



Published in final edited form as:

J Chem Theory Comput. 2013 April 9; 9(4): 2020–2034. doi:10.1021/ct3010485.

Improved Generalized Born Solvent Model Parameters for Protein Simulations

Hai Nguyen^{§,¥}, Daniel R. Roe^{§,#}, and Carlos Simmerling^{§,¥,*}

[§]Department of Chemistry, Stony Brook University, Stony Brook, New York 11794

[¥]Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, New York 11794

[#] Department of Medicinal Chemistry, University of Utah, Salt Lake City, UT, 84112

Abstract

The generalized Born (GB) model is one of the fastest implicit solvent models and it has become widely adopted for Molecular Dynamics (MD) simulations. This speed comes with tradeoffs, and many reports in the literature have pointed out weaknesses with GB models. Because the quality of a GB model is heavily affected by empirical parameters used in calculating solvation energy, in this work we have refit these parameters for GB-Neck, a recently developed GB model, in order to improve the accuracy of both the solvation energy and effective radii calculations. The data sets used for fitting are significantly larger than those used in the past. Comparing to other pairwise GB models like GB-OBC and the original GB-Neck, the new GB model (GB-Neck2) has better agreement to Poisson-Boltzmann (PB) in terms of reproducing solvation energies for a variety of systems ranging from peptides to proteins. Secondary structure preferences are also in much better agreement with those obtained from explicit solvent MD simulations. We also obtain near-quantitative reproduction of experimental structure and thermal stability profiles for several model peptides with varying secondary structure motifs. Extension to non-protein systems will be explored in the future.

Keywords

Implicit solvent; Generalized Born; solvation energy; GBSA; protein folding; perfect radii

Introduction

In order to accurately describe the properties of biomolecules in aqueous environment, solvent effects must be included in the Molecular Dynamics (MD) simulation. Solvation can

*To whom correspondence should be addressed. carlos.simmerling@stonybrook.edu.

Supporting Information Available: RMSD plots for trpzip2 training set, HP36 training set, HP36 test set, Lysozyme test set; RMSD histograms of trpzip2 and tc5b test sets; objective function evolution for several GA runs; salt bridge PMF plot for RAAE simulation with GB-Neck2; RMSD histogram of HP5F and tc5b trajectories at 300K; table of optimization results, tables of training and test sets, table of mbondi3 radii; table showing temperatures for REMD runs; tables showing average percent secondary structures and local conformational propensities from Ala10 and HP-1 REMD simulations. Structures for training and test sets as well as a Python script to change mbondi2 to mbondi3 radii are available on request.

be explicitly represented as atomistic solvent molecules or it can be implicitly represented by a model that calculates solvation effects using a continuum representation. Although implicit solvent is less realistic than explicit solvent model, it is still widely used¹ due to low computational cost, and many models directly provide solvation free energies as compared to the potential energies provided by explicit models. This has led to wide use in the drug discovery field of implicit solvent models in post-processing trajectories originally performed in explicit solvent.² In addition, the low viscosity in implicit solvent simulations can accelerate the rate of conformational sampling (such as protein folding) compared to explicit solvent.³

Solvation free energy can be decomposed into two terms for the polar and non-polar contributions. The present work focuses solely on the polar contribution. The non-polar term is often approximated by the equation $\Delta G_{np}=\gamma A$ where γ is the surface tension coefficient and A is the total solvent accessible area. The nonpolar term is frequently omitted in simulations due to the cost of calculating the surface area and its derivatives, and the fact that the magnitude of this term is typically much smaller than the polar contribution. Moreover, a simple solvent accessible surface area (SASA) based approximation that is commonly used to calculate the nonpolar term has several limitations.⁴ Chen et al.^{4a} have shown that this non-polar model tended to overestimate nonpolar interactions that shifted ensembles to non-native states. Despite these limitations, SASA-based approaches are widely used and available in the Amber program⁵, thus we evaluate the impact of their inclusion during simulations using our improved GB model.

Among all implicit solvent models, the Poisson- Boltzmann (PB) method⁶ is considered the most accurate model for calculating polar solvation energy in MD. However, the computational cost of solving the PB equation and its derivatives, particularly on massively parallel computers, is high enough that it is not widely used in MD simulations.⁷ Instead, most MD simulations use the GB equation (eq. 1), as was first introduced by Still et al.⁸ and subsequently modified by other groups.

$$\Delta G_{GB} = -\frac{1}{2} \left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_{out}} \right) \sum_{i,j} \frac{q_i q_j}{f_{ij}^{GB}(r_{ij})} \quad (1)$$

Here $f_{ij}^{GB} = \sqrt{r_{ij}^2 + R_i R_j \exp\left(\frac{-r_{ij}^2}{4R_i R_j}\right)}$; q_i, q_j are the partial charges of atom i and j ; r_{ij} is the distance between atom i and j ; ϵ_{in} and ϵ_{out} are interior and exterior dielectric constants respectively. R_i and R_j are the effective Born radii. It has been shown that accurate calculation of effective radii (or using 'perfect' radii calculated from PB method) is a key to close agreement between GB and PB solvation energies.⁹ The effective radius is normally calculated by eq. 2

$$R_i^{-1} = \rho_i^{-1} - I_i \quad (2)$$

where I_i is Coulomb integral derived from Coulomb Field Approximation (CFA)

$$I_i = \frac{1}{4\pi} \int_{\Omega, r > \rho_i} \frac{1}{r^4} d^3r \quad (3)$$

ρ_i is the intrinsic radius of the atom i and integral I_i is calculated over the volume Ω outside atom i but inside the molecule. I_i can be calculated numerically⁸ or analytically by using the pairwise descreening approximation (PDA) method introduced by Hawkins et al. (the GB-HCT model).¹⁰ Although GB-HCT is less computationally expensive than numerical methods,⁷ it tends to underestimate the effective radii of buried atoms.¹¹ A modification based on GB-HCT was proposed by Onufriev et al.¹² (GB-OBC), in which effective radii for buried atoms are scaled up by an adjustable empirical parameter (α, β, γ) set (eq. 4a).

$$R_i^{-1} = \tilde{\rho}_i^{-1} - \rho^{-1} \tanh(\alpha\psi - \beta\psi^2 + \gamma\psi^3) \quad (4a)$$

$$\text{where } \tilde{\rho}_i = \rho_i - \text{offset}, \psi = \tilde{\rho}_i I \quad (4b)$$

Importantly, these analytical models (GB-HCT and GB-OBC) use the van der Waals (VDW) surface to define the boundary between solvent and solute, instead of using more realistic but much more computationally demanding molecular surface (MS). Mongan et al.¹³ introduced a “neck” correction to make the space defined by the VDW boundary closer to that defined by MS boundary, particularly at small interatomic distances where finite size explicit water is typically excluded (GB-Neck).

$$I_{MS} = I_{vdw} + \int_{neck} \frac{1}{r^4} d^3r \quad (5)$$

where I_{vdw} is the integral I_i in eq. 3, using VDW volume for volume Ω . I_{MS} is then applied as I in eq. 4b.

All 3 of these PDA-based GB models have some advantages such as low computational cost,^{7, 13} and in particular efficient parallel scaling compared to explicit solvent models.¹⁴ These GB models have also been ported to GPU-based MD codes which accelerate MD up to 700 times faster than simulation on conventional CPUs.¹⁵ The advantage of speed, however, comes with less accuracy in these GB models. GB-HCT and GB-OBC have apparent limitations such as high alpha helical content¹⁶ and overly strong ion interactions compared to TIP3P explicit solvent simulations.^{16b, 17} Although GB-Neck introduced corrections to GB-OBC, this is not reflected in improved solvation energy accuracy.¹³ Additionally, Dill et al.^{16e} and Roe et al.^{16a} have shown that GB-Neck tends to destabilize native peptide/protein structures, likely due to imbalance between intramolecular hydrogen bonds and interaction with implicit solvent.

Our goals for improving the GB model are to give more accurate solvation energy and effective radii calculation compared to PB method; to reduce secondary structure and salt bridge bias, and to better reproduce experimental structures and thermal stability for small proteins and peptides. We hypothesize that at least some of these weaknesses could be improved by more rigorous fitting of the many empirical parameters in these models. Since

GB-Neck is more physically realistic than GB-HCT and GB-OBC, we decided to use it as the base model for our parameter refitting. The relatively poor performance of GB-HCT in many studies led us to omit it from the present comparisons.

In the original GB-Neck work,¹³ 8 parameters were optimized by fitting GB solvation energies to PB solvation energies for a set of proteins and peptides. The GB-Neck parameters include scaling factors S_x ($x=H, C, N, O$) that were initially introduced in GB-HCT by Hawkins et al.¹⁰ for analytically calculating the I integral in eq. 3, the $\{\alpha, \beta, \gamma\}$ set used in eq. 4a that was initially introduced in GB-OBC by Onufriev et al.,¹² and the neck scale factor S_{neck} introduced by Mongan et al.¹³ These describe properties related to gaps between atom pairs, and are thus likely dependent on size of the atoms involved. We therefore expanded the number of parameters from 8 to 18 (see method section) by making $\{\alpha, \beta, \gamma\}$ atomic number dependent and making *offset* (eq. 4b) a free parameter as well. We recognized that the significant increase in the number of free parameters in the model necessitated use of much larger training and test sets than used previously, and thus much of the present work focuses on development of a large and broad data set for training and testing.

In our training objective function, not only absolute solvation energy but also effective radii and relative solvation energy of peptide (or protein) conformations were included. PB solvation energies and ‘perfect’ radii of structures in the training set were used as benchmarks for fitting. The new GB-Neck parameter set (GB-Neck2) shows significant improvement in accuracy of calculating solvation energy and effective radii compared to GB-OBC and original GB-Neck model for this training set. Importantly, the improvement is clearly transferable to test sets having thousands of structures for various proteins and peptides, including molecules not used in training.

The final goal of a GB model is to approximate the results (structure, stability, salt bridge profile etc.) obtained from more expensive explicit solvent simulations; thus we performed simulations of several peptides in GB-Neck2 as well as in explicit water to test if the improved agreement of GB-Neck2 to PB results (solvation energy and effective radii) led to improved agreement of structural ensembles compared to those obtained from explicit solvent simulation. Overall, the GB-Neck2 model does a much better job in reproducing ensemble data from explicit water (such as alpha-helical stability) as compared to GB-OBC and comparable to the original GB-Neck, with the exception of propensity to form ion pairs (salt bridges). We found that although salt bridges were specifically included in our training by fitting to PB solvation energies, they tended to remain too strong in GB-Neck2 when comparing to TIP3P simulation. A possible explanation is that PB also has too-strong ion interactions compared to TIP3P, perhaps arising from our use of the same set of intrinsic Born radii in our GB and PB calculations for consistency.¹⁸ Salt bridge strength was thus adjusted in the same strategy as Geney et al.^{17a} and Shang et al.¹⁸ by empirically adjusting the Born radius of side-chain H^{N+} of Arg to reproduce salt bridge PMF of TIP3P simulation. Unlike those earlier studies, we also adjusted the Born radius of side-chain $O\epsilon$ of Glu (and $O\delta$ of Asp) to match PMF profiles of salt bridges and hydrogen bonds in TIP3P simulations. This radius modification was sufficient to reproduce the PMF of Lys salt bridge formation and we found no need to modify the Born radius of H^{N+} in Lys.

We also tested the ability of GB-Neck2 in combination with widely used ff99SB force field¹⁹ in reproducing experimental structure and thermal stability of different peptides from experiment by simulating a hairpin (HP5F)²⁰ system and a mini-protein with α , 3_{10} and polyproline helices and a small hydrophobic core (trp-cage variant tc5b).²¹ The effect of including a nonpolar solvation energy term in GB simulations was also tested. Although the agreement of melting temperature between simulation and experiment depends not only on the GB model but also on the protein force field, this testing is still valuable to confirm the robustness of the combination of specific GB model and force field. Dill et al.^{16e} evaluated various combinations of force fields and GB models for peptide and protein simulations and found that GB-OBC¹² with the ff96 force field was the best combination for this application. However, this GB model and this force field both have well-known flaws, and thus it is likely that the combination benefits from significant fortuitous error cancelation. In the present case, we use the ff99SB protein force field (FF),¹⁹ which has been shown by many studies to provide excellent results with explicit water.²² We find that this single combined protein FF + solvent model is able to quantitatively reproduce the experimental thermal stability behavior of two tested peptide models with different secondary structures. Taken together, our results lead us to recommend this combination for simulations of peptides and proteins.

