PROTEINS: Structure, Function, and Genetics 53:148–161 (2003)

Free Energy Landscape of Protein Folding in Water: Explicit vs. Implicit Solvent

Ruhong Zhou*
IBM T.J. Watson Research Center, Yorktown Heights, New York 10598 and Department of Chemistry, Columbia University, New York, New York 10027

ABSTRACT **The Generalized Born (GB) continuum solvent model is arguably the most widely used implicit solvent model in protein folding and protein structure prediction simulations; however, it still remains an open question on how well the model behaves in these large-scale simulations. The current study uses the β-hairpin from C-terminus of protein G as an example to explore the folding free energy landscape with various GB models, and the results are compared to the explicit solvent simulations and experiments. All free energy landscapes are obtained from extensive conformation space sampling with a highly parallel replica exchange method. Because solvation model parameters are strongly coupled with force fields, five different force field/solvation model combinations are examined and compared in this study, namely the explicit solvent model: OPLSAA/SPC model, and the implicit solvent models: OPLSAA/SGB (Surface GB), AMBER94/GBSA (GB with Solvent Accessible Surface Area), AMBER96/GBSA, and AMBER99/GBSA. Surprisingly, we find that the free energy landscapes from implicit solvent models are quite different from that of the explicit solvent model. Except for AMBER96/GBSA, all other implicit solvent models find the lowest free energy state not the native state. All implicit solvent models show erroneous salt-bridge effects between charged residues, particularly in OPLSAA/SGB model, where the overly strong salt-bridge effect results in an overweighting of a non-native structure with one hydrophobic residue F52 expelled from the hydrophobic core in order to make better salt bridges. On the other hand, both AMBER94/GBSA and AMBER99/GBSA models turn the β-hairpin in to an α-helix, and the α-helical content is much higher than the previously reported α-helices in an explicit solvent simulation with AMBER94 (AMBER94/TIP3P). Only AMBER96/GBSA shows a reasonable free energy landscape with the lowest free energy structure the native one despite an erroneous salt-bridge between D47 and K50. Detailed results on free energy contour maps, lowest free energy structures, distribution of native contacts, α-helical content during the folding process, NOE comparison with NMR, and temperature dependences are reported and discussed for all five models. Proteins 2003;53:148–161.**

Key words: GB models; protein folding; salt-bridge effect; explicit vs. implicit solvent; free energy landscape

INTRODUCTION

Understanding the protein folding mechanism is critical in molecular biology not only because it is one of the fundamental problems remaining in protein science, but also because several fatal diseases are directly related to protein folding/misfolding, such as the Alzheimer's disease, Mad Cow disease, and cystic fibrosis disease.[1] Despite the enormous efforts made by various groups,[2–7] the problem still remains largely unsolved. Experiments that probe proteins at different stages of the folding process have helped to elucidate kinetic mechanisms and the thermodynamic stabilities of folding.[8–11] However, many of the details of protein folding pathways remain unknown. Computer simulations performed at various levels of complexity ranging from simple lattice models to all-atom models with explicit solvent are used to supplement experiment and fill in some of the gaps in our knowledge about folding pathways. Because explicit solvent simulations typically require enormous CPU time, many recent studies have been carried out with implicit solvent models.[12–15] However, it is still largely an open question as to how well these implicit solvent models can predict the thermodynamics as well as the kinetics of protein folding. It will be very interesting to determine whether or not implicit solvent models can reproduce either the results from explicit solvent simulations or experimental results. From some of the preliminary results published recently,[16] we found that the free energy landscape of a β-hairpin in the implicit solvent model SGB is quite different from that in the explicit solvent model SPC using the same OPLSAA force field. Because the GB type model is probably the most popular implicit solvent model in protein folding simulations because of its fast speed compared with the more rigorous Poisson-Boltzmann solvers,[12] we would like to ask the same question for other variations of the GB model, particularly the implementations based on other force fields. Thus, this article will focus on GBSA (GB with Solvent Accessible Surface Area) model as implemented in AMBER package in the context of various AMBER force

*E-mail: ruhongz@us.ibm.com.

Received 13 January 2003; Accepted 6 April 2003

fields, AMBER param94[17] (named as AMBER94/GBSA), param96[18] (AMBER96/GBSA), and param99[19] (AMBER99/GBSA). The results from OPLSAA/SPC and OPLSAA/SGB will also be included for comparison.

The C-terminus β-hairpin of protein G will be used again as the sample system for this study. This β-hairpin has received much attention recently on both the experimental and theoretical fronts,[8–11,20–25] because it is believed to be one of the smallest naturally occurring systems that exhibit many features of a full-size protein and also because it is a fast folder (folds in about 6 μs). Understanding the folding of key protein secondary structures, such as the β-sheet and the α-helix, may provide a foundation for understanding folding in more complex proteins.

The breakthrough experiments by the Serrano[8,9] and Eaton groups[10,11] have recently established the β-hairpin from the C-terminus of protein G as the system of choice to study β-sheets in isolation. These pioneering experiments have inspired much theoretical work on this system using both explicit and implicit solvent models.[20–23,26] For example, Pande and coworkers[21] have studied the kinetics of this β-hairpin with the GB continuum solvent model and an OPLS united atom force field. Karplus and coworkers[22] have explored the free-energy landscape of the same system using a continuum solvent model EEF1[15] and the CHARMM force field (CHARMM19). Very recently explicit solvent simulations have also been used to study this system. Garcia and coworkers[23] have studied the free energy landscape of this system in explicit solvent using the AMBER force field (AMBER94) with a short cutoff (9 Å) in the electrostatic interactions. Zhou and coworkers[27] have also explored this system with the OPLSAA force field and explicit solvent with no cutoffs in the electrostatic interactions by utilizing the P3ME method.[28] Thus, it is of great interest to compare the free energy landscapes from continuum solvent models to those from the explicit solvent models, particularly for establishing the foundation on how well these implicit solvent models can reproduce the results from explicit solvent simulations or experiments.

In general, a poor simulation result could be due to a poor model, an insufficient sampling, or both. This is particularly true for protein folding simulations because of their enormous sampling requirements. In order to eliminate the sampling issue in this study, we used a highly parallel Replica Exchange Method (REM) to explore the free energy landscape. For the explicit solvent simulations,[27] a total of 64 replicas are used with temperature spanning 270 to 695 K. For the continuum solvent simulations, a total of 18 replicas are simulated with the same temperature coverage. Because the force field is normally parameterized at room temperature, we do not expect it to yield accurate results for higher temperatures; nevertheless, these high-temperature replicas permit the system to rapidly cross the energy barriers and thus lead to efficient sampling at the lower temperatures. It is found that the free energy landscapes for the continuum solvent models are quite different from that of the explicit solvent model. Some of the non-native states are heavily overweighted in the continuum solvent models compared to the explicit solvent model, and more importantly the lowest free energy structures from the continuum solvent models are not always the native β-hairpin structure. For example, it is found that the OPLSAA/SGB model over-stabilizes salt bridges, and both the AMBER94/GBSA and AMBER99/GBSA models turn the β-hairpin into an α-helix. Even though it was previously reported[23] that AMBER94 with explicit solvent model also showed a tendency to overpopulate the α-helix, the continuum solvent model GBSA seems to greatly enhance this tendency. These results indicate that even with the large successes of these GB type continuum solvent models in many fields, such as pKa calculations, surface electrostatic potentials, solvation free energies, and ligand-receptor bindings,[29–31] one must be cautious with large-scale simulations, which involve dramatic conformational changes, such as protein folding.

METHODOLOGY

The replica exchange method has been implemented in the context of the molecular modeling package IMPACT.[28,32] Replicas are run in parallel at a sequence of temperatures. Periodically, the configurations of neighboring replicas are exchanged and acceptance is determined by a Metropolis criterion that guarantees detailed balance. The acceptance criterion used in REM is identical to that in Jump Walking methods.[32] Because the high-temperature replica can traverse high-energy barriers, there is a mechanism for the low-temperature replicas to overcome the quasi ergodicity they would encounter in a one-temperature walk. The replicas themselves can be generated by Monte Carlo (MC), by Hybrid Monte Carlo (HMC),[33] or by molecular dynamics (MD) with velocity rescaling as used by Okamoto et al.[34] The HMC method, which uses MD to generate possible conformations, is often called "bad MD but good MC,"[33] because it typically uses a larger timestep than the normal MD. The HMC method is adopted in the OPLSAA/SGB simulation,[16] partly because that SGB does not provide the analytical surface area gradients for its cavity energy term.[14] The HMC method can help bypass this problem.[16] The current AMBER9x/GBSA (including AMBER94, AMBER96, and AMBER99) simulations as well as the explicit solvent simulation[27] all used MD for sampling with the velocity rescaling approach. Both approaches are probably equally good for this study, but in general HMC is more efficient for smaller systems like the peptide in implicit solvent because it can use a much larger timestep in its underlying MD (the time step scales as $O(1/\sqrt{N})$; thus, it is unfavorable for larger systems such as proteins in explicit solvent). It should be pointed out though that because all REM uses the Metropolis criterion for replica exchanges, they are essentially MC methods not MD methods. Thus, the MD timings reported here and also in previous REM studies[16,23,27,34] should not be taken as direct kinetic measurements.

The REM itself can be summarized as the following two-step algorithm:

(i) Each replica i ($i = 1,2, \ldots, M$) at fixed temperature T_m ($m = 1, 2, \ldots, M$), is simulated *simultaneously*

and *independently* for a certain number of MC or MD steps.

(ii) Pick a pair of replicas, and exchange them with the acceptance probability,

$$T(x_i|x_j) = \begin{cases} 1, & \text{for } \Delta \le 0, \\ \exp(-\Delta), & \text{for } \Delta > 0. \end{cases} \quad (1)$$

where $\Delta = (\beta_i - \beta_j)(V(x_j) - V(x_i))$, β_i and β_j are the two reciprocal temperatures, x_i is the configuration at β_i, and x_j is the configuration at β_j, and $V(x_i)$ and $V(x_j)$ are potential energies at these two configurations, respectively. After the exchange, go back to step (i).

All the replicas are run in parallel on M processors (M = 64 for explicit solvent, and M = 18 for implicit solvents); and in step (ii), only exchanges between neighboring temperatures are attempted because the acceptance ratio decreases exponentially with the difference of the two β's.

In the following paragraphs, we will briefly describe the GB model, which was first developed by Still and coworkers.[13] The GB model is probably the most widely used continuum solvent model in protein folding and protein structure prediction simulations. Many algorithmic improvements, reformulation, and reparametrizations with specific force fields have been proposed by various groups after the initial model.[14,29,30,35,36] Here we will discuss two implementations of the GB model. Surface GB (SGB) parameterized with OPLSAA by Friesner and coworkers[14] and GB with Solvent Accessible Surface Area (GBSA) parameterized with AMBER by Case and coworkers.[29,36] The main difference between SGB and GB is that SGB uses the surface integral rather than the volume integral for the so-called "single energy term" (see below, Eq. 5). Of course, the parametrizations in various GB models could be very different too. Friesner et al.'s SGB model has been parameterized with the OPLSAA force field to reproduce the experimental solvation free energies for about 200 small organic molecules[14,30]; on the other hand, the GBSA implemented in AMBER program by Case et al. is parameterized against the AMBER94 force field. The combination of GBSA with three different versions of the AMBER force field will be explored here: AMBER94/GBSA, AMBER96/GBSA, and AMBER99/GBSA.

In typical GB models the total solvation free energy[13,14] of a solute such as a protein is expressed as the sum of the "reaction field energy" U_{rxn}, and the "cavity energy" U_{cav}, such that

$$U_{GB} = U_{rxn} + U_{cav}. \quad (2)$$

The cavity energy U_{cav} is often estimated by the Solvent Accessible Surface Area (SASA). Both SGB and GBSA uses such a term for the cavity energy with the surface tension coefficient about 5.0 cal/mol/Å^2.[14,36] The total reaction field energy is also a sum of two terms (consult the literature[13,14] for more details), the so-called "single energy" U_{se} and "pair energy" U_{pe},

$$U_{rxn} = \sum_i U_{se}(q_i, \mathbf{r}_i) + \sum_{i<j} U_{pe}(q_i, q_j, \mathbf{r}_i, \mathbf{r}_j), \quad (3)$$

where the single energy U_{se} can be expressed as either a volume integral as in the original GB model[13] or a surface integral as in the SGB model[14]:

$$U_{se}^{GB} = \frac{3}{8\pi}(1/\epsilon_i - 1/\epsilon_o)\int_V \frac{q_k^2}{|\mathbf{R} - \mathbf{r}_k|^4}\, d^3\mathbf{R} \quad (4)$$

$$U_{se}^{SGB} = -\frac{1}{8\pi}(1/\epsilon_i - 1/\epsilon_o)\int_S \frac{q_k^2}{|\mathbf{R} - \mathbf{r}_k|^4}(\mathbf{R} - \mathbf{r}_k)\cdot\mathbf{n}(\mathbf{R})\, d^2\mathbf{R}, \quad (5)$$

where $\mathbf{r}_k$ is the position of charge q_k and $\mathbf{n}(\mathbf{R})$ is the surface normal, ϵ_i is the dielectric constant for the interior of the solute (for proteins, it is typically about 1.0–4.0 in SGB,[14]), and ϵ_o is the dielectric constant for outside water (78.5). It can be easily proved with Green's theorem that the above two equations, U_{se}^{GB} and U_{se}^{SGB}, are equivalent.[14] However, the volume integral scales as $O(N)$ for each atom, whereas the surface integral only scales as $O(N^{2/3})$; thus SGB scales more favorably for larger systems.[14]

The "pair energy", i.e., the pairwise screened Coulomb energy, on the other hand, is defined as

$$U_{pe} = -(1/\epsilon_i - 1/\epsilon_o)\frac{q_i q_j}{\sqrt{r_{ij}^2 + \alpha_{ij}^2 e^{-D}}}, \quad (6)$$

which is the same for both GB and SGB model.[14] The parameter $\alpha_{ij} = \sqrt{\alpha_i \alpha_j}$ (α_i and α_j are the Born radius) and parameter $D = r_{ij}^2/(2\alpha_{ij})^2$.

RESULTS AND DISCUSSION

The β-hairpin under study here is taken from the C terminus (Res. 41–56) of protein G (PDB 2gb1). The 16-residue β-hairpin is capped with the normal Ace and Nme groups, resulting in a blocked peptide sequence of Ace-GEWTYDDATKTFTVTE-Nme, with a total of 256 atoms. In the explicit solvent simulation, the solvated system has 1361 water molecules (SPC water, with density 1.0 g/cm^3) and 3 counterions (3 Na^+ ions) for neutralizing the molecular system (total 4342 atoms) with periodic boundary condition. In the continuum solvent models, no water molecules or counter ions are included, so the total system size is 256 atoms of the β-hairpin. A dielectric constant of 1.0 was used in the explicit solvent model as well as all the AMBER9x/GBSA models (default in AMBER9x/GBSA models[36]), and a dielectric constant of 2.0 was used in the OPLSAA/SGB model (recommended by the OPLSAA/SGB model[14]). It should be pointed out that assigning dielectric constants in continuum solvent models for proteins can be very tricky.[37,38] In this study, we have tried dielectric constants of 1.0, 2.0, and 4.0 for the solute in the OPLSAA/SGB model, and very similar free energy contour maps were obtained.[16] The OPLSAA/SPC and OPLSAA/SGB simulations are carried out with IMPACT package,[28,32] and AMBER9x/GBSA simulations are carried out with AMBER package.[17,19] A total of 64 replicas for the explicit solvent model and 18 replicas for

the implicit solvent models are simulated with the same temperatures range of 270 to 695 K. The explicit solvent simulation needs more replicas because there are many more atoms in the system, and the temperature gap in REM scales like $O(1/\sqrt{N})$. Following a similar approach as the previous explicit solvent simulation on this β-hairpin,[23] we used the native structure from the C-terminus of protein G (2gb1) as the initial starting point for equilibration. A standard protocol of 1000 steps of conjugate gradient minimization followed by a 100-ps MD equilibration is performed for each replica. The final configurations of each replica in the above equilibration process are then used as the starting points for each replica. A total production run of 3-ns MD in each replica is collected for all models. The replica exchanges are attempted every 400 fs, and protein configurations are saved every 80 fs.

The optimal temperature distribution in REM should be exponential, i.e., $T_n = T_0 \exp(kn)$, where T_n is the nth temperature and T_0 and k are constants that can be easily determined by running a few short trial simulations. In the current explicit solvent simulation, the 64 replicas are distributed at temperatures 270, 274, 278, . . . , 675, 685, and 695 K, and in the implicit solvent simulation the 18 replicas were distributed at 270, 282, 295, . . . , 611, 649, and 695 K. With these temperature distributions we obtain an acceptance ratio of approximately 30−40% for the adjacent replica exchanges in both explicit and implicit solvent simulations. We observe that the "temperature trajectory" for one replica visits all the temperatures many times during the 3-ns MD run, and at a given temperature all of the replicas are also visited many times during the same *MD* run, indicating that our temperature series are reasonably optimized. It should be pointed out that a 3-ns MD might not be enough to fully equilibrate the system particularly in the explicit solvent even with 64 REM replicas (total about 200-ns MD). However, we did notice that there are many tens of "transitions" (exchanges/jumps) followed by extensive relaxation between various free energy states (defined below). If there are only a few transitions between free energy states, one might worry about the noncomplete equilibration, here the many tens of transitions indicate that a reasonable equilibration might be achieved. Of course, a more rigorous proof will need two or more simulations starting from different configurations to see if they converge to the same result. However, this is probably beyond the current capacity (the 3-ns MD with explicit solvent already takes about 1.5 months on 64 processors of IBM SP2 Power3-375MHz clusters, or equivalently 8 processor-years).

The free energy landscape is determined by calculating the normalized probability distribution function $P(X)$ from a histogram analysis on the conformations sampled.[23] Because the potential of mean force (PMF) $W(X)$, or equivalently the free energy, is related to this probability distribution function through the relation

$$P(X) \equiv Z^{-1}\exp(-\beta W(X)), \tag{7}$$

where X is a specified set of reaction coordinates and Z is the partition function, the relative free energy change corresponding to a change in reaction coordinates can be easily obtained from,

$$W(X_2) - W(X_1) = -kT \ln \frac{P(X_2)}{P(X_1)}. \tag{8}$$

In the previous preliminary work[27] we determined the free energy surfaces for the β-hairpin in explicit water for various reaction coordinates,[27] including the number of β-strand hydrogen bonds, the hydrophobic core radius of gyration, the fraction of native contacts, the radius gyration of the entire peptide, the principal components,[39] RMSD, and found the number of β-strand hydrogen bonds, N_{HB}^{β}, and the hydrophobic core radius gyration, Rg^{core}, to be very informative for this small β-hairpin. Thus we use these two reaction coordinates to compare the free energy contour maps for the explicit and the implicit solvent models. The number of β-strand hydrogen bonds N_{HB}^{β} is defined as the number of backbone-backbone hydrogen bonds excluding the two at the turn of the hairpin[27] (total 5 such hydrogen bonds[20,26,27]). A hydrogen bond is counted if the distance between two heavy atoms (N and O in this case) is less than 3.5 Å and the angle N—H · · · O is larger than 150.0 degree. The hydrophobic core radius gyration Rg^{core} is the radius of gyration of the side-chain atoms on the four hydrophobic residues, W43, Y45, F52, and V54. In the following subsections, we also calculate the number of residues in the β-sheet or α-helix format. This is done with the program STRIDE[40] downloaded from web http://www.stride.org, which uses both hydrogen bond energies and dihedral angles in assigning secondary structure to each residue.[40]

Free Energy Landscape

We first compare the free energy landscapes from the five force field/solvation models. Figure 1 shows the free energy contour maps with the two reaction coordinates, N_{HB}^{β} and Rg^{core}. As shown previously,[27] the explicit solvent model OPLSAA/SPC in general agrees quite well with experiment near room temperature; thus we use it as a benchmark for comparisons with continuum solvent models. Because some of the preliminary results of OPLSAA/SPC and OPLSAA/SGB models have been published,[16,27] we will mainly focus on the AMBER9x/GBSA results, with brief reviews on OPLSAA/SPC and OPLSAA/SGB for completeness.

The free energy surface from OPLSAA/SPC reveals several interesting features of this β-hairpin folding in explicit water: (1) there are four states in this two reaction coordinate representation at 310K: the native folded state (F), the unfolded state (U), and two intermediates: a "molten globule" state, called state H[20,23,27] (hydrophobic core formed but no beta strand H-bonds yet) and a partially folded state (P); (2) the native state F has the lowest free energy, i.e., the most heavily populated structure is the native structure; (3) the overall shape of the free energy contour map is an "L" shape, which indicates the folding process is driven by hydrophobic core collapse not by hydrogen bond zipping, because otherwise a more "diagonal" shape should be found.[27] These findings are

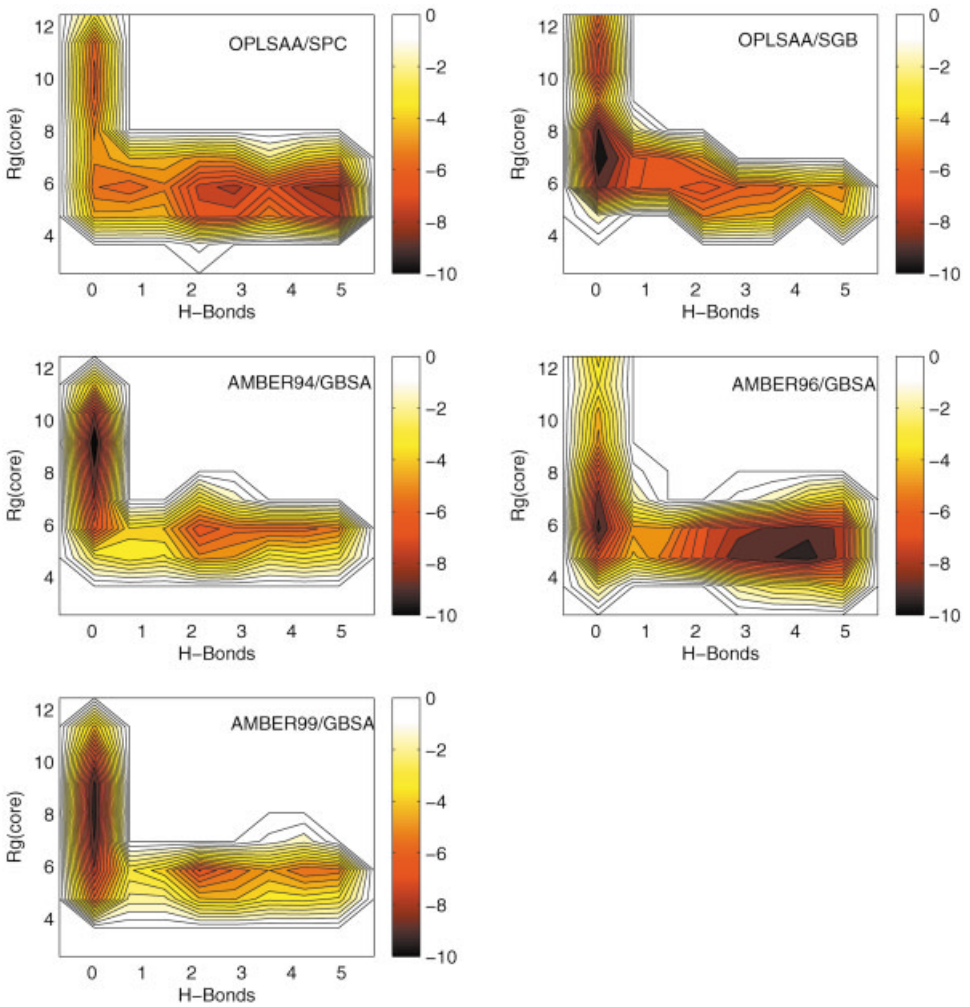

Fig. 1. Comparison of the free energy contour maps versus the number of beta strand H-bonds N_{HB}^{β} and the hydrophobic core radius gyration Rg^{core}, (a) OPLSAA/SPC, (b) OPLSAA/SGB, (c) AMBER94/GBSA, (d) AMBER96/GBSA and (e) AMBER99/GBSA. A hydrogen bond is counted if the distance between two heavy atoms (N and O in this case) is $<$3.5 Å and the angle N–H $\cdots$ O is $>$150.0°. The free energy is in units of kT, and contours are spaced at intervals of 0.5 kT.

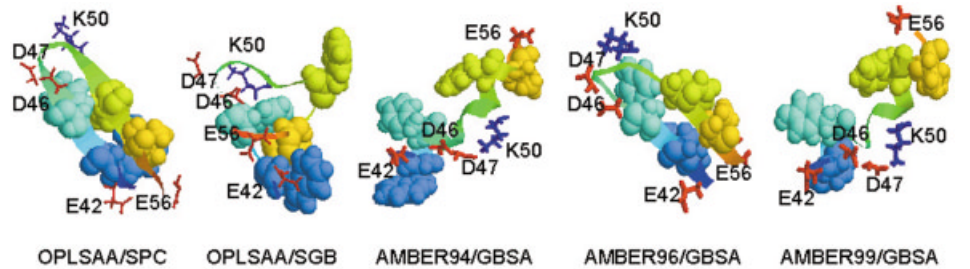

Fig. 2. Comparison of the most popular structures at the lowest free energy state for various models, (a) OPLSAA/SPC, (b) OPLSAA/SGB, (c) AMBER94/GBSA, (d) AMBER96/GBSA, and (e) AMBER99/GBSA. The hydrophobic residues (W43, Y45, F52, and V54) are represented by spacefill and charged residues (E42, D46, D47, K50, and E56) are represented by sticks with positively charged residues colored blue and negatively charged residues colored red, and the rest are represented by ribbons.

generally consistent with both experimental and theoretical results from others,[8,10,20,22,23,26] but there are still some differences. One important difference from experiment is in the folding mechanism. Eaton and coworkers[10,11] developed a helix-coil-type model to provide a structural interpretation for their equilibrium and kinetic data from temperature jump and TRP fluorescence quench experiments. The model suggests a mechanism in which

folding is initiated at the turn and propagates toward the tails by forming hydrogen bonds one by one (i.e., H-bond zipping mechanism), so that the hydrophobic core, from which most of the stabilization derives, form relatively late in the process.[22]

Our simulation shows no evidence of this hydrogen-bond driven model. Instead, our simulation supports a hydrophobic-core driven mechanism. This is also found by Pande et al.[20] using the CHARMM22 force field with explicit solvent TIP3P water and Garcia et al.[23] using AMBER94 force field with TIP3P water. Of course, the free energy barrier from state H to state P is very small, <0.8 kT (at 310 K, this is about 0.48 kcal/mol). This implies the transition from state H to state P and F is probably smooth at 310 K, and the hydrogen bond zipping could occur nearly simultaneously with the hydrophobic core collapse as also suggested by Thirumalai and coworkers.[26]

The free energy contour map from OPLSAA/SGB is shown in Figure 1(b). The only difference between this model and the above OPLSAA/SPC model is in solvation. Surprisingly, the free energy contour map from OPLSAA/SGB is quite different from that of OPLSAA/SPC: (1) the native state (N_{HB} = 4–5 and $Rg^{core} \approx 5.8$ Å) is no longer the lowest free energy state in OPLSAA/SGB; (2) the most heavily populated state, state H (named after OPLSAA/SPC), has no meaningful β-strand hydrogen bonds ($N_{HB} \approx 0$), and also it has a slightly higher radius of gyration for the hydrophobic core, $Rg^{core} \approx 7.0$ Å (representative structures will be shown in next subsection). It has about 2.92 kcal/mol (4.75 kT) lower free energy than the native state. The lower free energy of state H will be shown to be related to the fact that SGB overstablizes the salt bridges between charged residues.

Figure 1 (c–e) shows the free energy contour maps for AMBER9x/GBSA models. Interestingly, AMBER94/GBSA and AMBER99/GBSA show very similar free energy contour maps, even though AMBER99[19] has all parameters refit from AMBER94[17] (in contrast, AMBER96[18] has only some backbone torsion parameters refit from AMBER94). The AMBER94 force field by Cornell et al.[17] is probably still the most widely used version of the AMBER force field. It is clear from Figure 1 (c,e) that (1) the native state is again not the lowest free energy state in either AMBER94/GBSA or AMBER99/GBSA and (2) the lowest free energy state, H state, has a much higher Rg^{core}, about 9.0 Å in AMBER94 and 8.1 Å in AMBER99/GBSA (for comparison, it is about 5.8 Å in OPLSAA/SPC and 7.0 Å in OPLSAA/SGB). The H state also has about 2.95 and 2.60 kcal/mol (4.80 and 4.22 kT) lower free energy than the corresponding native state in AMBER94/GBSA and AMBER99/GBSA, respectively. As will be shown later, the larger core radius gyration is due to the fact that the β-hairpin has been turned into an α-helix in both AMBER94/GBSA and AMBER99/GBSA. The overall shape of the free energy contour map still remains an "L" shape; however, because the lowest free energy state is no longer the native state, the "folding mechanism" means something totally different now. It shows more of turning a stretched chain or a folded β-hairpin into an α-helix by making local backbone hydrogen bonds along the α-helix axis.

On the other hand, AMBER96/GBSA surprisingly shows a very similar free energy contour map as compared to the explicit solvent model. The AMBER96 force field is a derivative of the AMBER94 force field with the only changes in the torsional parameters. This is the only continuum solvent model examined here that mimics the explicit solvent simulation, even though the β-hairpin populations at low temperatures are not as good as the explicit solvent model when compared with experiment; for example, at 282 K the β-hairpin population from estimation of the fraction of native contacts is 57% in AMBER96/GBSA compared with 74% in OPLSAA/SPC and ~80% in experiment. Nevertheless, similar features can be seen from the free energy contour maps of AMBER96/GBSA and OPLSAA/SPC: both show the native folded state F having the lowest free energy; both show the intermediate H state[20,23,27] having a core radius gyration ~5.8 Å; and both show an "L" shape contour map. The reason why AMBER96/GBSA works reasonable well compared with the explicit solvent model and experiment, but not AMBER94 or AMBER99, might be mainly due to the backbone torsional parameters used in various AMBER force fields. Both AMBER94 and AMBER99 seem to have backbone ϕ and ψ torsional parameters too favorable for α-helix conformations. There might also exist an incorrect balance between the gas phase force field parameters and GBSA solvation parameters (more discussions below). Of course, there are still some noticeable differences in contour maps of AMBER96/GBSA and OPLSAA/SPC; for example, the H state in AMBER96/GBSA has a significantly lower free energy than OPLSAA/SPC. The free energy difference between the H state and F state is only about 0.16 kcal/mol (0.26 kT) in AMBER96/GBSA, whereas it is ~1.76 kcal/mol (2.89 kT) in OPLSAA/SPC. For comparison, this difference is −2.92, −2.95, and −2.60 kcal/mol in OPLSAA/SGB, AMBER94/GBSA, and AMBER99/GBSA, respectively, where the H states have much lower free energies than the native states. In addition, the free energy barrier from the H state to F state increases from about 0.8 kT in OPLSAA/SPC to about 4.4 kT in AMBER96/GBSA, which means that the transition from H state to F state will be much harder in AMBER96/GBSA. As will be shown in next subsection, there are also slight differences in the lowest free energy structures.

Lowest Free Energy Structure

In order to understand the free energy landscape and folding mechanism better, we also examined the structures at the lowest free energy state carefully. The structures belonging to the free energy basins are partitioned into clusters defined such that a structure belongs to a cluster if it has an RMSD no larger than 1 Å from at least one other structure in that cluster. This clustering algorithm allows us to determine the unique structures in a free energy basin and also the populations in each cluster bin. Figure 2 shows the most populated structures from the lowest free energy states of OPLSAA/SPC, OPLSAA/

SGB, AMBER94/GBSA, AMBER96/GBSA, and AMBER99/GBSA, respectively. The hydrophobic residues (W43, Y45, F52, and V54) are represented by spacefill and charged residues (E42, D46, D47, K50, and E56) are represented by sticks with positively charged residues colored blue and negatively charged residues colored red, and the rest are represented by ribbons. The lowest free energy structure in the explicit solvent model, OPLSAA/SPC, is basically the native β-hairpin structure; thus again the focus will be on structures from continuum solvent models. Figure 2(b) shows the most popular structure of OPLSAA/SGB (from state H). Two interesting observations emerge from the comparison of this structure with the native structure: (1) the hydrophobic residue F52 is expelled from the hydrophobic core in OPLSAA/SGB, while it is well packed with other three hydrophobic residues (W43, Y45, V54) in the explicit solvent model. In other words, all four hydrophobic residues are well packed in the explicit solvent but in the continuum solvent SGB, only three residues W43, Y45, V54 are packed; (2) in the explicit solvent, the side chains of charged residues extend completely into the solvent, i.e., are fully solvated; whereas in the continuum solvent SGB model, the charged residues are clustered to form salt bridges between opposite charges. For example, D46 and D47 form two salt bridges with K50 near the β-hairpin turn, and the C-terminal end residue E56 also swings toward K50 in order to get closer to the positive charge. The net effect of this salt bridge formation brings the oppositely charged residues, two near the β-hairpin turn (D46, D47) and one from the C-terminal end (E56) into closer contact with residue K50, thereby expelling the hydrophobic residue F52 (in the middle of the same β-strand as K50) from the hydrophobic core. This suggests that the balance between electrostatic interactions and the hydrophobic interactions is no longer preserved in SGB. The electrostatic interactions between the charged residues (salt bridges) overwhelm the hydrophobic interactions between the four hydrophobic core residues.

Figure 2(c) shows the lowest free energy structure of AMBER94/GBSA. The most astonishing observation is that the β-hairpin has been turned into an α-helix. Obviously, the AMBER94 force field biases the structure to form an α-helix. It has been previously reported that the AMBER force field has a tendency to overestimate the α-helix because of some backbone torsion parameters[23,41]; for example, Garcia *et al.*[23] found that there is ~15–20% α-helix from clustering for this same β-hairpin at 282 K using AMBER94 and an explicit solvent model TIP3P. However, what seems not previously known is that the continuum solvent model GBSA dramatically enhances this α-helix tendency. Following a similar approach as Garcia et al.,[23] the α-helix population is estimated to be 77% by clustering H state with $(N_{HB}^{\beta}, Rg^{core}) = (0.0 \pm 1.0, 9.0 \pm 1.0)$. Another method will be used in the following subsection to estimate the α-helix population by directly counting the α-helical residues, and the results are found to be very similar. Second, the hydrophobic core is completely broken. The four hydrophobic residues, W43, Y45, F52, and V54, are separated from each other in order to fit in an α-helix. Another interesting observation is that even though the salt-bridge effect is not as strong as in the OPLSAA/SGB case, the charged sidechains still show a tendency to get closer in order to have stronger electrostatic interactions; for example, residues K50 and D47 exhibits a clear salt bridge. This is found in all the three AMBER9x/GBSA models as shown in Figure 2(c–e). The overall salt bridge strength seems to be smaller though in AMBER9x/GBSA compared with OPLSAA/SGB. The overly strong salt-bridge effect in SGB might be related to the fact that the van der Waals surface, but not the more rigorous Connolly surface,[42] was used in SGB model[14] in the surface integral for the single energy in Eq. 5. This will make the loss in single energies not strong enough to overcome the gain in Coulombic interactions when two oppositely charged groups associate in SGB, because the overlaps of the van der Waals surfaces are typically less than the overlaps of the Connolly surfaces when two charges approach each other. On the other hand, the less serious salt-bridge effect in AMBER9x/GBSA models could also be because the AMBER force fields have so strong backbone torsional parameters that overrun the salt-bridge effects. Garcia and coworkers have tried to turn off some of these backbone torsions completely by setting them to zero, the α-helix content does indeed decrease significantly (personal communication).

The structure from AMBER99/GBSA shown in Figure 2(e) is very similar to that from AMBER94/GBSA discussed above, except that the α-helix turns are not as elegant as those in AMBER94/GBSA (see more structures below in Fig. 3). Similarly, the α-helix population is estimated to be 61%, with the H state population clustering using $(N_{HB}^{\beta}, Rg^{core}) = (0.0 \pm 1.0, 8.1 \pm 1.0)$. Finally, Figure 2(d) shows the most populated structure in AMBER96/GBSA, which is basically the same as the native β-hairpin structure except that there is a salt bridge between charged residues D47 and K50. Of course, the four hydrophobic core residues are also not packed as well as the native structure; for example, the hydrophobic residues W43 and F52 are slightly unpacked compared with the native structure. Nevertheless, this is the only native-like structure found in the lowest free energy state in all four continuum solvent models studied here.

Figure 3 shows a few other representative structures for the three bad models, OPLSAA/SGB, AMBER94/GBSA, and AMBER99/GBSA. These are the structures with highest populations from structure clustering analysis. Some common features can be found among them: (1) no β-strand hydrogen bonds exist; (2) the hydrophobic cores are all destroyed in favor of either helical structures as in AMBER94/GBSA, and AMBER99/GBSA or overstablized salt-bridged structures as in OPLSAA/SGB; (3) they all come from the so-called H state.[20,23,27] Interestingly, there is a slight difference in the α-helix structures from AMBER94/GBSA and AMBER99/GBSA. The AMBER94/GBSA structures show much better α-helices, whereas the α-helices in AMBER99/GBSA structures are all somewhat distorted. This reflects the effort made in AMBER99 as to fix the

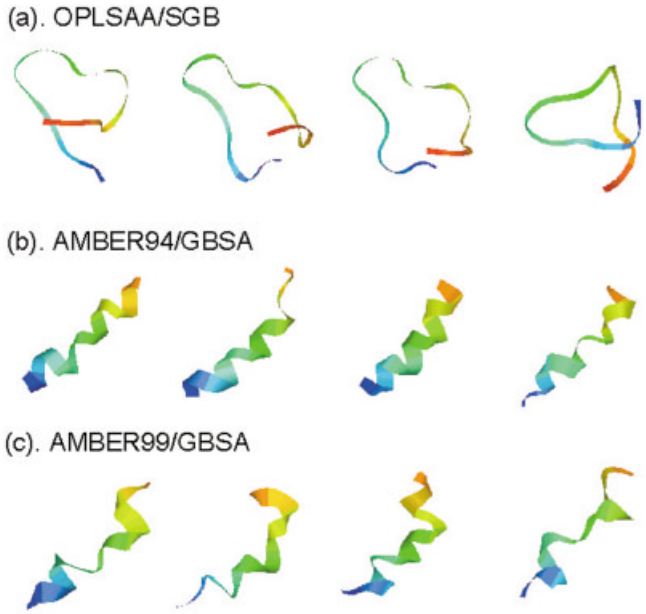

Fig. 3. Other representative structures in the lowest free energy basin (state H) in implicit solvent models OPLSAA/SGB, AMBER94/GBSA, and AMBER99/GBSA. All of them show very different structures from the native one. See text for more details.

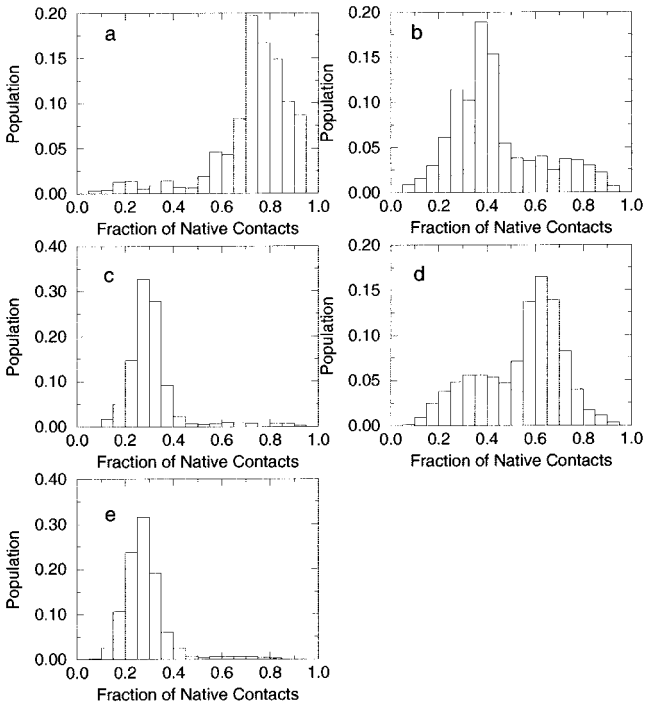

Fig. 4. Comparison of the population histogram at various fraction of native contacts for (a) OPLSAA/SPC, (b) OPLSAA/SGB, (c) AMBER94/GBSA, (d) AMBER96/GBSA, and (e) AMBER99/GBSA. In the explicit solvent model OPLSAA/SPC, the most heavily populated states have ~70–80% of native contacts, whereas in all implicit solvent models except for AMBER96/GBSA, the most heavily populated states have much less native contacts. The AMBER96/GBSA model shows a decent population of native contacts.

α-helix tendency problem, even though it is not very successful as shown in this case.

Native Contacts

Another way to look at this free energy landscape is to profile the conformational distribution with the fraction of native contacts. Figure 4 shows the population histogram at various fraction of native contacts for all the five models at 310 K. In explicit solvent OPLSAA/SPC, the most heavily populated state shows 70–80% of the native contacts, whereas in the implicit solvent OPLSAA/SGB, the most heavily populated H state, has only ~30–40% of the native contacts. The AMBER94/GBSA and AMBER99/GBSA show an even smaller percentage of native contacts, only 20–30%, in the most heavily populated H states. Again, only the AMBER96/GBSA shows a good fraction of native contacts at the most heavily populated state among all implicit solvent models. The heavily populated decoy-like structures shown in Figure 3 for OPLSAA/SGB, AMBER94/GBSA, and AMBER99/GBSA are consistent with this analysis. The number of native contacts formed in the H states of OPLSAA/SGB, AMBER94/GBSA, and AMBER99/GBSA are significantly less than those in the native state. As mentioned earlier, for the OPLSAA/SGB model, we have also tried other dielectric constants, such as $\epsilon_i = 1.0$ and 4.0, for the peptide, but the fraction of native contacts in the most heavily populated states hardly changes. It improves only slightly with a higher dielectric constant, increasing 2–3% from $\epsilon_i = 2.0$ to 4.0, which is still way too low compared with the explicit solvent model. This indicates that at least for SGB the erroneous salt-bridge effect and the imbalance between polar and nonpolar interactions in the implicit solvent model are not eliminated. Thus, the problem cannot be easily fixed by increasing the overall protein dielectric constant, although it is possible that introduction of a much larger dielectric screening of the charged residue interactions alone might well do the trick. For the AMBER94/GBSA and AMBER99/GBSA cases, the problem is worsen by the coupling of backbone torsion parameters, which results in too much α-helical content.

α-Helix Content

Whether or not there exist significant intermediate α-helical structures during the folding process is of great interest and also under heavy debate recently. As mentioned earlier, Garcia et al.[23] found that significant helical content exists (15–20%) at low temperatures using the AMBER94 force field with TIP3P explicit water and also that these conformations are only slightly unfavorable energetically with respect to hairpin formation at biological temperatures. Pande *et al.*[29] also found significant helical intermediates at 300 K using OPLS united atom force field with GBSA continuum solvent model. These authors speculated that significant helical content was not found in earlier simulations because of insufficient samplings and not found in the experiment because of the limited time resolution.[21,23] However, we argue that the significant α-helix content might be mainly due to the artifacts of the protein force fields. Detailed analysis of the

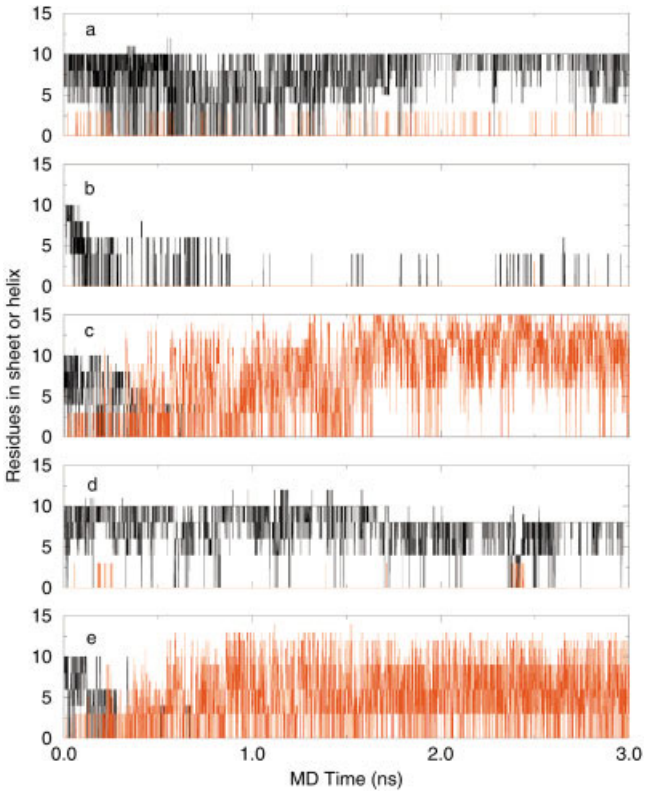

Fig. 5. Number of residues in α-helix (red) and β-sheet (black) format at various temperatures as determined by the program STRIDE,[40] (a) OPLSAA/SPC, (b) OPLSAA/SGB, (c) AMBER94/GBSA, (d) AMBER96/GBSA, and (e) AMBER99/GBSA. Both AMBER94/GBSA and AMBER99/GBSA show enormous α-helical content.

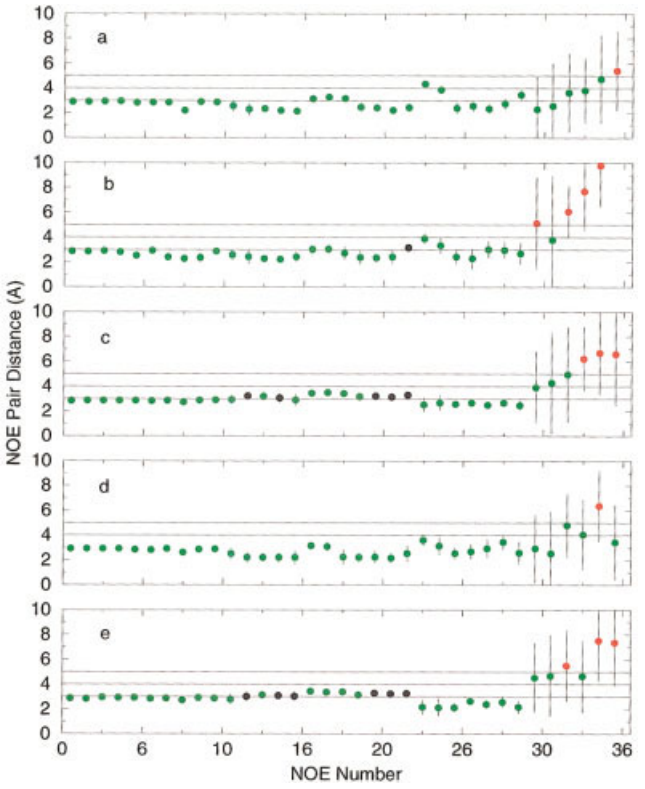

Fig. 6. Comparison of the simulation results with NMR measurements of NOE for various solvation models. (a) OPLSAA/SPC, (b) OPLSAA/SGB, (c) AMBER94/GBSA, (d) AMBER96/GBSA, and (e) AMBER99/GBSA. The distances are calculated by $R_{AVG} = \langle R_{HH}^{-6} \rangle^{-1/6}$ and averaged over the entire ensemble obtained from the replica-exchange method at 310 K. The error bars correspond to the variance around R_{AVG}. Three horizontal lines at 3, 4, and 5 Å are drawn to indicate the upbound for observing the strong, medium, and weak/very weak NOE signals. The agreement with the NOE strength is indicated by color: red for not observable in the NOE signal, black for observable in the NOE signal but with a wrong strength, and green for observable in the NOE signal and also with a correct strength.

number of β-sheet and α-helix residues during the folding process is performed with the program STRIDE[40] for all the five models.

Figure 5 shows the "trajectories" of the number of α-helix residues (colored red) and β-sheet residues (colored black) during the MD or HMC simulations. It is clear that both OPLSAA/SPC and OPLSAA/SGB simulations show very little α-helical content. The number of helical residues, including both α-helix and 3_{10}-helix, are ≤3. Also, only 1–2% of the conformations show any helical content. Furthermore, almost all of the helices we found are 3_{10}-helices near the original beta-turn (Residues 47–49). Very few conformations are found to have helix residues in places other than the original β-turn. All the other temperatures examined from 270 to 695 K show similar results. This minimal helix content is in contrast to recent results of Garcia et al.[23] and Pande et al.[21] but seems to agree with experiment better, because no evidence has been found in experiment for the significant helical content.[8–11] It is interesting to notice that the α-helical contents from OPLSAA/SGB agree very well with OPLSAA/SPC, even though the free energy contour maps are quite different. This, along with the following AMBER9x/GBSA results, suggests that the helical content is mainly determined by the protein force field, but not by the solvation model in this case. This might make sense since helix formation is

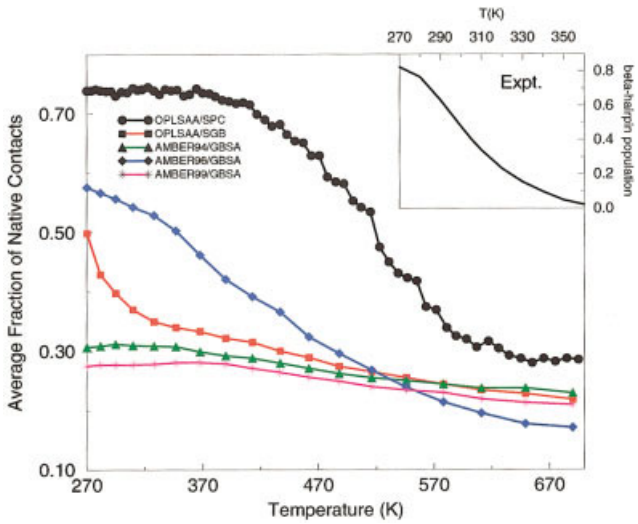

Fig. 7. Comparison of the temperature dependence of the β-hairpin population for various models. The β-hairpin population is estimated with the average fraction of native contacts. The experimental results[10] are also shown in the inset for comparison.

mainly driven by local hydrogen bonds (local in residue sequence) and largely determined by torsional potentials; although β-sheets involve global interactions (global in residue sequence) and both hydrophobic interactions and hydrogen bonds contribute. Thus, β-sheet formation might be more influenced by solvation models, whereas the formation of helices is largely determined by protein force fields.

On the other hand, both AMBER94/GBSA and AMBER99/GBSA turned the β-hairpin into an α-helix, as clearly shown in Figure 5(c,e). After some initial equilibration period, most of conformations show some α-helical residues. If we count the peptide as an α-helix if ≥4 residues are in the helical format, the helix population is 85% in AMBER94/GBSA and 59% in AMBER99/GBSA, generally consistent with the earlier estimates with the H state populations. Meanwhile, only very few residues were found to be in the β-sheet format in both AMBER94/GBSA and AMBER99/GBSA. These findings differ significantly from the results of OPLSAA/SPC and OPLSAA/SGB simulations. As speculated earlier,[16] this suggests that the helical content is mainly determined by the protein force field and less by the solvation model. Of course, it seems that the continuum solvent model makes it easier for the peptide to be in the α-helical format, because the corresponding explicit solvent simulation by Garcia et al. using AMBER94 and a TIP3P water model shows much less α-helical content. The reason for a greater α-helix tendency could be that the continuum solvent model GBSA further lowers the energy of the α-helical format relative to the β-hairpin format by overestimating the GB single energies (cf. Eq. 5) for partial charges because they are less overlapped in α-helix format than the β-sheet format. Further investigation might be worthwhile here. Overall, the large helical content in AMBER94 and AMBER99 force fields must be related to the artifacts of AMBER force fields as found previously as well.[23,41]

Interestingly, AMBER96/GBSA seems to have this problem fixed. Similar to OPLSAA/SPC, the number of helical residues, including both the α-helix and the 3_{10}-helix, is typically ≤3, and only <1% of the conformations exhibit helical content at 310 K. Thus, AMBER96/GBSA gives quite reasonable results on the α-helical content compared with explicit solvent simulation and experiment, and the problem of overestimation of the α-helix seems to have been fixed in AMBER96. It is worth noting, however, that in a very recent study[43] the AMBER96/GBSA fails to form an α-helix for a designed 20-residue mini-protein, called Trp-cage mini-protein.[44] The Trp-cage mini-protein has a short α-helix in residues 2–9 and a 3_{10}-helix in residues 10–12, which form a Trp-cage for residue Trp6 along with the rest of the mini-protein.[44] The AMBER96/GBSA does not form the α-helix,[43] whereas both AMBER94/GBSA and AMBER99/GBSA successfully formed this helix and also folded this mini-protein into a very reasonable structure[43,45] (OPLS/GBSA also folded this mini-protein to a reasonable structure by Pande et al.[46]). Thus, it seems that AMBER96 might have overcorrected the alpha helix tendency problem and will miss real alpha helices in some other proteins. Another evidence is that Garcia et al.[47] found that AMBER96 also fails to form an α-helix for the alanine-rich peptide *Ace–A_5(AAARA)$_3$–NMe* with explicit solvent TIP3P; on the other hand, the AMBER94 (with TIP3P water) predicts too much α-helical content for the same peptide, and a modified AMBER94 with torsion parameters for ϕ and ψ set to zero gives a much better agreement with experiment.[47] This indicates further optimizations of AMBER torsional parameters might be still needed to resolve this α-helix vs. β-sheet balance problem. Garcia et al.'s modification might serve as a *bare* potential that can be further optimized by perturbation.[47]

NOE Comparison

There are NOE measurements available for this β-hairpin alone in solution, and it is of interest to compare simulation results directly to the raw experimental data. The NMR experiment results are from Serrano group.[8,9] The NMR measurements provided a set of NOE constraints that the β-hairpin structure must satisfy. There are 35 unambiguous NOE proton pairs seen in NMR experiment (Fig. 2 in Ref. 8) with also the NOE strengths measured (s, strong; m, medium; w, weak; vw, very weak): 1. H_α-NH E42 (w); 2. H_α-NH T44 (w); 3. H_α-NH Y45 (w); 4. H_α-NH D46 (w); 5. H_α-NH D47 (m); 6. H_α-NH A48 (m); 7. H_α-NH T49 (m); 8. H_α-NH K50 (m); 9. H_α-NH T55 (w); 10. H_α-NH E56 (w); 11. H_α G41-NH E42 (m); 12. H_α E42-NH W43 (s); 13. H_α T44-NH Y45 (m); 14. H_α Y45-NH D46 (s); 15. H_α D46-NH D47 (s); 16. H_α D47-NH A48 (m); 17. H_α A48-NH T49 (m); 18. H_α T49-NH K50 (m); 19. H_α T51-NH F52 (m); 20. H_α T53-NH V54 (s); 21. H_α V54-NH T55 (s); 22. H_α T55-NH E56 (s); 23.NH Y45-NH D46 (vw); 24. NH D46-NH D47 (vw); 25. NH D47-NH A48 (m); 26. NH A48-NH T49 (m); 27. NH T49-NH K50 (m); 28. NH K50-NH T51 (w); 29. NH T55-NH E56 (vw); 30. H_α Y45-H_α F52 (vw); 31. H_α W43-H_α V54 (vw); 32. H_α K50-3H Y45 (vw); 33. H_α Y45-2H F52 (vw); 34. H_β Y45-5H F52 (vw); 35. H_β W43-4H F52 (vw). The question we try to address here is whether or not the calculated proton pair distances fall within the distance range of typical NOEs of the experimentally observed strength (2–3 Å for strong, 2–4 Å for medium, and 2–5 Å for weak and very weak NOEs). Figure 6 compares the calculated distances with the NOEs seen experimentally for all the five models. The distances are calculated by $R_{AVG} = \langle R_{HH}^{-6} \rangle^{-1/6}$ and averaged over the entire ensemble obtained from the replica-exchange method at 310 K. The error bars correspond to the variance around R_{AVG}. Three horizontal lines at 3, 4, and 5 Å are drawn to indicate the upbound for observing the strong, medium and weak/very weak NOE signals. The agreement with the NOE strength is also indicated by color in the figure following Pande et al.'s notation,[21] red for not observable in the NOE signal, black for observable in the NOE signal but with a wrong strength, and green for observable in the NOE signal and also with a correct strength. The NOE pairs 1–29 are proton pairs within the same residue or neighbor residues. Except for some mismatches for strong NOE signals in AMBER94/GBSA and AMBER99/GBSA, these 1–29 pairs are all visible in NOE signal in all five

models, as shown in Figure 6. The NOE pairs 30–35 are for protons in residues separated by many residues, which provides the real structural information for the folded β-hairpin. Both OPLSAA/SPC and AMBER96/GBSA predict these 6 proton pairs within or very close to 5 Å, except for pair 34, H_β Y45-5H F52, in AMBER96/GBSA, which shows a 6.3 ± 2.8 Å distance. This is consistent with the earlier finding that the hydrophobic core packing in the most populated structure of AMBER96/GBSA is not as good as the native structure. Although the other three models, OPLSAA/SGB, AMBER94/GBSA, AMBER99/GBSA, all show many pairs having distances much larger than 5 Å, which are beyond the NOE signal observation range. Particularly, the OPLSAA/SGB model shows a distance of 13.8 ± 2.9 Å for pair 35, H_β W43-4H F52 (not shown in Fig. 6 because it is off the scale of graph). This is related to the fact that the most populated structure in OPLSAA/SGB has the residue F52 expelled from the hydrophobic core. Pande et al.[21] also calculated these NOE distances using their folded ensembles from many short kinetics runs. The calculations were done using a 1-ns window after the β-hairpin is "folded." Even though we can not check the salt bridges in these "folded structures" from the data reported in the article,[21] it seems the agreement with NMR experiment for pairs 30–35 are not that good because most pairs show much larger distances than 5 Å, even though these are "folded ensemble structures" (Fig. 9 in Pande et al.[21]). These pair distances in their figure look very similar to those shown in our Figure 6(c,e) for AMBER94/GBSA and AMBER99/GBSA. If Pande et al. had used the entire ensemble, the results would be even worse. However, as we already know that the lowest free energy structures from AMBER94/GBSA and AMBER99/GBSA are very bad. Of course, Pande et al.[21] are mainly interested in the folding kinetics of the β-hairpin. As long as these short kinetics runs pass the initial lag phases into a single exponential region, the approach works fine.[48] The "folded structure" is used as an indicator to show the protein has been "folded" in these short kinetics runs[21]; thus, they may stop once a folded structure is found or the MD run is terminated. This is slightly different from the thermodynamics we are trying to address here. The not so good structures in Pande et al.'s simulation should be related to the continuum solvent model GBSA, as the authors also pointed out. It was found that the C_α protons between the hydrophobic core residues were particularly troublesome.[21] The strands of the folded hairpin are slightly out of plane, which results in a higher than expected separation between the C_α protons of the core residues. As mentioned above, if the entire ensemble had been used, the OPLS/GBSA results by Pande et al.[21] would be even worse, but we think the entire ensemble should be used, because the NMR experiment also "measures" the entire ensemble.

Temperature Dependence

Even though the explicit solvent OPLSAA/SPC model gives very reasonable results at low temperatures, the temperature dependence is not quite correct. The β-hairpin populations at higher temperatures are way too high, and the folding transition temperature is also way too high. This is found to be true with CHARMM and AMBER force fields as well with explicit solvent models.[27] This should not be too surprising given that most of the modern force fields are parameterized at room temperature. Nevertheless, we include the temperature dependence data here for all five models for completeness and also more importantly, to provide data for force field developers to improve the models.

The β-hairpin populations at various temperatures are calculated with the average fraction of native contacts and the results are compared with the experimental populations from the TRP fluorescence yield measurements.[10] Thirumalai et al.[26] have used the average fraction of native contacts to estimate the β-hairpin population, and here we follow the same approach. Figure 7 shows the comparison of the β-hairpin populations at various temperatures for the five different models. The experimental populations are shown in the inset for comparison. The fluorescence yield experiment shows a β-hairpin population of ~80% at a low temperature of 282 K, whereas the calculation shows a 74% population for OPLSAA/SPC, 43% for OPLSAA/SGB, 31% for AMBER94/GBSA, 57% for AMBER96/GBSA, 57% for AMBER96/GBSA, and 28% for AMBER99/GBSA. As expected, the β-hairpin population is seriously underestimated in all the implicit solvent models except AMBER96/GBSA, where it shows a decent 57% at 282 K. Although the explicit solvent model simulation predicts populations in reasonable agreement with experiment near the biological temperatures, it overestimates populations at higher temperatures.[27] The folding transition temperature is estimated to be about 470 K for OPLSAA/SPC, and 430 K for AMBER96/GBSA. Both are too high compared with the experiment. The transition temperatures for OPLSAA/SGB, AMBER94/GBSA, and AMBER99/GBSA are not very meaningful because there are no meaningful decays in the β-hairpin population with temperature. It should be pointed out that the population at very high temperatures, such as above 500 K, are not zero, particularly in OPLSAA/SPC case. This is because of the way the population is calculated.[26] Even at very high temperatures, the unfolded hairpin structures still have some native contacts locally (native contact defined as C_α–C_α distance < 6.5 Å[26] for nonadjacent residues). The higher populations in OPLSAA/SPC at high temperatures are also partly because the NVT ensemble used. The volume will increase at higher temperatures, which typically favors the unfolding. This will decrease the β-hairpin population at higher temperatures, but the populations near room temperature, which we care most, should not be affected much (preliminary NPT simulations support this, data not shown). As stated earlier, we listed these results here not for head-to-head comparisons with experiments because we should not expect them to be perfect at higher temperatures, but rather for force field developers to use these data to improve the parameters to include some kind of temperature dependence in the future.

Finally, let's briefly discuss the possible fixes. First, one might expect that the erroneous formation of salt bridges in continuum solvent models, particularly in OPLSAA/SGB, is probably exacerbated by the fact that counter ions are not included in implicit solvent models. In explicit solvent simulations, there are three counter ions (3 Na^+), which will somewhat neutralize the negatively charged residues (E42, D46, D47, and E56), so that these negative charges may be partially screened, thus reducing the direct electrostatic interactions that would lead to a salt bridge. However, we think this is probably a small effect because the Na^+ are free ions in solution. It should be pointed out that such salt effects can be *explicitly* included in the Poisson-Boltzmann equations,[49] but are not easily to be formulated rigorously in GB-type models. Thus the fix has to be either included as a Debye-Huckel screening term in the pair-energy as used by Case et al.[29] or implicitly included in other parameters such as atomic radii for charged ions like $NH3^+$ and COO^- as used by Friesner et al.[14] Second, and more importantly, in the continuum solvent models the charged groups might not be screened enough. These charged groups interact with each other through the GB pair-energy (Eq. 6) and the normal pairwise Coulombic interaction as well as implicitly through the GB single energy (Eq. 5), which of course is mainly for the interaction between the charges and the surrounding water.

A small solute dielectric constant (1.0–4.0 by Friesner et al.[14] and Case et al.[29]) will typically result in a large polar electrostatic energy for proteins (or so-called reaction field energy,[14] which includes both the single and pair energies), which the charged residues normally contribute most. This large polar electrostatic energy can easily overrun the nonpolar cavity energy, which is typically estimated from the solvent accessible surface area and is independent of the dielectric constant. The salt-bridge effect is thus largely due to the imbalance of these two polar and nonpolar energy terms. It is also found that in SGB the gain of the pairwise electrostatic energies when two charge groups approach each other is not offset enough by the loss of the single energies, which amplifies the salt-bridge effects. Thus it seems more favorable for oppositely charged residues to come close together instead of being hydrated as they would be in explicit water. One way to fix this is to invoke a stronger dielectric screening (a much larger dielectric constant) for charged residues as suggested by Warshel *et al.* in another context.[50] There is some experimental evidence for this too.[38] Another possible approach to fix this problem is to introduce a penalty function between oppositely charged residues as was suggested by Jacobson and Friesner in connection with their loop geometry optimizations (personal communication). Of course the introduction of a larger dielectric constant for charged residues similarly gives rise to a penalty, albeit a different one than that introduced by Jacobsen and Friesner. Some initial testing of the penalty function indicates that it partially fixes the problem. We will address this question in more detail in a separate publication because it involves a complete refitting of the model.

Finally, the AMBER9x/GBSA results clearly indicate that the implicit solvent model parameters, such as atomic radii, are strongly coupled with the gas phase force field parameters, such as the backbone torsional parameters. An appropriate balance is needed between these two in order to have an accurate model. It seems that the new AMBER99 force field still does not completely fix the infamous α-helix overestimation problem. There will be an even newer model AMBER02, param 2002, coming out with polarizability. We are going to test it once it is fully published. We have noticed that Garcia et al.[47] have modified the AMBER94 force field when studying the analinerich peptides *Ace–*Ala_{21}*–NMe* and *Ace–*$A_5(AAARA)_3A$*–NMe*. The torsion parameters for ϕ and ψ were set to zero. The modified AMBER94 gives a much better agreement with experiment in the α-helical content.[47] This modification might also serve as a starting point for further optimization.[47]

CONCLUSION

The free energy landscapes of a β-hairpin folding in water using both explicit and implicit solvent models are studied in this article. Five different models are explored: the explicit solvent model: OPLSAA/SPC, and implicit solvent models: OPLSAA/SGB, AMBER94/GBSA, AMBER96/GBSA, and AMBER99/GBSA. A highly parallel replica exchange method consisting of 64 and 18 replicas for explicit and implicit solvent models, respectively, has been used with temperature spanning from 270 to 695 K. The major conclusions are summarized in the following.

Surprisingly, the free energy landscapes of the implicit solvent models are quite different from that of the explicit solvent model, except for AMBER96/GBSA. Only AMBER96/GBSA gives a somewhat reasonable free energy contour map when compared with the explicit solvent model. All the other models, OPLSAA/SGB, AMBER94/GBSA, and AMBER99/GBSA give significantly deviant free energy contour maps. Even though we can not extrapolate the exact free energy contour map from the limited experimental data, we have some evidence that the explicit solvent model gives a reasonable free energy contour map near biological temperature, for example, it predicts the correct native structure; it gives a reasonable population at low temperatures; it agrees with experiment on α-helical content; and it reproduces all NOEs from NMR experiment. On the other hand, those bad implicit solvent models, OPLSAA/SGB, AMBER94/GBSA and AMBER99/GBSA all show heavily overweighted non-native states and predict non-native structures to be the lowest free energy structures. The clustering analysis of the most heavily populated states reveals more details about the problems in implicit solvent models. All implicit solvent models show erroneous salt-bridge effects, particularly in SGB, where the overly strong salt bridges between charged residues amplify the imbalance between the polar electrostatic interaction and the non-polar hydrophobic interaction, and result in a most heavily populated structure with one hydrophobic residue F52 expelled from the hydrophobic core. In GBSA models with AMBER94 and AMBER99,

the implicit solvent model dramatically increases the α-helical content, ~70–80% in AMBER94 and 60% in AMBER99, which is much larger than the 15–20% of α-helical content found previously by Garcia et al. with AMBER94 and an explicit solvent model TIP3P. This indicates that GBSA greatly enhances the α-helix tendency in AMBER force fields. It is also found that the presence or absence of helical content seems to be mainly determined by the protein force fields, but not by solvation models. Furthermore, the β-hairpin population is underestimated in all implicit solvent models at low temperatures; for example, at 282K it is estimated to be only 43% in OPLSAA/SGB, 31% in AMBER94/GBSA, 57% in AMBER96/GBSA, and 27% in AMBER99/GBSA, compared with 74% in the explicit solvent OPLSAA/SPC and about 80% in experiment.

These results indicate that people should be cautious when applying these GB-type continuum solvent models to large-scale protein-folding simulations, even though they are very successful in many fields, such as surface electrostatic potentials, solvation free energies, and ligand-receptor bindings.[29–31] The balance between the polar electrostatic and nonpolar cavity interaction might not be well preserved for charged systems when large conformational changes are involved such as protein folding. This imbalance between polar and nonpolar terms in implicit solvent models might be also related to the recent findings of the absence of the desolvation free energy barriers in implicit solvent models, such as GB and EEF1.[51–53] We have suggested several possible fixes for this problem, and work has been initiated along some of these lines.

ACKNOWLEDGEMENTS

I thank Jed Pitera for helping with the AMBER simulations, and Bruce Berne, William Swope, and Huafeng Xu for many useful discussions. I would also like to thank Angel Garcia and Vijay Pande for helpful comments.

REFERENCES

1. Fersht AR. Structure and mechanism in protein science. New York: W.H. Freeman and Company 1999.
2. Dill KA, Chan HS. From Levinthal to pathways to funnels. Nat Struc Biol 1997;4:10.
3. Brooks CL, Gruebele M, Onuchic JN, Wolynes PG. Chemical physics of protein folding. Proc Natl Acad Sci USA 1998;95:11037.
4. Zhou Y, Karplus M. Interpreting the folding kinetics of helical proteins. Nature 1999;401:400.
5. McCammon JA, Wolynes PG, Eds. Current opinion in structural biology. London: Current Biology Press; 2002.
6. Wales DJ, Scheraga HA. Global optimization of clusters, crystals and biomolecules, Science 1999;285:1368.
7. Snow CD, Nguyen H, Pande VS, Gruebele M. Absolute comparison of simulated and experimental protein-folding dynamics. Nature 2002;420:102.
8. Blanco FJ, Rivas G, Serrano L. A short linear peptide that folds in a native stable β-hairpin in aqueous solution. Nat Struct Biol 1994;1:584.
9. Blanco FJ, Serrano L. Folding of protein g b1 domain studied by the conformational characterization of fragments comprising its secondary structure elements. Eur J Biochem. 1995;230:634.
10. Munoz V, Thompson PA, Hofrichter J, Eaton WA. Folding dynamics and mechanism of β-hairpin formation. Nature 1997;390:196.
11. Munoz V, Henry ER, Hofrichter J, Eaton WA. A statistical mechanical model for β-hairpin kinetics. Proc Natl Acad Sci USA 1998;95:5872.
12. Honig B, Nicholls A. Classical electrostatics in biology and chemistry. Science 1995;268:1144.
13. Still W, Clark Tempczyk Anna, Hawley Ronald C, Hendrickson Thomas. Semi-analytical treatment of solvation for molecular mechanics and dynamics. J. Am Chem Soc 1990;112:6127.
14. Ghosh A, Rapp CS, Friesner RA. Generalized born model based on a surface integral formulation. J Phys Chem 1998;102:10983.
15. Lazaridis T, Karplus M. Effective energy function for proteins in solution. Proteins 1999;35:133.
16. Zhou R, Berne BJ. Can a continuum solvent model reproduce the free energy landscape of a β-hairpin folding in water? Proc Natl Acad Sci 2002;99:12777.
17. Cornell W, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A second generation force field for the simulation of proteins, nucleic acids, and orgnaic molecules. J Am Chem Soc 1995;117:5179.
18. Kollman PA, Dixon R, Cornell W, