicit models. This difference was averaged over all conformations used in this study. The colors depict *relative* force magnitude ranging from the smallest (green) to median (white) to the largest (red). A. Nonpolar solvation force difference magnitudes for  $\sigma_s = 0.93$  with optimized  $\gamma$  and  $p$  parameters; differences colored relative to a maximum average error of  $0.257 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ . B. Nonpolar solvation force difference magnitudes for  $\sigma_s = 1.40$  with optimized  $\gamma$  and  $p$  parameters; differences colored relative to a maximum average error of  $0.455 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ . C. Nonpolar solvation force difference

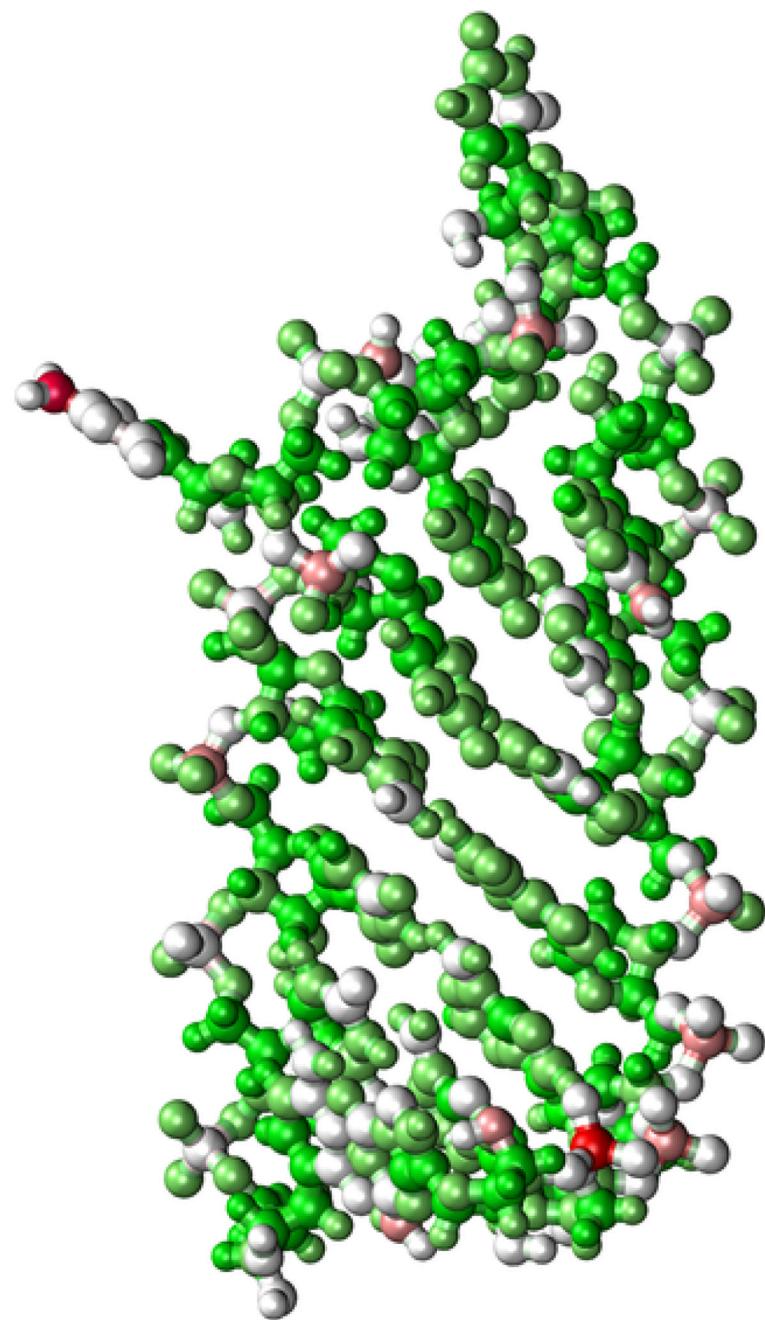
magnitudes with  $\sigma_s$ ,  $\gamma$ , and  $p$  parameters from a previous study [47] for a low-charge density protein; differences colored relative to a maximum average error of 0.494 kcal mol<sup>-1</sup> Å<sup>-1</sup>.

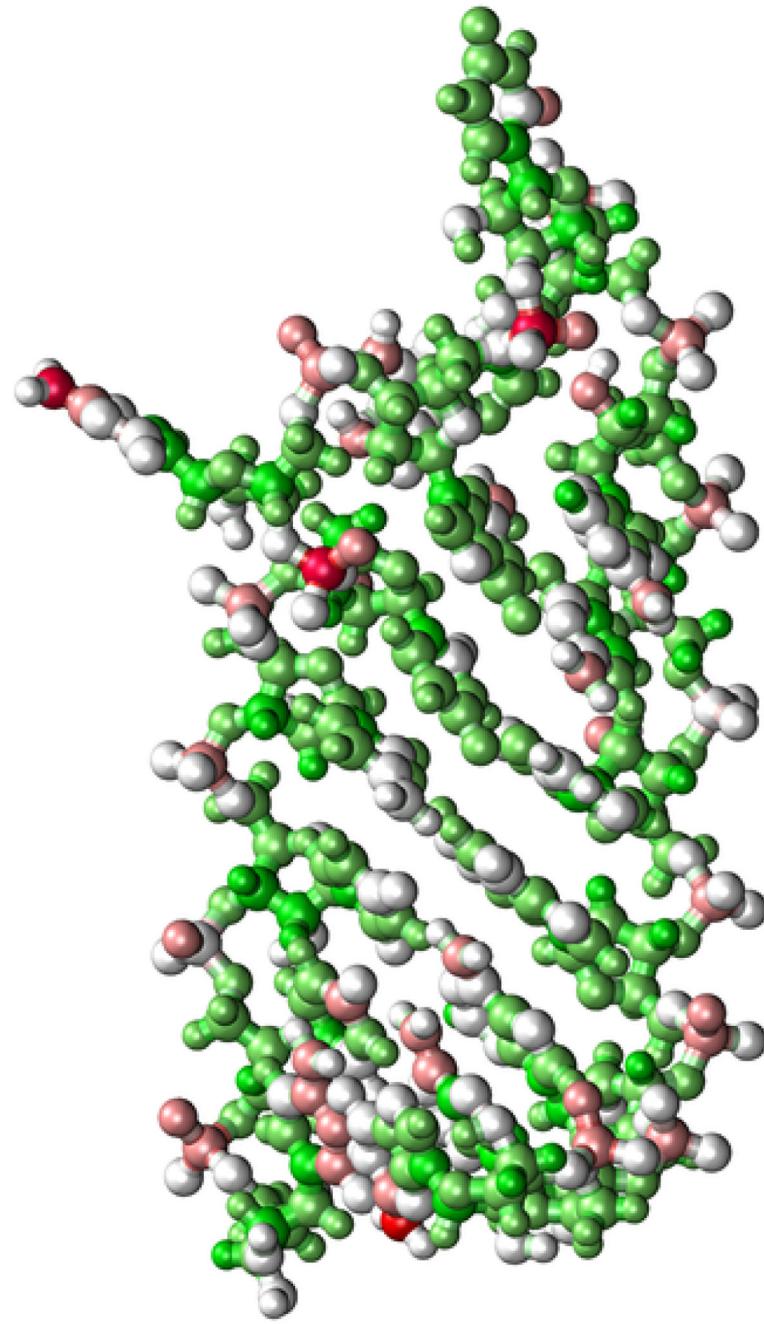


**Figure 6.**

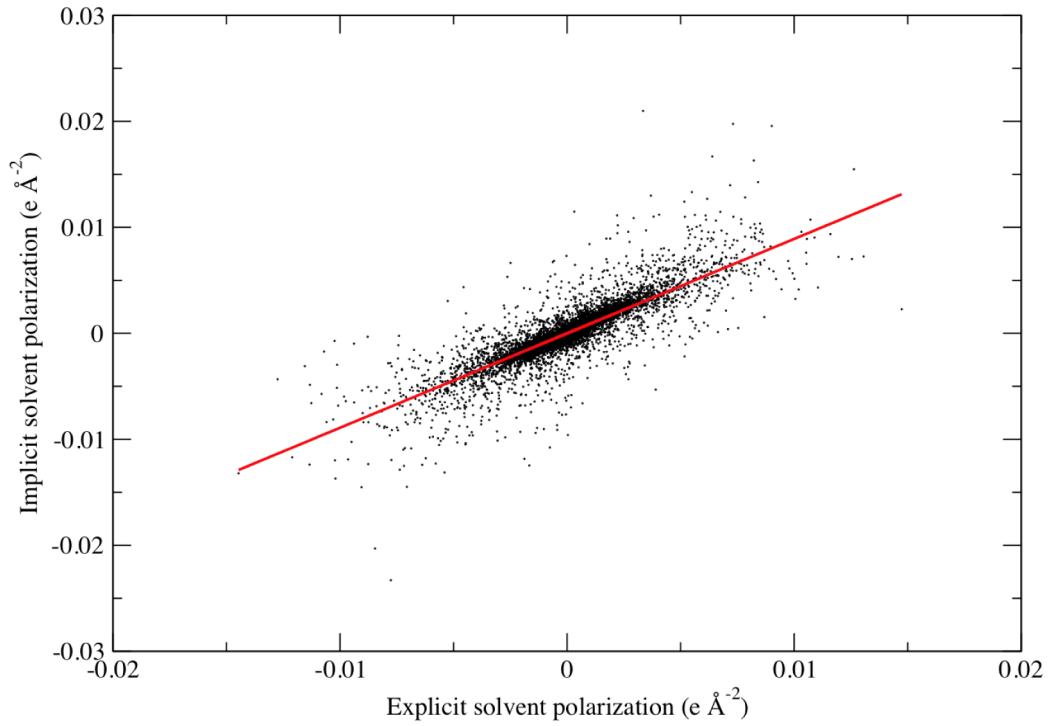
Comparison of polar solvation forces between implicit and explicit solvent models. The red line represents the best linear regression of the data with a slope of  $r = 0.997$  and a Pearson correlation coefficient of  $R = 0.800$ .





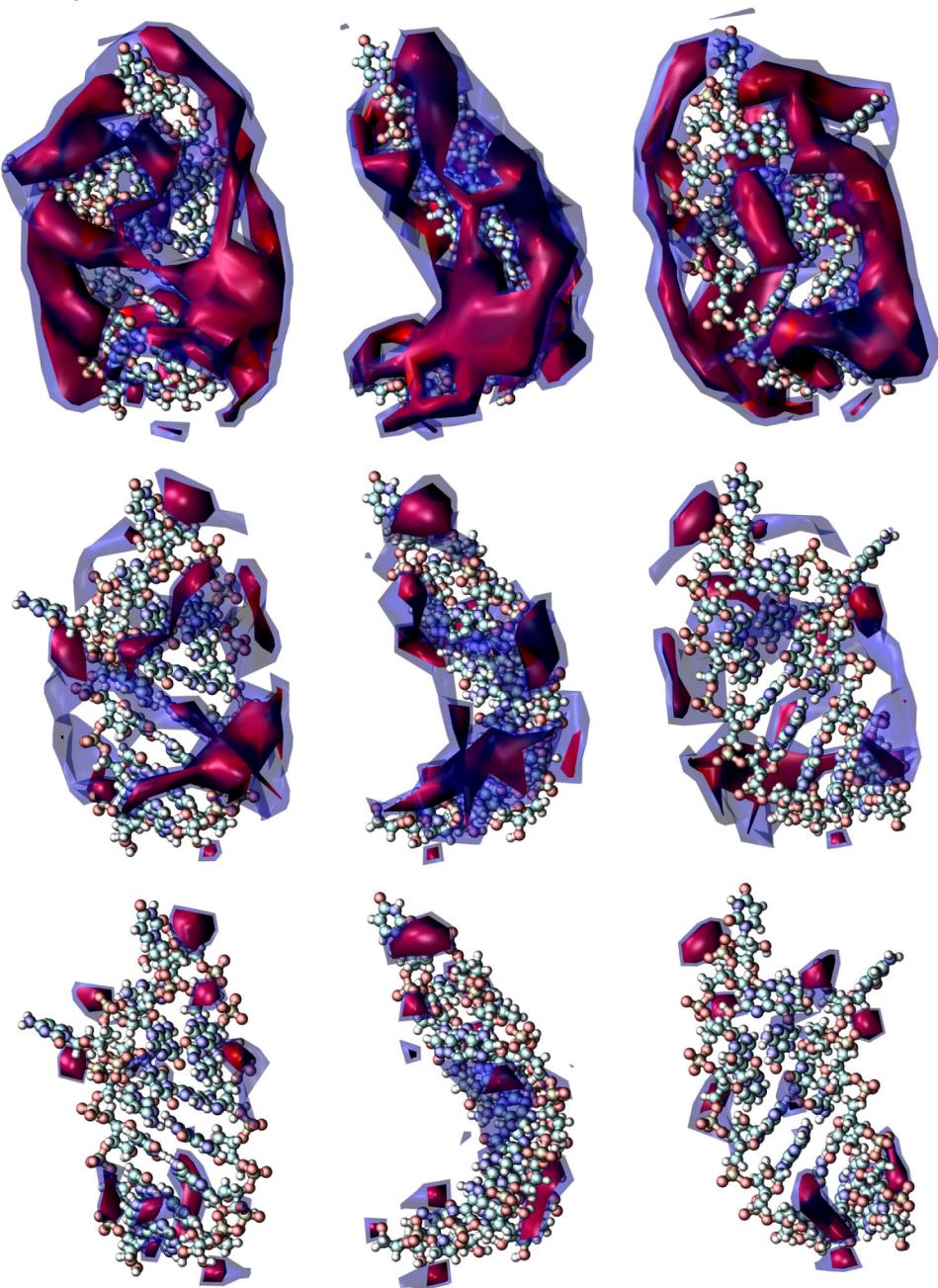
**Figure 7.**

The distribution of polar solvation forces and solvation force differences on the IRE structure. This figure depicts IRE conformation II colored by the polar solvation force magnitude averaged over all conformations used in this study. The colors depict relative force magnitudes ranging from the smallest (green) to median (white) to the largest (red). A. Explicit solvent model polar solvation force magnitudes colored to the largest average force magnitude of  $31.3 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ . B. Implicit solvent model polar solvation force magnitudes colored to the largest average force magnitude of  $36.4 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ . C. Magnitudes of the *difference* in implicit and explicit polar solvation forces colored to the largest average force difference magnitude of  $16.0 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ .



**Figure 8.**

Comparison of solvent polarization between implicit and explicit solvent models. The red line represents the best linear regression of the data with a slope of  $r = 0.910$  and a Pearson correlation coefficient of  $R = 0.846$ .

**Figure 9.**

The distribution of polarization around the IRE structure. This figure depicts IRE conformation II surrounded by isocontours of the relative polarization magnitude *calculated for conformation II*. The blue isocontour depicts 30% of the maximum polarization magnitude while the red isocontour depicts 40% of the maximum polarization magnitude. Each row depicts the same structure rotated by 90° clockwise around the vertical axis. A. Explicit solvent model polarization magnitude. B. Implicit solvent model polarization magnitude. C. Magnitude of polarization difference between explicit and implicit models.

**Table 1**

Pairwise all-atom root-mean-squared deviation values between IRE conformation cluster median snapshots. All values in Å.

Cluster	I	II	III	IV
I	0	1.76	2.23	5.27
II	1.76	0	1.20	3.97
III	2.23	1.20	0	4.03
IV	5.27	3.97	4.03	0

**Table 2**

Optimal parameters for the best-fit nonpolar solvent radii and associated quality of fit metrics (see text for definitions). Parameter 99% confidence intervals are presented in square brackets. The “SASA only” and related labels refer to the model of the nonpolar cavity creation term and are described in detail in the manuscript text.

Parameters	SASA only	SAV only	SASA and SAV
$\sigma_s$ , Å	2.2 [1.24 - 2.20]	0.93 [0.84 - 1.02]	0.93 [0.84 - 1.02]
$\gamma$ , cal mol <sup>-1</sup> Å <sup>-2</sup>	28.75 [23.52 - 33.98]	—	-0.35 [-2.06 - 1.39]
$p$ , cal mol <sup>-1</sup> Å <sup>-3</sup>	—	73.49 [70.92 - 76.06]	73.92 [69.37 - 78.35]
$r$	0.55	0.97	0.97
$R$	0.73	0.98	0.98
$\chi^2$ , kcal <sup>2</sup> mol <sup>-2</sup> Å <sup>-2</sup>	0.0325	0.0021	0.0021

**Table 3**

Parameters for specific nonpolar solvent radii and associated quality of fit metrics (see text for definitions). Parameter 99% confidence intervals are presented in square brackets. Parameters without a \* are the best fit results for that  $\sigma_s$ ; parameters with a \* are from the previous work of Wagoner et al [47] and included to demonstrate the extent of parameter transferability between systems. The “SASA only” and related labels refer to the model of the nonpolar cavity creation term and are described in detail in the manuscript text.

Parameters	SASA only	SAV only	SASA and SAV	
Chosen $\sigma_s$ , Å	1.4	1.4	1.25	1.4
$\gamma$ , cal mol <sup>-1</sup> Å <sup>-2</sup>	36.53 [30.22 - 42.84]	—	0*	1.86 [-0.38 - 4.14]
$p$ , cal mol <sup>-1</sup> Å <sup>-3</sup>	—	60.37 [57.22 - 62.52]	55*	58.25 [52.66 - 63.70]
$r$	0.53	0.94	0.79	0.94
$R$	0.71	0.97	0.97	0.97
$\chi^2$ , kcal <sup>2</sup> mol <sup>-2</sup> Å <sup>-2</sup>	0.0342	0.0045	0.0055	0.0044