Materials and Methods

Training set for parameter fitting

We first designed test sets of between ~3,500 - 103,000 structures of each protein or peptide, and then took a subset of the structures for the training set. The subset was selected in a way that gives both training and test set similar absolute solvation energy root-mean-square-deviation (*abs_e*) between GB-OBC and PB solvation energies. For example, the Ala10 test set had 50000 structures with *abs_e* of 1.12 kcal/mol between GB-OBC and PB. The Ala10 training set had only 413 structures with *abs_e* of 1.14 kcal/mol. This reassures us that a small number of structures could represent a desired quality metric (*abs_e* in this case) of a larger number of structures. The assumption is tested by evaluating the model using the full test set, which was impractical during training due to the large number of parameter variations that were tested.

An overall summary of the training sets and their contributions to the training objective function is given in Table 1. These sets are discussed in more detail below.

Roe et al.^{16a} used REMD simulations of Ala10 peptide to quantify the helical bias in GB-OBC and GB-HCT models. We used this system in our test and training sets; the training set Ala10_set_1 has 480 structures extracted from TI and REMD trajectories from Roe et al.^{16a} We first introduced 50 alpha and 50 hairpin structures from TI trajectories and then added 10 structures from each of the 20 most populated clusters sampled at 300K in 50 ns REMD using TIP3P²³ explicit water, as well as single representative structures for the next 180 clusters.

Okur et al.^{17b} used the peptide sequence RAAE (Arg-Ala-Ala-Glu) to evaluate salt bridges in the GB-OBC model. Our RAAE set has 200 structures taken from the 300K trajectory of

TIP3P REMD simulation from Okur et al.^{17b} We chose structures uniformly sampling the salt bridge distance ($C\zeta$ of Arg and $C\delta$ of Glu) ranging from 3.6 Å to 14.5 Å with nearly equal interval of 0.05 Å.

We also added structures for two peptides having different secondary structures: β -hairpin (trpzip2, PDB ID: 1LE1)²⁴ and α -helix (3Ai3)²⁵; these also have more complicated side chains than Ala10 and RAAE. The trpzip2 set had 413 structures from Okur et al.²⁶ Those structure ensembles were chosen from cluster analysis of MD (or REMD) simulation trajectories, giving various types of backbone structures from helix, hairpin, PPII, and coil. The backbone RMSD for those structures to native trpzip2 is presented in figure S1. The 3Ai3 set had 200 structures of the peptide sequence Ac-YGG-(KAAAA)₃-K-NH₂, one of the helical peptides studied by simulation and NMR in Song et al.²⁵ We chose structures first by clustering the first 50ns of 300K trajectory data from GB-HCT REMD simulation of 3Ai3²⁵ and then picking 200 structures from the 20 most populated clusters (10 structures / cluster).

Because Ala10, RAAE, trpzip3, 3Ai3 were small peptides, we also added HP36 mini-protein²⁷ structures to train for structures having a hydrophobic core. These were extracted from the first 75 frames of 300K MD simulation in TIP3P from Wickstrom et al.²⁸ (the backbone RMSD to X-ray structure (PDB ID: 1YRF)²⁹) is given in figure S2).

Structure sets described above were used for training solvation free energy as compared to PB data. We also included two structure sets, Ala10_set_2 and HP1113, to train for effective Born radii. Ala10_set_2 has 200 Ala10 structures (50 structures for each alpha helix, hairpin, left handed helix (“left”), PPII) which were extracted from trajectories of TI calculations from Roe et al.^{16a} We added additional large protein structures to evaluate the effective radii underestimation of deeply buried atoms.¹¹ The HP1113 set has 6 large proteins having various secondary structure types, with PDB ID codes 1TSU,³⁰ 1BDD,³¹ 1UBQ,³² 1AEL,³³ 1FKG,³⁴ 3GB1³⁵ (details in table S2).

Test sets for evaluating the new model

We designed two test set types. Test set type I (Ala10, trpzip2, 3Ai3, RAAE, HP36 sets) had proteins or peptides for which a smaller set of structures were included in the training stage. This tests the extension of the model to a broader set of structures (thousands rather than hundreds). Test set type II had proteins or peptides for which structures were not included in solvation energy training (tc5b, DPDP, HIV1-PR, Lysozyme), testing the transferability to entirely different molecular systems. The large numbers of structures were chosen for testing to ensure local variation in structure as well as alternative folds. Summary for the test sets is given in table S3

Test set type I—Ala10 set has 50000 Ala10 structures taken from 50 ns of 300K trajectory of REMD simulation in TIP3P.^{16a} Trpzip2²⁴ set has ~80000 structures having RMSD to native structure ranging from 0.2 to 7.6 Å (figure S3) which were taken from TIP3P and GB simulations.²⁶ The 3Ai3 set had 49000 structures taken from 49ns of 300K trajectory of GB-HCT REMD simulation.²⁵ The RAAE set had 50000 structures from 50ns of 300K trajectories of TIP3P REMD simulation of RAAE peptide.^{17b} The HP36²⁷ set had

3500 structures extracted from the first 35ns (skipping every 10 frames) of TIP3P MD simulation at 300K from Wickstrom et al.²⁸ (figure S4).

Test set type II—Test set type II has 4 structure sets representing 4 different protein types, which are a helical mini-protein (trp-cage tc5b variant),²¹ a small peptide having 3-stranded β -sheet (DPDP),³⁶ a larger, mainly helical protein (lysozyme)³⁷ and a larger protein having mainly β -sheet (HIV-1 protease).³⁰ The tc5b set had 103000 structures having backbone RMSD from 0.3 to 8.0 Å to native TC5b (figure S5), which were extracted from TIP3P and GB simulation ensembles.^{17a} The DPDP set had 50000 structures from 150 ns GB-HCT REMD simulation trajectory at 300K.³⁸ HIV-1 PR had 1427 structures having closed, semi-open and wide open conformations of HIV-1 PR protein which were extracted from 600 ns trajectory at 300K of 1TSU³⁰ in TIP3P (C α RMSD to closed X-ray structure is given in figure S6)³⁹. The lysozyme set had 1000 structures taken from first 30 ns of 300K trajectory of TIP3P MD simulation (backbone RMSD to experimental native structure (PDB ID: 1IEE³⁷) is given in figure S7).¹⁹

PB calculations and intrinsic radii—All PB calculations were performed using Delphi v2 and v4⁴⁰ with grid spacing of 0.25 Å and solvent probe of 1.4 Å (different Delphi versions were used due to their availability in computer clusters). We selected Delphi based on comparison of performance of different PB solvers for systems of the type studied in the present work.⁷ Interior and exterior dielectric constants of 1.0 and 78.5 respectively were used for solvation energy calculation. Calculation of ‘perfect’ radii was done as described by Onufriev et al.⁹, using the same PB parameters as for our PB solvation energy calculations, with the exception that an exterior dielectric constant of 1000 was used, as suggested by Sigalov et al.⁴¹ The original GB-Neck suggested use of the bondi radii set,⁴² however it was shown by Dill et al.^{16e} that this combination tended to destabilize protein native structure. Onufriev et al.¹² showed that GB-OBC worked quite well with mbondi2 radii, consistent with our previous observations with this combination.^{18, 43} We reasoned that GB-Neck was an improvement of GB-OBC model and thus, mbondi2, instead of bondi, should be a good starting radii set. For consistency, the same radii were used in GB and PB. We therefore used mbondi2 intrinsic Born radii set¹² and charge set from ff99SB¹⁹ in all PB and GB solvation energy calculations. Radii adjustment will be discussed below.

Fitting parameters and procedure

In the original GB-Neck model,¹³ Mongan et al. fit only 8 parameters: S_x ($x=H, C, N, O$), $\{\alpha, \beta, \gamma\}$ and S_{neck} . We refit 18 parameters by allowing $S_x, \alpha, \beta, \gamma$ to vary for H, C, N, and O. We tried several parameter combinations for Sulfur (S) atom and found that S parameters have insignificant effect on correlation between GB and PB solvation energies due to a small number of S atoms in protein molecule. Thus, the S parameters were arbitrarily chosen to be the same as the ones for Oxygen (O). S_x is a scaling parameter originally introduced by Hawkins et al. in the pairwise GB-HCT model¹⁰ to avoid double counting of overlapping VDW volume. S_x was conventionally considered to range from 0 to 1, but the search space was extended greater than 1 in the original GB-Neck paper.¹³ In our optimization, we also extended the range of S_x to [0.0, 2.0]. $\{\alpha, \beta, \gamma\}_x$ ($x=H, C, N, O$) are adjustable parameters used in eq. 4a. Onufriev et al.¹² and Mongan et al.¹³ used one set of $\{\alpha, \beta, \gamma\}$ for all atoms,

but we allowed different elements to adopt their own parameter to allow for atomic-size dependence of the interstitial gaps for which these parameters empirically correct. We chose [0.0-10.0] as potential range for these parameters. We also attempted fitting with one parameter set for all elements like Onufriev et al.¹² or Mongan et al.¹³ did, but found no significant improvement in solvation energy calculation compared to GB-OBC and GB-Neck (data not shown). The *offset* parameter (eq. 4b) was originally used by Still et al.⁸ to decrease atomic radii to maximize the agreement between GB and experimental solvation energies for a set of small molecules, and it has been used in several GB models such as GB-HCT,¹⁰ GB-OBC¹² and GB-Neck¹³ as a conventional constant (*offset*=0.09). In our study, we treated *offset* as an adjustable parameter with possible range of [-0.2, 0.2]. S_{neck} is the scaling factor introduced by Mongan et al.¹³ to avoid the overlap of neck regions in nearby pairs, and thus reducing the over calculating of neck integral (eq. 5). We kept the original range of [0.0, 1.0] for S_{neck} . In summary, the search range of each parameter set is $S_x \in [0.0, 2.0]$, $\{\alpha_x, \beta_x, \gamma_x\} \in [0.0, 10.0]$ ($x=H, C, N, O$), *offset* $\in [-0.2, 0.2]$ and $S_{neck} \in [0.0, 1.0]$.

GB MD simulation biases have been shown to correlate with differences between GB and PB solvation energies.^{16a} In this work, we define “absolute solvation energy root-mean-square-deviation” (*abs_e*) as RMSD of solvation energies for a set of conformations, where the error is the difference between GB and PB energies. “Relative solvation energy root-mean-square-deviation” (*rel_e*) was calculated as the RMSD for GB and PB energy differences for all pairs of structures. “Effective radii root-mean-square-deviation” (*eff_rad_rmsd*) is defined as RMSD of GB effective Born radii from those calculated using PB (‘perfect’ radii).⁹ Typically, only *abs_e* is considered when optimizing GB models.¹²⁻¹³ However, we consider *rel_e* and *eff_rad_rmsd* as important additional targets. The *rel_e* is included since it is the relative energy of alternate conformations that determines thermodynamic populations, such as those sampled in MD simulations. The *eff_rad_rmsd* is included since Onufriev et al.⁹ showed that the best agreement between GB and PB solvation energy is obtained when using ‘perfect’ radii from PB calculation in GB, later confirmed by Honig et al.^{16d} We set our objective function for training as the sum of weighted *abs_e*, *rel_e* and *eff_rad_rmsd*. The objective function is shown in eq. 6, where w_i is weighting factor and x_i is contribution of each component i (each set in Table 1).

$$obj_funct = \sum_i w_i x_i. \quad (6)$$

We weighted *abs_e*, *rel_e*, and *eff_rad_rmsd* so that they contributed roughly equally to the objective function. We first minimized the objective function with $w_i = 1$ for all contributions and calculated how far each contribution value could be decreased from those calculated by GB-OBC model. The *abs_e* values for trpzip2, 3Ai3, HP36 decreased a few kcal/mol but the *abs_e* of Ala10_set_1 and RAAE set decreased only ~0.5 kcal/mol. Thus, we used $w_i = 1$ for *abs_e* of trpzip2, 3Ai3 and HP36 sets while used $w_i = 10$ for *abs_e* of Ala10_set_1 and RAAE set. We set other weighting factors in a similar way. Weighting factors for different contributions are summarized in Table 1. Our choice of weighting factors is not unique and thus others could choose different w_i .

The search space for fitting 18 parameters is vast, thus we did not expect to locate the global minimum for our objective function. Our goal instead is to have a parameter set showing significant improvement in solvation energy and effective radii calculation relative to PB when comparing to GB-OBC and original GB-Neck models, and one that simultaneously accounts for many aspects of the training data. We used the local search method UOBYQA⁴⁴, which is an unconstrained minimization method that does not require objective function derivatives, and allows optimization with large number of variables. It took about 1 minute for the objective function calculation and we spent about 20 days (~30000 function evaluations) for each optimization. Because of the computational expense, we performed only 5 optimization runs, each run starting with an initial random guess and resulting in different final objective functions and parameter sets (Table S1). Due to the relative few number of runs, we attempted to determine if additional or independent optimization with a different approach would provide improved objective function values. We employed a parallel Genetic Algorithm (GA)⁴⁵ (code implemented by Metcalfe et al.).⁴⁶ The GA is one of the most popular global search methods⁴⁷ and, in principle, well suited for this task because it is likely that some of the parameters are weakly coupled, and thus mating of genes with independent improvements located in different parameters could be productive. GA options such as mutation and crossover rate were set to default values⁴⁶. Each optimization run had population size of 120 and the objective function was allowed to be evaluated up to 2500 generations. We performed 31 runs in total. Initial populations of most runs were randomly created. Parameters of GB-OBC, GB-Neck and previous UOBYQA results were also included in some runs as initial guesses.

Structures used for testing the new parameters in MD simulations

All of the fitting described above was performed relative to PB solvation energy calculations. This is consistent since both lack description of the hydrophobic and van der Waals components of the aqueous solvation, and thus by design our fitting did not modify the electrostatic component of GB to empirically correct for these missing terms as it would if we fit directly to reproduce data from explicit solvent simulations, which would likely lead to reduced transferability. However, we hypothesized that improved agreement with PB would also result in improved agreement with explicit water.^{16a} To test this hypothesis, we compared simulations generated with the re-optimized GB-Neck parameter set (named GB-Neck2) with those from TIP3P explicit water for several systems.

Ser-Ala-Ala-Glu Model Peptide (SAAE)—SAAE was used to compare hydrogen bond (H γ of Ser and O ϵ of Glu) PMFs between GB and TIP3P models. The potential of mean force for hydrogen bond formation was a useful independent measure of model quality, but was also performed because one of our intermediate models showed much too high propensity of forming such interactions as compared to MD in explicit water. Since the solvation energy profile matched that in PB using the same intrinsic Born radii (data not shown), we built on our previous work^{17a, 18} that showed adjusting intrinsic radii could improve fit to explicit solvent data. Additionally, the strong salt bridge interaction between side chains of Asp (or Glu) and Arg could be adjusted by modifying the radius of either the H^{N+}(Arg) or the carboxyl oxygen^{17b}. Thus, the SAAE model was built to compare the H-

bond PMFs of GB models to TIP3P model and allow adjustment the radii of carboxyl oxygen atoms independent from subsequent adjustment of H^{N+} to refine salt bridge strength.

We performed 3 REMD simulations using TIP3P, GB-OBC, GB-Neck and 2 REMD simulations for GB-Neck2 with original and modified carboxyl oxygen intrinsic radius (1.5 and 1.4 Å respectively). GB-OBC and GB-Neck runs were used as controls. SAAE was solvated in a truncated octahedron box with 8 Å buffer by using 459 TIP3P water molecules. TIP3P REMD simulation was performed for 60 ns while GB REMD simulations were extended to 50 ns. Because Glu had two symmetric $O\epsilon$ in the side chain, we used PMFs of distance between $H\gamma$ (Ser) and $C\delta$ (Glu) to define PMFs of H-bond instead of distance between $H\gamma$ (Ser) and $O\epsilon$ (1 and 2) of Glu. The choice of $C\delta$ (Glu) for PMF calculating was consistent with previous reports.^{17a, 17b, 18}

Arg-Ala-Ala-Glu Model Peptide (RAAE)—RAAE was used as test system due to its small size and the availability of TIP3P data from Okur et al.^{17b} Okur and coworkers^{17b} demonstrated that Arg salt bridge strength in RAAE was 2.5-3 kcal/mol stronger in GB-OBC than in TIP3P. Shang et al.¹⁸ later corrected this GB-OBC overestimation by reducing radii of H^{N+} (Arg) from 1.3 Å to 1.1 Å. We had initially hypothesized that simply including RAAE structures in training would help reduce salt bridge strength, and performed GB-Neck2 simulations using original mbondi2 radii. However, the stability was still significantly overestimated compared to TIP3P, and we thus performed several simulations with various combinations of modified radii to identify a value that better reproduced the PMF in TIP3P. Particularly, we performed REMD simulations of GB-Neck2 by using unmodified mbondi2 H radii and also using 4 modified radii for H^{N+} : using a radius of 1.4 Å $O\epsilon$ (Glu) with 1.3 Å, 1.2 Å, 1.17 Å or 1.1 Å for H^{N+} (Arg). 300 K trajectories from TIP3P and GB-OBC mbondi2 REMD simulations from Okur et al.,^{17b} GB-OBC 1.1 H^{N+} (Arg) REMD simulation from Shang et al.¹⁸ and GB-Neck mbondi2 REMD simulation were used for comparison with GB-Neck2 modified mbondi2 simulations. RAAE protocols and initial structures were taken from Okur et al.^{17b} We used one 40ns run for GB-Neck and GB-Neck2 simulations.

Lys-Ala-Ala-Glu Model Peptide (KAAE)—In addition to Arg, Lys can also participate in salt bridge interactions. We compared the KAAE salt bridge PMF of GB-Neck2 simulation to that from TIP3P simulation. As with RAAE, it was built in a helical backbone conformation to allow favorable salt bridge orientation.^{17b} We performed REMD simulations for TIP3P, GB-OBC, GB-Neck as controls and two simulations for GB-Neck2 (with mbondi2 and with mbondi3 (Table S4)) to see which simulation of GB-Neck2 could best reproduce TIP3P salt bridge PMF.

The RAAE protocol was adopted for KAAE. All simulations of GB models were run up to 50 ns while TIP3P simulation was run for 30ns. The distance between $N\zeta$ of Lys and $C\delta$ of Glu was defined as salt bridge distance. A large solvation buffer length was defined to minimize periodicity artifacts in the PMF^{17b}, using a truncated octahedron box with 16 Å buffer and 2433 TIP3P water molecules.

The SAAE, RAAE and KAAE peptides were built using the tleap program in Amber10 with acetylated and amidated N- and C-termini. The radii obtained from these optimizations is denoted mbondi3 (table S4).

Ala10 Model Peptide—Alanine decapeptide (Ace-Ala₁₀-NH₂) was used to compare secondary structure content (DSSP)⁴⁸ and local structural propensities between GB and TIP3P simulations following Roe et al.^{16a} To test if the improvement observed in our training would translate to better secondary structure balance in MD, we repeated Roe's protocol with GB-Neck2. DSSP and local structure propensities from GB-Neck simulation were compared with the ones from TIP3P, GB-OBC and GB-Neck models. Eight replicas were used for REMD simulations, starting from extended conformations. One 50ns REMD run was performed for GB-Neck2 and compared to GB-OBC, GB-Neck and TIP3P data from Roe et al.^{16a}

HP-1 Model Peptide—Because Ala10 structures were used in fitting GB-Neck2 parameters, we desired an independent test of the change in helical propensity. HP-1 is good candidate since it is nearly the same size as Ala10 and showed moderate α -helix content in TIP3P REMD simulation⁴⁹. HP-1 (MLSDDEFKAVFGM) is adopted from the N-terminal helix of HP36, a 36-residue helical subdomain of the villin headpiece. As with Ala10, we compared DSSP and local conformational propensities between GB simulations (GB-OBC, GB-Neck and GB-Neck2) and TIP3P. 300K trajectories from TIP3P and GB-OBC REMD simulations were taken from Wickstrom et al.⁴⁹ We performed 2 REMD simulations up to 50ns for GB-Neck and GB-Neck2. Nonhelical structures extracted from TIP3P REMD simulation were used as initial structures for REMD. Because HP-1 peptide has Lys salt bridge potential, we the optimized mbondi3 radii set (table S4) for GB-Neck2 simulation.

We further tested the robustness of GB-Neck2 by evaluating the ability to reproduce the experimental thermal stability for 2 small peptides for which experiments indicate different secondary structure motifs different from the unstructured Ala10 and helical HP-1: a hairpin structure (HP5F),²⁰ and the trp-cage tc5b mini-protein²¹ that has alpha and 3-10 helix as well as a PPII strand and a small hydrophobic core. The short length and the availability of experimental melting temperature (and melting curve for tc5b) made these two sequences ideal for testing. For each protein, we performed 2 REMD simulations starting from folded and linear structures. Simulated melting curves for those proteins were generated by calculating fraction folded (the fraction of the number of frames having native structure over the total number of frames from simulation) versus temperature. When comparing melting temperature between simulation and experiment, it is important to note that the agreement depends on not only on the solvation model but also the protein force field. As discussed above, we employed the widely used ff99SB force field, but disagreement with experiment can arise from many sources outside GB model accuracy so one must use caution when interpreting the results.

HP5F model peptide—HP5F²⁰ is a short peptide with sequence KKYTWNPATGKFTVQE.²⁰ We first simulated an extended structure in GB-Neck2 using REMD, and then extracted the representative structure for most populated cluster at 300K. This representative “folded” structure was used to initiate a second independent REMD run.

REMD simulations were run to 150 ns, 75 ns and 90 ns for GB-OBC, GB-Neck and GB-Neck2 respectively. We also performed an additional 70 ns run for GB-Neck2 with a SASA (solvent accessible surface area) based nonpolar solvation term ($gbsa = 1$ in Amber) to test the effect. An experimental atomic structure of HP5F has not been reported, but as it is expected to adopt the same fold as the GB1p peptide,^{20, 50} we used the GB1p backbone for calculating RMSD during HP5F trajectories. The GB1p structure was derived from the C-terminal hairpin of protein G (PDB ID: 3GB1,³⁵ residue 41-56). Residues 2 to 15 were chosen for RMSD to avoid the flexible termini. Structures having backbone RMSD smaller than 2.0 Å were defined as folded. We chose this cutoff based on the position of the minimum separating folded and unfolded regions in the simulated RMSD histogram (figure S8). A full experimental melting curve for HP5F was not available, thus we compared our results to the experimental melting temperature and folded population at 298K.²⁰

Trp-Cage tc5b model peptide—The tc5b²¹ variant of trp-cage is a 20 residue peptide having sequence of NLYIQWLKDGGPSSGRPPPS.²¹ The first model from the NMR structure ensemble, and a linear structure built by tleap, were used as starting structures for 2 REMD simulations for each GB model (340 ns, 240 ns, 160 ns and 72 ns for GB-OBC, GB-Neck, GB-Neck2 and GB-Neck2 with nonpolar solvation term, respectively). Different GB models have different simulation lengths since they have different time scale for convergence (having small error bars from two runs). As with HP5F, we defined folded structures by using a backbone RMSD cutoff of 2.0 Å based on the RMSD histogram (figure S9). RMSD to native tc5b was calculated for backbone atoms from residue 3 to 18 to avoid flexible termini.^{17a} We compared melting curves from GB-OBC, GB-Neck, GB-Neck2 and GB-Neck2 SASA to the ones from NMR and CD experiments.²¹

Protocols for simulations and data analysis

REMD Simulation Protocols—All simulations used to compare GB simulations to TIP3P simulations and experiments presented in Results were carried out with AMBER 10⁵ and the ff99SB force field.¹⁹ The AMBER 10 code was modified to support GB-Neck2; it is now available in AMBER version 11 or later by specifying $igb = 8$. All simulations used REMD⁵¹ for enhancing sampling. The time step was 2 fs for all REMD simulations. SHAKE⁵² was used for constraining all bonds to hydrogen. For small protein/peptide simulations, we did not employ a surface-area based nonpolar solvation term. Temperature was controlled by using Berendsen thermostat⁵³ in TIP3P²³ simulations with a time constant of 1.0 ps⁻¹, or by using a Langevin thermostat in GB simulations with a collision frequency of 1.0 ps⁻¹. Unless noted, GB-Neck simulation used mbondi2 radii,¹² GB-OBC simulations used mbondi2 with 1.1 H^{N+} (Arg)¹⁸ while GB-Neck2 simulation used mbondi3 (Table S4). Further details of each simulation are given in the context. In explicit solvent simulations, peptide models were solvated with TIP3P²³ water in a truncated octahedron box. PME⁵⁴ was used for treating long range electrostatic interactions and nonbonded interaction cutoff was 8 Å. No cutoff was used in GB simulations.

Exchanges in REMD simulations were attempted every 1 ps. 32 replicas were used for TIP3P REMD while only 8 replicas were needed for Ala10, HP-1, tc5b, HP5F GB REMD and 6 replicas were used for GB REMD of RAAE, SAAE, KAAE. Temperature

distributions were chosen to give 15-25% exchange success, with actual temperatures reported in Table S5. The TIP3P REMD simulation protocol was adopted from Okur et al.^{17b} For REMD simulations of RAAE, SAAE, KAAE model peptides, backbone atoms were restrained with weak positional restraints (1.0 kcal/mol*Å) to the starting helical conformation, as discussed in Okur et al.^{17b} There were two runs for tc5b and HP5F REMD simulations, starting from extended and folded conformations. We discarded first 25 ns of tc5b trajectories and 40 ns of HP5F trajectories to avoid initial structure bias. The error bars in these 2 cases were calculated from two runs. For other REMD simulations (Ala10, HP-1, RAAE) only 1 run was performed since the convergence time under these conditions had already been reported.^{16a, 17b, 49} In the case of SAAE and KAAE, we assumed that converged simulation time for side chain sampling should be comparable to that reported for RAAE^{17b}. For Ala10, HP-1, SAAE, RAAE and KAAE REMD simulations, the first 10 ns of each run was discarded and error bars were estimated from first and second half of data.

Data analysis—PMFs were calculated based on the assumption of Boltzmann-weighted populations. Data were extracted from histograms of RMSD or distance, using $\Delta G = -RT \ln(N_i/N_0)$ where N_0 was the population of the most populated bin and N_i was the population of i^{th} bin. Calculation of RMSD, DSSP⁴⁸ and ϕ/ψ values were done using the ptraj program in Amber10. For proteins taken from the Protein Data Bank, all ligands, water molecules and ions were removed and missing hydrogen atoms were added by tleap program in Amber10. Local secondary structure assignment for Ala10 was previously defined by Roe et al.^{16a} based on ϕ/ψ angle values (alpha (-70°/-25°), left (50°/30°), PP2 (-70°/150°), or extended (-150°/155°)). We retained this definition for HP-1 to be consistent with Ala10.

Cluster analysis—Cluster analysis was done by the Moil-View program⁵⁵ following the protocol described by Okur et al.^{17b} We used a similarity cutoff of 2.5 Å for all backbone atoms of Ala10, trpzip2, and 3Ai3 and HP5F trajectories.

Results and Discussion

Parameter fitting

The 18 parameters were refit to minimize the objective function. As stated in Methods, we performed 5 runs for UOBYQA in which each run started from initial random parameters and each UOBYQA run converged at different local minima (Table S1). We then performed 31 runs for parallel GA in which most of runs started with random parameters while some runs started by including in initial population GB-OBC parameters, GB-Neck parameters or parameter sets from UOBYQA runs. Attempting to vary GA parameters such as population size and mutation rate were not successful in producing better objective function values than those from UOBYQA. Figure S10 shows the objective function evolution during optimization.

Objective function values for GB-OBC, GB-Neck and GB-Neck2 are provided in Table 1. The best objective function was achieved through UOBYQA optimized parameters; these are denoted hereafter as GB-Neck2 (Table 2). The objective function of GB-Neck2 is 274, which is much smaller than 381 and 445 for GB-OBC and GB-Neck models, respectively.

It should be noted that although the scaling parameters S_X were initially introduced to correct for overlap of van der Waals spheres and so might be expected to remain less than or equal to 1.0, there is no formal reason that they cannot be greater than 1.0, as pointed out by Hawkins et al.¹⁰ Indeed, it can be argued that since the purpose of the majority of the parameters introduced into the GB formalism is to allow a better fit to higher levels of theory, the overall agreement of the model is more important than assigning a physical meaning to the parameters. When the S_X values are considered free parameters it allows them to correct for other errors in the model, such as those introduced by the CFA.

Comparison with PB solvation energies and effective radii

Results on training sets—Table 1 shows the contributions to the objective function of the absolute solvation energies, relative solvation energies, and effective Born radii. For small systems like Ala10 or RAAE where most atoms are solvent-exposed, these pairwise GB models perform reasonably well,^{16a} and only modest improvement in these metrics is obtained with refitting. Particularly, *abs_e* and *rel_e* to PB calculation of GB-Neck2 for Ala10_set1 are 0.8 and 1.0 kcal/mol, compared to 1.1 and 1.6 kcal/mol of GB-OBC or 2.7 and 2.3 kcal/mol of GB-Neck, respectively. In larger systems like trpzip2, 3Ai3 or HP36, more substantial improvement was seen in GB-Neck2. For example, *abs_e* for 3Ai3 was reduced from 10.1 (GB-Neck) or 7.6 (GB-OBC) to 3.7 kcal/mol (GB-Neck2). For a given molecule, obtaining more accurate GB absolute solvation energies was easier than relative solvation energy. HP36, for instance, has *abs_e* of 4.3 kcal/mol (GB-Neck2), which is 24.1 kcal/mol lower than *abs_e* of GB-Neck (85.2 % reduction in error), while *rel_e* of GB-Neck2 for this training set is only improved by 0.3 kcal/mol (5.9 % reduction). This result suggests that refitting leads to improvements in systematic error of GB-Neck across all conformations (see Figure 1). GB-Neck2 also has better agreement with PB in calculating effective Born radii. The GB-Neck2 *eff_rad_rmsd* to ‘perfect’ radii of Ala10 (0.10 Å) was smaller than GB-OBC and GB-Neck (0.16 Å and 0.19 Å respectively). Once again, the improvement is more significant for larger systems. The *eff_rad_rmsd* for the protein HP1113 set is 1.47 Å for GB-Neck2 as compared to 1.82 Å for GB-OBC or 2.27 Å for GB-Neck.

Overall, there is improvement of absolute solvation energy, relative solvation energy and effective radii calculation for GB-Neck2 for all training sets as compared to GB-OBC and original GB-Neck model.

Results on test sets: In this section, we employ larger test sets to gauge the transferability of the new parameters. As stated in Methods, we designed two test categories: type I and II. Type I had a peptide / protein system that was used in training, but with many more conformations, while test set type II had entirely different molecules than those in the training sets.

Comparison with PB solvation energy: The *abs_e* and *rel_e* to PB data are presented in table 3. The trend observed for the type I test sets is consistent with results for the training data, indicating that the structure variation was sufficient in the training data to permit application to more structure variety while retaining the improvement compared to the older

GB models. For instance, *abs_e* for Ala10 test set from GB-OBC, GB-Neck and GB-Neck2 are 1.1, 2.2 and 1.0 kcal/mol respectively while *rel_e* are 0.7, 0.7 and 0.5 kcal/mol for GB-OBC, GB-Neck and GB-Neck2. For more complex molecules, the test set results closely match those from the training set: the *abs_e*, for example, of trpzip2 test set from GB-OBC, GB-Neck and GB-Neck2 are 9.2, 8.4 and 3.2 kcal/mol which are close to 9.5, 8.3, 2.8 kcal/mol for trpzip2 training set, respectively.

Results for type II test sets (table 3) indicate that the improvements are transferable to independent systems, with lower *abs_e* and *rel_e* for GB-Neck2 as compared to GB-OBC and GB-Neck. There is little improvement for very small proteins like tc5b and DPDP. However, larger proteins show quite dramatic improvement. For example, *abs_e* of GB-Neck2 for the AIDS drug target HIV1-PR was 17.2 kcal/mol, eliminating 85% - 87% of the error as compared to GB-OBC (115.0 kcal/mol) or GB-Neck (133.1 kcal/mol). Additionally, *rel_e* of GB-Neck2 for HIV1-PR was 16.8 kcal/mol, significantly improved as compared to 20.1 kcal/mol error with GB-OBC and GB-Neck. Since relative energies control the equilibrium populations, this improvement would be expected to have a significant impact on the ensemble sampled in MD simulations.

Comparison with PB ‘perfect’ radii: In order to test the transferability in improvement of effective Born radii from training to testing stage, we randomly extracted 100 tc5b structures having backbone RMSD to native structure smaller than 2.5 Å to compare effective radii from GB to PB. Since calculating PB ‘perfect’ radii for large proteins is computationally expensive,⁹ we chose tc5b as a system large enough to have buried atoms and small enough to be computationally tractable. In addition, native-like structures were chosen to have a wide range of effective radii from atoms in molecule's surface to deeply buried atoms. The inverse of effective radii is used in calculating forces, thus it makes sense to compare inverse effective radii.¹³ The RMSD between GB and PB inverse effective radii for GB-OBC, GB-Neck and GB-Neck2 are 0.068, 0.052 and 0.054 respectively. GB-Neck2 and GB-Neck have nearly the same RMSD and the performance of these models is somewhat better than GB-OBC.

Figure 1 shows 2D histograms for the TC5b set of inverse effective radii from GB models compared to inverse of ‘perfect’ radii derived from PB. The effective radii of buried atoms (lower left region) were still underestimated in GB-OBC while GB-Neck and GB-Neck2 had less degree of underestimation. However all three models seemed to overestimate effective radii of atoms near surface of the molecules (upper right region). GB-Neck2 is somewhat improved for effective radii calculation of atoms in the middle region of the plot in which the most populated bins lie close to the diagonal. For atoms in this region, GB-OBC and GB-Neck tend to overestimate effective radii, perhaps leading to the dramatic improvement in systematic error with GB-Neck2 seen in Tables 1 and 3.

In summary, the results from the test sets confirm the improvement from the new parameters is transferable from a set of structures used for training to different set of structures not used in training as well as to entirely different proteins.

Comparison with explicit water MD: hydrogen bonds, salt bridges and secondary structure—A particular goal of this work was to reduce the errors in secondary structure bias and salt bridge strength previously reported for GB models compared to results from explicit water models such as TIP3P. The results presented in previous sections showed significant improvement of GB-Neck2 model in calculating solvation energy and effective Born radii when using PB as a benchmark. It is of interest to determine if better match to PB also results in improved correspondence of GB with reference simulations in explicit water, which tend to be much more computationally demanding. We hypothesized that fitting to the more accurate PB continuum water model could help improve agreement between GB and explicit water.

We next tested GB-Neck2 to see if this improvement could translate to improvement in balancing secondary structure populations and improving salt bridge strength. The comparison between GB and TIP3P simulations, however, does not solely depend on the GB model. Firstly, performance of a GB model is heavily affected by the intrinsic Born radius set that defines the boundary between solute and solvent. Secondly, GB only calculates the polar part of solvation free energy and simulation results also depend on the accuracy of non-polar solvation contributions, such as cavity formation and van der Waals interactions with solvent. Since the simple SASA-based non-polar solvation approach currently available in Amber has known limitations,^{4a-c} the main focus of this work is improved agreement in polar solvation free energy. However, in order to roughly estimate the affect that including this term has on results, simulations were performed with and without the commonly used surface area based non-polar term for the HP5F and tc5b systems

Strength of hydrogen bonds and salt bridges—The salt bridge is formed by oppositely charged side chains of Arg (or Lys) and Glu (or Asp). Conventionally, the ion pair interaction could be adjusted by changing the radii of H^{N+} of Arg. For example, Geney et al.^{17a} and Shang et al.¹⁸ empirically decreased the radius of H^{N+} of Arg from 1.3 Å to 1.1 Å to match the salt bridge PMF from GB to that from TIP3P simulation. However, we recognize that the radii (and hence desolvation penalty) of the carboxyl oxygen atoms can also be modified to change the balance of desolvation and Coulombic contributions. One approach to determine which group to adjust is to examine carboxyl interactions in the absence of a positively charged partner. We first chose a simple peptide system, SAAE, to investigate H-bond strength between the ionized side chain of Glu with the side chain of Ser. Original mbondi2 radii¹² were used for all GB models. Figure 2 shows the distance PMFs of $H\gamma$ (SER) and $C\delta$ (Glu) for TIP3P, GB-OBC, GB-Neck, and GB-Neck2 models. The H-bond is thermodynamically unstable in all cases, meaning that the H-bonded distance is a local and not the global free energy minimum. All of these GB models fail to reproduce the solvent-separated local minimum near 5 Å; such behavior is expected for continuum models. The H-bond in GB-Neck2 is 0.7 kcal/mol stronger than in TIP3P, while GB-Neck and GB-OBC H-bond strength are comparable to TIP3P. We empirically decreased the carboxyl oxygen radii from 1.5 Å to 1.4 Å to reproduce the profile obtained in TIP3P. The modified carboxyl oxygen radii should be applied to charged carboxyl groups in Asp and Glu sidechains as well as C-terminal residues.

Having carboxyl oxygen radii of Glu sidechain adjusted, we next investigated the salt bridge formed by Arg and Glu. We originally included a set of RAAE structures in our training set (Table 1) for explicitly training the salt bridge. The fitting resulted in modest improvement in solvation energy calculation for these ion pairs as compared to GB-OBC and GB-Neck. We therefore expected improved agreement between salt bridge PMFs from GB-Neck2 and TIP3P REMD simulations for this RAAE system. Salt bridge PMFs from REMD runs for different GB models and TIP3P are shown in Figure 3A, with variation of the Arg H^{N+} Born radii in the mbondi2 set; PMFs for GB-OBC, GB-Neck, GB-OBC 1.1 Å H^{N+} , GB-Neck2 with 1.4 Å $O\epsilon$ and GB-Neck2 with 1.4 Å $O\epsilon$ + 1.17 Å H^{N+} are shown. All GB profiles have a global minimum slightly shifted from the one in TIP3P due to the difference in salt bridge geometry between GB and TIP3P, discussed in more detail by Okur et al.^{17b} With standard mbondi2 radii (1.5 Å $O\epsilon$ of Glu and 1.3 Å H^{N+} of Arg) the PMF indicates salt bridges from GB-Neck2 simulation are ~3.5 kcal/mol stronger than in TIP3P, significantly worse than the 1.0 kcal/mol and 2.0 kcal/mol stronger with GB-OBC and GB-Neck, respectively. This implies that fitting to PB solvation energies did not help improve salt bridge profile (RMSD between GB and PB absolute energies for RAAE test set (table 3) is 1.4 kcal/mol). We thus hypothesize that PB with mbondi2 radii may also have too strong salt bridge compared to TIP3P, as indicated by Shang et al.¹⁸ With new carboxyl oxygen radii (1.4 Å) fit to SAAE PMFs and standard H^{N+} radii (1.3 Å), the salt bridge with GB-Neck2 is still ~2.0 kcal/mol stronger than in TIP3P (Figure S11). Thus, radii of H^{N+} of Arg were empirically reduced from 1.3 Å to 1.17 Å to match the TIP3P PMF curve (Figure 3). The PMF from GB-Neck2 with 1.17 Å H^{N+} (Arg) also matches well to that from GB-OBC with modified H^{N+} radii as reported in Shang et al.¹⁸, suggesting that modification of this radius is a general way to improve salt bridges in GB models. The physical justification for adjusting these radii is discussed in detail by Geney et al.^{17a}

We next addressed whether primary amines (Lys and N-term) needed comparable corrections to Arg. Figure 3B shows the PMFs for KAAE from GB-OBC, GB-Neck and GB-Neck2 (all with mbondi2) and GB-Neck2 with mbondi3 radii. As discussed above, none of the GB models reproduce the solvent-separated minimum seen with explicit water. In GB-Neck2 with mbondi2 radii, the salt bridge was ~1.0 kcal/mol stronger than TIP3P while the GB-OBC salt bridge was ~0.5 kcal/mol stronger. In contrast, the salt bridge with GB-Neck mbondi2 was ~0.5 kcal/mol weaker than in TIP3P. GB-Neck2 with mbondi2 and modified carboxyl oxygen showed near-quantitative match to TIP3P PMF, suggesting that our carboxyl changes were sufficient and no adjustment of radii is needed for H^{N+} of Lys side chain or N-terminal amines.

The new radii set with modified carboxyl oxygen and Arg H^{N+} is denoted mbondi3 (Table S4). Overall, mbondi3 appears to be the best radii set for use with GB-Neck2 in reproducing TIP3P PMFs of salt bridge interactions. In the remainder of this work manuscript, all simulations of GB-Neck2 used mbondi3 intrinsic Born radii unless noted otherwise.

Evaluating α -helical bias

Ala10 Model Peptide: Roe et al.^{16a} showed that the ability of a GB model to reproduce PB solvation energies for Ala10 was well correlated with the extent of helical bias obtained in

simulations compared to TIP3P simulation. We therefore hypothesize that our new GB model, with better agreement to PB, should also better reproduce secondary structure preferences as compared to TIP3P. Roe et al.^{16a} quantified the accuracy by comparing DSSP and local conformational propensity between GB and TIP3P simulations. We repeated these analyses for our GB-Neck2 model, using GB-OBC and GB-Neck results as controls (Figure 4, with numerical data provided in Table S6).

GB-Neck2 has reduced alpha and turn content as compared to GB-OBC (4.4 % vs. 10.1 % in OBC for alpha content; 16.2 % vs. 25.5 % in OBC for turn content). However, the original GB-Neck still has somewhat better agreement to TIP3P data (1.4 % vs. 2.5 % in TIP3P for alpha content; 4.6 % vs. 2.9 % in TIP3P for turn content). Although 3-10 helix content was reduced for GB-Neck2, the population is still somewhat too large compared (9.3 % in GB-Neck2 and 12.7 % in OBC vs. 2.9 % in TIP3P). GB-Neck2 also has higher preference for residue to sample the helical region of the Ramachandran map (30.3 % in GB-Neck2 and 22.6 % in GB-Neck vs. 6.2 % in TIP3P). Although GB-Neck2 better reproduced absolute and relative solvation energies for Ala10 training set and Ala10 test set than GB-Neck, this improvement seems not to transfer to better agreement with TIP3P simulation. There might be several reasons for this. First, the mbondi2 radii set (mbondi3 is the same as mbondi2 for systems that do not have Arg, Glu, Asp or charged C-termini) was not specifically optimized for use with GB-Neck, and the improved agreement to TIP3P for this combination may be fortuitous cancellation of error. This same cancellation of error may make the performance of GB-Neck better than PB in this particular case; however it is difficult to get converged REMD data when using PB solvation, and such calculations are out of the scope of the present work. In addition, the small improvement in energy compared to PB may not be enough to improve structure results compared to TIP3P simulation for this system. This seems reasonable since we have seen significant improvement for larger systems like HP5F or tc5b, which will be shown below.

HP-1 Model Peptide: Because Ala10 structures were used in training GB-Neck2, we repeated the same analyses as we did for Ala10 but for a different peptide system (HP-1) to confirm the results in balancing secondary structure (Figure 5). Furthermore, unlike Ala10, HP-1 is known to adopt modest helical content in solution.⁴⁹ Similar trends to Ala10 were observed in DSSP data and local alpha content (Table S6). Particularly, the alpha content from GB-Neck was slightly smaller than TIP3P content while GB-Neck2 alpha content was somewhat larger (23.8 % in GB-Neck2 vs. 18.9 % in GB-Neck vs. 21.6 % in TIP3P). GB-OBC had too much alpha content (43.9 %). Although all GB models had close average turn content compared to TIP3P, GB-Neck and GB-Neck2 had better agreement as indicated by DSSP. This shows that performance of GB-Neck and GB-Neck2 on alpha content is somewhat system dependent, likely due to the role of side chain interactions in helix formation of HP-1.⁴⁹ However, the trend remains that GB-Neck tends to destabilize alpha conformations, as demonstrated above and as previously reported by Dill et al.^{16e} and Roe et al.^{16a} Overall, the good performance of GB-Neck2 in balancing secondary structure can be transferred from training system (Ala10) to testing system (HP-1).

Folding of HP5F and tc5b: Comparison with experimental melting temperature: The above GB simulation results were compared to TIP3P simulations using the same protein force field. However, one of the main purposes of improving a GB model is to get closer agreement between computational and experimental data, particularly for simulations that are currently difficult or intractable in explicit water. However, such comparisons are more complex than the comparison between GB and TIP3P simulation because they also depend on the protein force field used. Deviations from experiment may not be a result of weakness in the GB part of the model, and accurate reproduction of experimental data could arise from fortuitous cancellation of error and may not provide proof of an accurate solvent model. Nonetheless, the comparison to experiment provides a useful measure of the quality one might expect from this particular combination. For the purpose of this testing, we used the combination of the GB-Neck2, mbondi3 radii and the ff99SB force field.¹⁹ This widely adopted force field was used since it has been shown to well balance secondary structure.^{19, 22c, 56}

We compared equilibrium thermal stability between different GB models and experiment (NMR or CD) for HP5F²⁰ and tc5b,²¹ which adopt different structure motifs (hairpin and helix-turn-PPII).^{17a, 57} Simulations with GB-Neck2 were also repeated including a SASA-based nonpolar solvation term in order to ascertain its impact on results.

Figure 6A shows the simulated melting curves for HP5F for GB-OBC, GB-Neck, GB-Neck2 and GB-Neck2 SASA models compared to experiment data. The melting temperature and fold population of GB-Neck2 at 298 K (317K and 74%, respectively) are in excellent agreement with experimentally determined values (326 K and 82%).²⁰ For the tc5b mini-protein, the melting curves for GB-Neck2 and NMR and CD experiments²¹ are shown in **figure 6B**. GB-Neck2 predicts a melting temperature of 302 K, which is again close to the experimental value of 315 K²¹ and to the reported value of 321 K from TIP3P REMD simulation⁵⁸ with ff99SB force field. The excellent agreement between GB-Neck2 simulation and experiment is promising since several groups reported significantly elevated simulated melting temperatures for tc5b.⁵⁹ Pitera et al.^{59a} reported a melting temperature of ~400K from REMD simulation of GB-HCT model + ff94 force field. Zhou et al.^{59b} also obtained a melting temperature above 400K when using TIP3P model + OPLS-AA force field. Compared to GB-Neck2 simulations, GB-OBC and GB-Neck significantly underestimate melting temperature for both testing systems (GB-OBC: ~307 K and ~264 K; GB-Neck: <275K and ~290K for HP5F and tc5b respectively). GB-Neck especially destabilizes the native hairpin even at very low temperature.

GB-Neck2 runs with and without the non-polar term both produce reasonable estimations of melting points for HP5F and tc5b (317 K and 335 K for HP5F; 302 K and 324 K for tc5b for simulations with and without nonpolar term respectively). Inclusion of the SASA-based non-polar term provides small increases in stability but does not dramatically impact the results for these systems. It is likely that use of a better non-polar model (such as that in AGBNP2⁶⁰) could improve results even further, however that is beyond the scope of the current work, which focuses on the polar component of solvation.

Conclusion

Pairwise GB solvation models remain desirable due to their high computational efficiency, but many weaknesses have been reported. We propose a new parameter set for the GB-Neck model, obtained by making several key parameters that relate to interstitial cavities dependent on chemical element. Adding more parameters called for use of a much larger training set than employed in the past, therefore we developed conformation libraries containing thousands of structures for peptide and protein sequences of various lengths and structure propensities. Our objective function for training included absolute and relative solvation free energies compared to PB, as well as accuracy of effective Born radii of the atoms. While our fitting significantly improved the model compared to previous ones, it is possible that even more extensive fitting of the same training data could result in further improvement of the model. This can be re-visited when such studies become more feasible. Since we optimized the parameters to improve the overall performance rather than focus on physical meaning of individual parameters, it is possible that some cancellation of error exists in our model, potentially limiting transferability to systems unlike those studied in this work.

Final empirical adjustments were made to some of the intrinsic radii to improve agreement with explicit solvent simulations. These modifications help GB-Neck reproduce the H-bond and salt bridge PMFs of TIP3P simulations. The new GB-Neck2 model not only shows better results for the training systems, but for a variety of tests systems that measure solvation free energy, secondary structure propensity and even thermal stability profiles compared to experimental data. Thus the combination of GB-Neck2 model, radii set, force field used here is recommended for future study of peptide or protein simulations.

Our GB-Neck2 model shows significant improvement in solvation energy and effective radii calculation as compared to GB-OBC and GB-Neck. This model, however, is still based on the CFA integral calculation which has been shown to overestimate effective radii,⁶¹ compared to much slower numerical models such as GBMV⁶² or GB-R6⁶¹ using non-CFA integrals. Through parameter fitting, our approach thus has attempted to empirically compensate for the CFA as much as possible. Onufriev et al.⁶³ recently developed an analytical form of GB-R6 (named AR6) but the resulting accuracy was substantially decreased from the numerical form (NR6), and performed worse than GB-Neck2 on our training and test sets.⁶⁴ We believe that our strategy in fitting parameters, as well as use of the training and test sets we have developed, could help to improve the performance of AR6 and future solvation models.

Our results also show that despite not including a nonpolar term, GBNeck2 is still able to improve agreement to TIP3P as well as experiment, and it is likely that further improvement will be seen with the addition of a more accurate term for nonpolar solvation free energy.

Future work will include optimization of additional parameters for nucleic acid simulations.⁶⁵

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was funded by NIH grant (GM R01 079383). This research utilized resources at the New York Center for Computational Sciences at Stony Brook University/Brookhaven National Laboratory which is supported by the U.S. Department of Energy under Contract No. DEAC02-98CH10886 and by the State of New York. This research was also supported by an allocation of advanced computing resources provided by the National Science Foundation. The computations were performed on Kraken at the National Institute for Computational Sciences (<http://www.nics.tennessee.edu/>). We thank Alexey Onufriev for stimulating discussion about GB models in general as well as the GB R6 model. HN thanks Yi Shang, Fangyu Ding, Lauren Wickstrom, Asim Okur, and Kun Song for providing peptide/protein trajectories in which references are noted. HN also thanks Yi Shang and Christina Bergonzo for critical reading of the manuscript.

References

1. Feig M, Brooks CL. Recent advances in the development and application of implicit solvent models in biomolecule simulations. *Curr. Opin. Struc. Biol.* 2004; 14(2):217–224.
2. Wang W, Donini O, Reyes CM, Kollman PA. BIOMOLECULAR SIMULATIONS: Recent Developments in Force Fields, Simulations of Enzyme Catalysis, Protein-Ligand, Protein-Protein, and Protein-Nucleic Acid Noncovalent Interactions. *Annu. Rev. Bioph. Biom.* 2001; 30(1):211–243.
3. Zagrovic B, Pande V. Solvent viscosity dependence of the folding rate of a small protein: Distributed computing study. *J. Comput. Chem.* 2003; 24(12):1432–1436. [PubMed: 12868108]
4. a Chen J; III, C. L. B. Implicit modeling of nonpolar solvation for simulating protein folding and conformational transitions. *Phys. Chem. Chem. Phys.* 2008; 10(4):471–481. [PubMed: 18183310] b Levy RM, Zhang LY, Gallicchio E, Felts AK. On the Nonpolar Hydration Free Energy of Proteins: Surface Area and Continuum Solvent Models for the Solute–Solvent Interaction Energy. *J. Am. Chem. Soc.* 2003; 125(31):9523–9530. [PubMed: 12889983] c Wagoner JA, Baker NA. Assessing implicit models for nonpolar mean solvation forces: The importance of dispersion and volume terms. *Proc. Natl. Acad. Sci. USA.* 2006; 103(22):8331–8336. [PubMed: 16709675] d Chen J, Brooks CL. Critical Importance of Length-Scale Dependence in Implicit Modeling of Hydrophobic Interactions. *J. Am. Chem. Soc.* 2007; 129(9):2444–2445. [PubMed: 17288425]
5. Case DA, Darden TA, Cheatham TE, Simmerling CL, Wang J, Duke RE, Luo R, Crowley M, Walker RC, Zhang W, Merz KM, Wang B, Hayik S, Roitberg A, Seabra G, Kolossvary I, Wong KF, Paesani F, Vanicek J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Yang L, Tan C, Mongan J, Hornak V, Cui G, Mathews DH, Seetin MG, Sagui C, Babin V, Kollman PA. AMBER 10. 2008
6. Gilson MK, Davis ME, Luty BA, McCammon JA. Computation of electrostatic forces on solvated molecules using the Poisson-Boltzmann equation. *J. Phys. Chem-us.* 1993; 97(14):3591–3600.
7. Feig M, Onufriev A, Lee MS, Im W, Case DA, Charles L, Brooks I. Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures. *J. Comput. Chem.* 2004; 25(2):265–284. [PubMed: 14648625]
8. Still WC, Tempczyk A, Hawley RC, Hendrickson T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* 1990; 112(16):6127–6129.
9. Onufriev A, Case DA, Bashford D. Effective Born radii in the generalized Born approximation: The importance of being perfect. *J. Comput. Chem.* 2002; 23(14):1297–1304. [PubMed: 12214312]
10. Hawkins GD, Cramer CJ, Truhlar DG. Pairwise solute descreening of solute charges from a dielectric medium. *Chem. Phys. Lett.* 1995; 246(1-2):122–129.
11. Onufriev A, Bashford D, Case DA. Modification of the Generalized Born Model Suitable for Macromolecules. *J. Phys. Chem. B.* 2000; 104(15):3712–3720.
12. Onufriev A, Bashford D, Case DA. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins: Struct., Funct., Bioinf.* 2004; 55(2): 383–394.

13. Mongan J, Simmerling C, McCammon JA, Case DA, Onufriev A. Generalized Born Model with a Simple, Robust Molecular Volume Correction. *J. Chem. Theory Comput.* 2007; 3(1):156–169. [PubMed: 21072141]
14. Tsui V, Case DA. Molecular Dynamics Simulations of Nucleic Acids with a Generalized Born Solvation Model. *J. Am. Chem. Soc.* 2000; 122(11):2489–2498.
15. Friedrichs MS, Eastman P, Vaidyanathan V, Houston M, Legrand S, Beberg AL, Ensign DL, Bruns CM, Pande VS. Accelerating molecular dynamic simulation on graphics processing units. *J. Comput. Chem.* 2009; 30(6):864–872. [PubMed: 19191337]
16. a Roe DR, Okur A, Wickstrom L, Hornak V, Simmerling C. Secondary Structure Bias in Generalized Born Solvent Models: Comparison of Conformational Ensembles and Free Energy of Solvent Polarization from Explicit and Implicit Solvation. *J. Phys. Chem. B.* 2007; 111(7):1846–1857. [PubMed: 17256983] b Zhou R. Free energy landscape of protein folding in water: Explicit vs. implicit solvent. *Proteins: Struct., Funct., Bioinf.* 2003; 53(2):148–161. c Nymeyer H, Garcia AE. Simulation of the folding equilibrium of alpha-helical peptides: A comparison of the generalized Born approximation with explicit solvent. *Proc. Natl. Acad. Sci. USA.* 2003; 100(24):13934–13939. [PubMed: 14617775] d Zhu J, Alexov E, Honig B. Comparative Study of Generalized Born Models: Born Radii and Peptide Folding. *J. Phys. Chem. B.* 2005; 109(7):3008–3022. [PubMed: 16851315] e Shell MS, Ritterson R, Dill KA. A Test on Peptide Stability of AMBER Force Fields with Implicit Solvation. *J. Phys. Chem. B.* 2008; 112(22):6878–6886. [PubMed: 18471007]
17. a Geney R, Layten M, Gomperts R, Hornak V, Simmerling C. Investigation of Salt Bridge Stability in a Generalized Born Solvent Model. *J. Chem. Theory Comput.* 2006; 2(1):115–127. b Okur A, Wickstrom L, Simmerling C. Evaluation of Salt Bridge Structure and Energetics in Peptides Using Explicit, Implicit, and Hybrid Solvation Models. *J. Chem. Theory Comput.* 2008; 4(3):488–498. c Ruhong Z, Bruce JB. Can a continuum solvent model reproduce the free energy landscape of a beta-hairpin folding in water? *Proc. Natl. Acad. Sci. USA.* 2002; 99(20):12777–12782. [PubMed: 12242327]
18. Shang Y, Nguyen H, Wickstrom L, Okur A, Simmerling C. Improving the description of salt bridge strength and geometry in a Generalized Born model. *J. Mol. Graphics Modell.* 2011; 29(5):676–684.
19. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* 2006; 65(3):712–725.
20. Fesinmeyer RM, Hudson FM, Andersen NH. Enhanced Hairpin Stability through Loop Design: The Case of the Protein G B1 Domain Hairpin. *J. Am. Chem. Soc.* 2004; 126(23):7238–7243. [PubMed: 15186161]
21. Neidigh JW, Fesinmeyer RM, Andersen NH. Designing a 20-residue protein. *Nat. Struct. Mol. Biol.* 2002; 9(6):425–430.
22. a Fadrná E, Špačková N. a. Sarzyńska J, Koča J, Orozco M, Cheatham TE, Kulinski T, Šponer J. i. Single Stranded Loops of Quadruplex DNA As Key Benchmark for Testing Nucleic Acids Force Fields. *J. Chem. Theory Comput.* 2009; 5(9):2514–2530. b Showalter SA, Brüschweiler R. Validation of Molecular Dynamics Simulations of Biomolecules Using NMR Spin Relaxation as Benchmarks: Application to the AMBER99SB Force Field. *J. Chem. Theory Comput.* 2007; 3(3):961–975. c Showalter SA, Brüschweiler R. Quantitative Molecular Ensemble Interpretation of NMR Dipolar Couplings without Restraints. *J. Am. Chem. Soc.* 2007; 129(14):4158–4159. [PubMed: 17367145] d Lange OF, van der Spoel D, de Groot BL. Scrutinizing Molecular Mechanics Force Fields on the Submicrosecond Timescale with NMR Data. *Biophys. J.* 2010; 99(2):647–655. [PubMed: 20643085]
23. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 1983; 79(2):926–935.
24. Cochran AG, Skelton NJ, Starovasnik MA. Tryptophan zippers: Stable, monomeric β -hairpins. *Proc. Natl. Acad. Sci. USA.* 2001; 98(10):5578–5583. [PubMed: 11331745]
25. Kun, Song; Stewart, JM.; Fesinmeyer, RM.; Andersen, NH.; Simmerling, C. Structural insights for designed alanine-rich helices: Comparing NMR helicity measures and conformational ensembles from molecular dynamics simulation. *Biopolymers.* 2008; 89(9):747–760. [PubMed: 18428207]

26. Okur A, Strockbine B, Hornak V, Simmerling C. Using PC clusters to evaluate the transferability of molecular mechanics force fields for proteins. *J. Comput. Chem.* 2003; 24(1):21–31. [PubMed: 12483672]
27. McKnight CJ, Matsudaira PT, Kim PS. NMR structure of the 35-residue villin headpiece subdomain. *Nat. Struct. Mol. Biol.* 1997; 4(3):180–184.
28. Wickstrom L, Bi Y, Hornak V, Raleigh DP, Simmerling C. Reconciling the Solution and X-ray Structures of the Villin Headpiece Helical Subdomain: Molecular Dynamics Simulations and Double Mutant Cycles Reveal a Stabilizing Cation Interaction. *Biochemistry.* 2007; 46(12):3624–3634. [PubMed: 17338549]
29. Chiu TK, Kubelka J, Herbst-Irmer R, Eaton WA, Hofrichter J, Davies DR. High-resolution x-ray crystal structures of the villin headpiece subdomain, an ultrafast folding protein. *Proc. Natl. Acad. Sci. USA.* 2005; 102(21):7517–7522. [PubMed: 15894611]
30. Prabu-Jeyabalan M, Nalivaika EA, King NM, Schiffer CA. Structural Basis for Coevolution of a Human Immunodeficiency Virus Type 1 Nucleocapsid-p1 Cleavage Site with a V82A Drug-Resistant Mutation in Viral Protease. *J. Virol.* 2004; 78(22):12446–12454. [PubMed: 15507631]
31. Gouda H, Torigoe H, Saito A, Sato M, Arata Y, Shimada I. Three-dimensional solution structure of the B domain of staphylococcal protein A: comparisons of the solution and crystal structures. *Biochemistry.* 1992; 31(40):9665–9672. [PubMed: 1390743]
32. Vijay-Kumar S, Bugg CE, Cook WJ. Structure of ubiquitin refined at 1.8 Å resolution. *J. Mol. Biol.* 1987; 194(3):531–544. [PubMed: 3041007]
33. Hodsdon ME, Cistola DP. Ligand Binding Alters the Backbone Mobility of Intestinal Fatty Acid-Binding Protein as Monitored by ¹⁵N NMR Relaxation and ¹H Exchange†. *Biochemistry.* 1997; 36(8):2278–2290. [PubMed: 9047330]
34. Holt D, Luengo J, Yamashita D, Oh H, Konialian A, Yen H, Rozamus L, Brandt M, Bossard M. Design, synthesis, and kinetic evaluation of high-affinity FKBP ligands and the X-ray crystal structures of their complexes with FKBP12. *J. Am. Chem. Soc.* 1993; 115(22):9925–9938.
35. Kuszewski J, Gronenborn AM, Clore GM. Improving the Packing and Accuracy of NMR Structures with a Pseudopotential for the Radius of Gyration. *J. Am. Chem. Soc.* 1999; 121(10):2337–2338.
36. Schenck HL, Gellman SH. Use of a Designed Triple-Stranded Antiparallel β-Sheet To Probe β-Sheet Cooperativity in Aqueous Solution. *J. Am. Chem. Soc.* 1998; 120(19):4869–4870.
37. Sauter C, Otalora F, Gavira J-A, Vidal O, Giege R, Garcia-Ruiz JM. Structure of tetragonal hen egg-white lysozyme at 0.94 Å from crystals grown by the counter-diffusion method. *Acta Crystallogr. D.* 2001; 57(8):1119–1126. [PubMed: 11468395]
38. Roe DR, Hornak V, Simmerling C. Folding Cooperativity in a Three-stranded [β]-Sheet Model. *J. Mol. Biol.* 2005; 352(2):370–381. [PubMed: 16095612]
39. Ding, F. Ph.D. diss. State University of New York; Stony Brook: 2010. Exploring the Structure and Dynamics of HIV-1 PR by MD Simulations.. (Publication No. AAT 3422802)
40. Gilson MK, Sharp KA, Honig BH. Calculating the electrostatic potential of molecules in solution: Method and error assessment. *J. Comput. Chem.* 1988; 9(4):327–335.
41. Sigalov G, Scheffell P, Onufriev A. Incorporating variable dielectric environments into the generalized Born model. *J. Chem. Phys.* 2005; 122(9):094511–094515. [PubMed: 15836154]
42. Bondi A. van der Waals Volumes and Radii. *J. Phys. Chem.* 1964; 68(3):441–451.
43. Hornak V, Okur A, Rizzo RC, Simmerling C. HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. *Proc. Natl. Acad. Sci. USA.* 2006; 103(4):915–920. [PubMed: 16418268]
44. Powell MJD. UOBYQA: unconstrained optimization by quadratic approximation. *Math. Program.* 2002; 92(3):555–582.
45. Forrest S. Genetic algorithms: principles of natural selection applied to computation. *Science.* 1993; 261(5123):872–878. [PubMed: 8346439]
46. Metcalfe TS, Charbonneau P. Stellar structure modeling using a parallel genetic algorithm for objective global optimization. *J. Comput. Phys.* 2003; 185(1):176–193.
47. Leardi R. Genetic algorithms in chemometrics and chemistry: a review. *J. Chemometr.* 2001; 15(7):559–569.

48. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*. 1983; 22(12):2577–2637. [PubMed: 6667333]
49. Wickstrom L, Okur A, Song K, Hornak V, Raleigh DP, Simmerling CL. The Unfolded State of the Villin Headpiece Helical Subdomain: Computational Studies of the Role of Locally Stabilized Structure. *J. Mol. Biol.* 2006; 360(5):1094–1107. [PubMed: 16797585]
50. Blanco FJ, Rivas G, Serrano L. A short linear peptide that folds into a native stable [beta]-hairpin in aqueous solution. *Nat. Struct. Mol. Biol.* 1994; 1(9):584–590.
51. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* 1999; 314(1-2):141–151.
52. Ryckaert J-P, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 1977; 23(3):327–341.
53. Berendsen HJC, Postma JPM, Gunsteren W. F. v. DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 1984; 81(8):3684–3690.
54. Darden T, York D, Pedersen L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 1993; 98(12):10089–10092.
55. Simmerling C, Elber R, Zhang J. MOIL-View - A Program for Visualization of Structure and Dynamics of Biomolecules and STO - A Program for Computing Stochastic Paths. *Modelling of Biomolecular Structures and Mechanisms*. 1995:241–465.
56. Wickstrom L, Okur A, Simmerling C. Evaluating the Performance of the ff99SB Force Field Based on NMR Scalar Coupling Data. *Biophys. J.* 2009; 97(3):853–856. [PubMed: 19651043]
57. a Simmerling C, Strockbine B, Roitberg AE. All-Atom Structure Prediction and Folding Simulations of a Stable Protein. *J. Am. Chem. Soc.* 2002; 124(38):11258–11259. [PubMed: 12236726] b Hsieh M-J, Luo R. Balancing Simulation Accuracy and Efficiency with the Amber United Atom Force Field. *J. Phys. Chem. B.* 2010; 114(8):2886–2893. [PubMed: 20131885]
58. Day R, Paschek D, Garcia AE. Microsecond simulations of the folding/unfolding thermodynamics of the Trp-cage miniprotein. *Proteins: Struct., Funct., Bioinf.* 2010; 78(8):1889–1899.
59. a Pitera JW, Swope W. Understanding folding and design: Replica-exchange simulations of “Trp-cage” miniproteins. *Proc. Natl. Acad. Sci. USA.* 2003; 100(13):7587–7592. [PubMed: 12808142] b Zhou R. Trp-cage: Folding free energy landscape in explicit water. *Proc. Natl. Acad. Sci. USA.* 2003; 100(23):13280–13285. [PubMed: 14581616] c Paschek D, Hempel S, García AE. Computing the stability diagram of the Trp-cage miniprotein. *Proc. Natl. Acad. Sci. USA.* 2008; 105(46):17754–17759. [PubMed: 19004791]
60. Gallicchio E, Paris K, Levy RM. The AGBNP2 Implicit Solvation Model. *J. Chem. Theory Comput.* 2009; 5(9):2544–2564. [PubMed: 20419084]
61. Mongan J, Svrcek-Seiler WA, Onufriev A. Analysis of integral expressions for effective Born radii. *J. Chem. Phys.* 2007; 127(18):185101. [PubMed: 18020664]
62. Michael SL, Freddie R, Salsbury, Charles LB III. Novel generalized Born methods. *J. Chem. Phys.* 2002; 116(24):10606–10614.
63. Aguilar B, Shadrach R, Onufriev AV. Reducing the Secondary Structure Bias in the Generalized Born Model via R6 Effective Radii. *J. Chem. Theory Comput.* 2010; 6(12):3613–3630.
64. Onufriev, AV. 2010. (private communication)
65. a Lyne PD, Lamb ML, Saeh JC. Accurate Prediction of the Relative Potencies of Members of a Series of Kinase Inhibitors Using Molecular Docking and MM-GBSA Scoring. *J. Med. Chem.* 2006; 49(16):4805–4808. [PubMed: 16884290] b Guimarães CRW, Cardozo M. MM-GB/SA Rescoring of Docking Poses in Structure-Based Lead Optimization. *J. Chem. Inf. Model.* 2008; 48(5):958–970. [PubMed: 18422307]

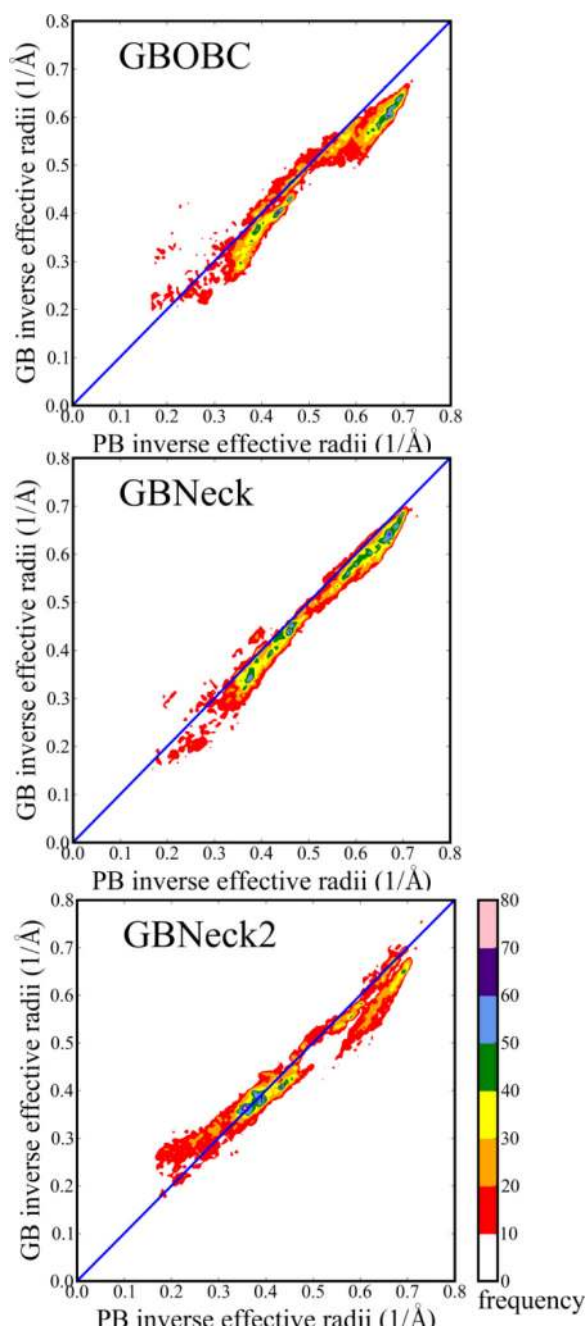


Figure 1. 2D histograms of inverse effective Born radii of each GB model versus PB ‘perfect’ radii for tc5b. Perfect agreement is shown by the diagonal line. The color indicates the frequency (number of atoms) in each bin.

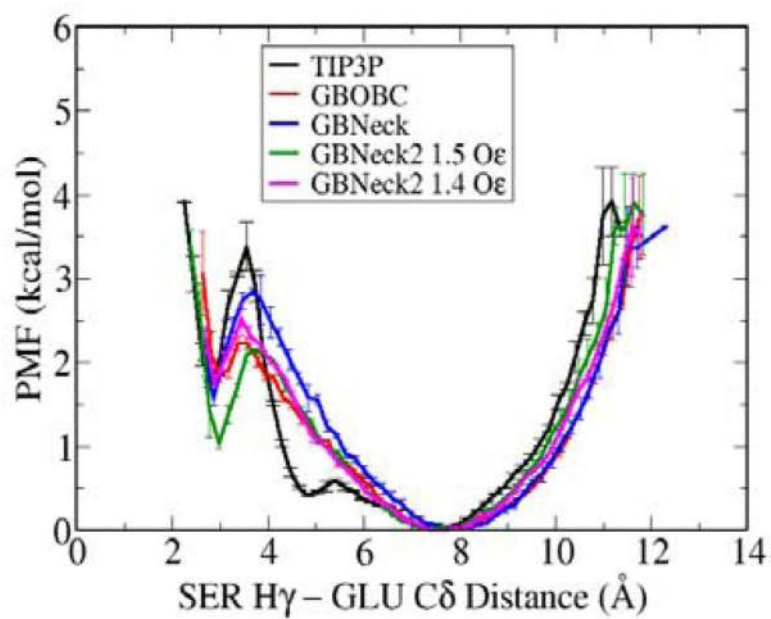


Figure 2. PMFs for side chain H-bond formation in the SAAE model peptide for various solvent models. The 2 GB-Neck2 curves used different Born radii for the Glu side chain carboxyl oxygen atoms, indicated in Å in the legend.

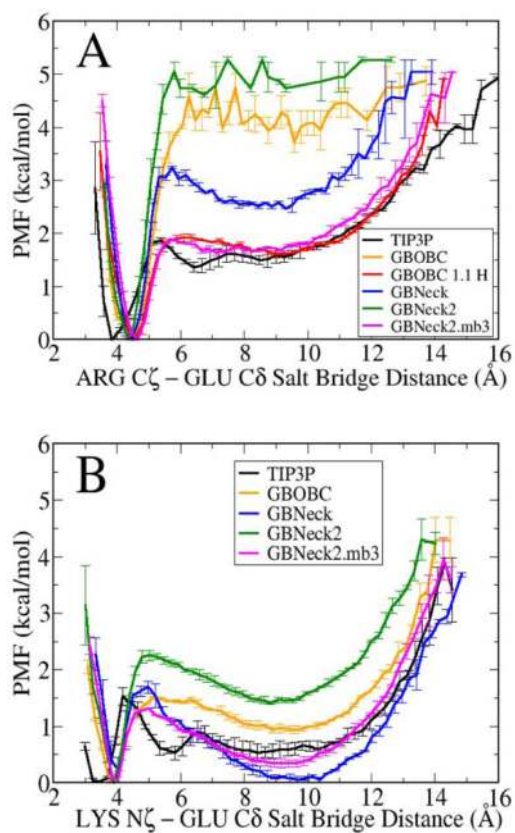


Figure 3. Salt bridge PMFs for various solvent models. Panel A shows the PMF profiles for RAAE (Arg salt bridge) while panel B shows PMFs for KAAE (Lys salt bridge). GB-OBC, GB-Neck and GB-Neck2 used original mbondi2 radii set while GB-OBC 1.1 H^{N+} used mbondi2 with modified H^{N+}(Arg). GB-Neck2.mb3 used the optimized radii set denoted mbondi3 (Table S4).

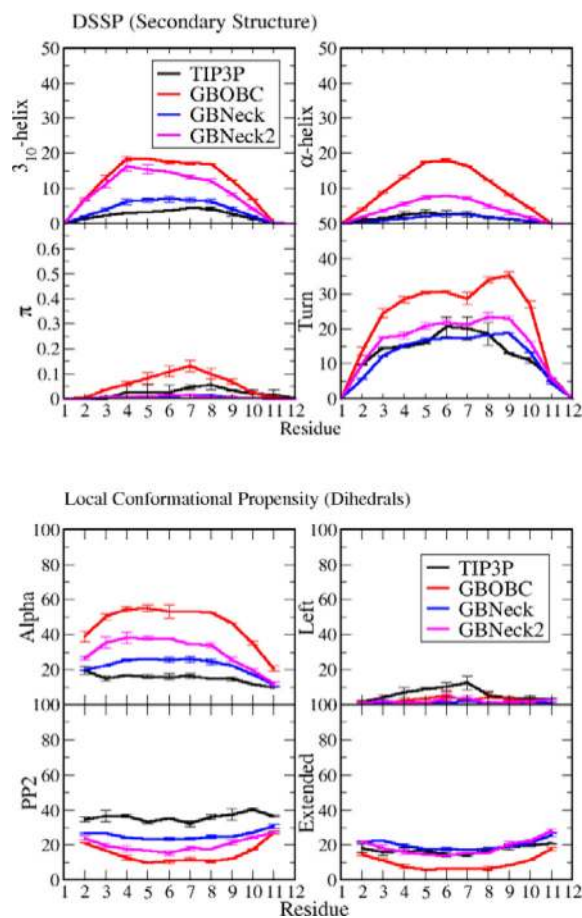


Figure 4. Secondary structure (upper) and local conformational propensities (lower) for each residue of Ala10 at 300K from REMD simulations using different solvent models.

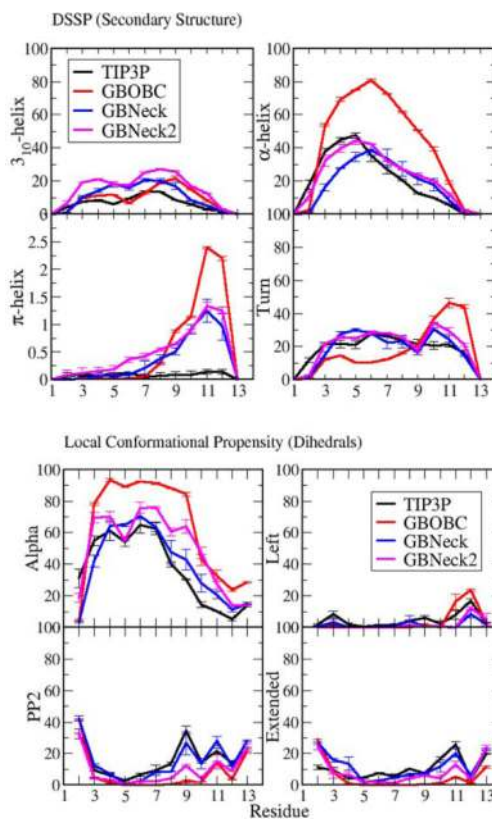


Figure 5. Secondary structure (upper) and local conformational propensities (lower) at 300K for each residue of HP-1, obtained from REMD simulations using different solvent models.

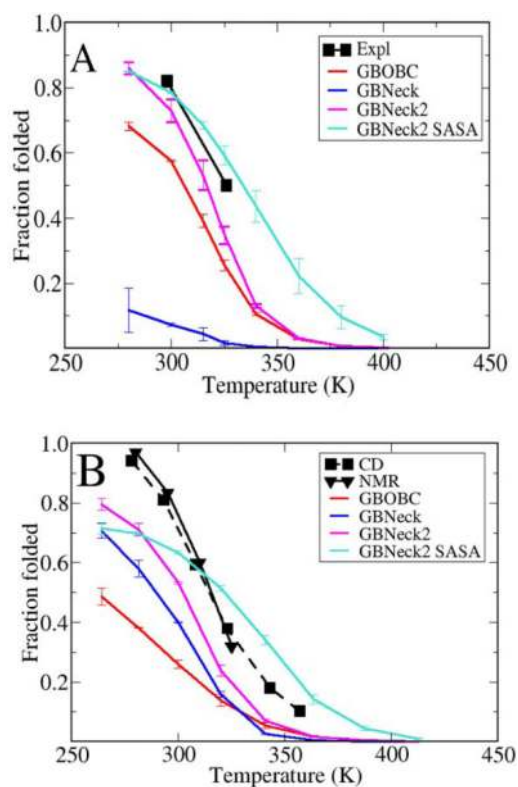


Figure 6. Panel A and B show the thermal stability profiles for the HP5F and tc5b respectively in GB-OBC, GB-Neck and GB-Neck2 (with and without SASA) REMD simulations, compared to experimental data.²⁰⁻²¹

abs_e (kcal/mol), rel_e (kcal/mol) and eff_rad_rmsd (Å) to PB results for each training set after optimization, compared to GB-OBC and GB-Neck models. w is the weighting factor for each component. The objective function for training is the sum of the weighted contributions from each column.

Table 1

	Ala10_set_1	Ala10_set_1	Ala10_set_1	trpzip2	trpzip2	3Ai3	3Ai3	RAAE	RAAE
	abs_e	rel_e	abs_e	rel_e	abs_e	rel_e	abs_e	abs_e	rel_e
	$w=10$	$w=10$	$w=10$	$w=1$	$w=10$	$w=1$	$w=10$	$w=10$	$w=10$
GB-OBC	1.1	1.6	9.5	4.8	7.1	4.5	1.5	1.5	1.3
GB-Neck	2.7	2.3	8.3	7.0	10.6	4.1	1.1	1.1	1.1
GB-Neck2	0.8	1.0	2.8	3.9	3.7	3.7	0.9	0.9	1.2

	HP36	HP36	HP1113	Ala10_set_2	obj_funct
	abs_e	rel_e	eff_rad_rmsd	eff_rad_rmsd	
	$w=1$	$w=10$	$w=50$	$w=250$	
GB-OBC	21.6	6.5	1.8	0.16	381.2
GB-Neck	28.3	5.1	2.3	0.19	444.7
GB-Neck2	4.3	4.8	1.5	0.10	273.8

Table 2

Optimized parameters for GB-Neck2 model.

Parameter	Value	Parameter	Value	Parameter	Value
S_H	1.426	α_H	0.788	α_N	0.503
S_C	1.059	β_H	0.799	β_N	0.317
S_N	0.734	γ_H	0.437	γ_N	0.193
S_O	1.061	α_C	0.734	α_O	0.868
<i>offset</i>	0.195	β_C	0.506	β_O	0.877
S_{neck}	0.827	γ_C	0.206	γ_O	0.388

abs_e and *rel_e* (kcal/mol) between each GB and PB calculation for type I and II test sets, shown for multiple GB models. Type II test sets are indicated in bold.

Table 3

	GB-OBC	GB-Neck	GB-Neck2	GB-OBC	GB-Neck	GB-Neck2
(A) <i>abs_e</i> (kcal/mol)						
Ala10	1.1	2.2	1.0	0.7	0.7	0.5
Trpzip2	9.2	8.4	3.2	1.6	1.9	1.2
3Ai3	7.2	10.6	4.0	2.1	2.0	1.9
RAAE	1.3	1.6	1.4	0.6	0.7	0.5
HP36	21.3	29.7	6.6	6.0	6.0	5.4
tc5b	7.4	13.4	5.3	1.8	2.6	1.8
DPDP	3.4	12.7	3.6	2.0	2.2	1.9
HIV1-PR	115.0	133.1	17.2	HIV1-PR	20.1	16.8
Lysozyme	72.2	88.4	13.1	Lysozyme	13.5	11.9
(B) <i>rel_e</i> (kcal/mol)						
Ala10			1.0	0.7	0.7	0.5
Trpzip2			3.2	1.6	1.9	1.2
3Ai3			4.0	2.1	2.0	1.9
RAAE			1.4	0.6	0.7	0.5
HP36			6.6	6.0	6.0	5.4
tc5b			5.3	1.8	2.6	1.8
DPDP			3.6	2.0	2.2	1.9
HIV1-PR			17.2	HIV1-PR	20.1	16.8
Lysozyme			13.1	Lysozyme	13.4	11.9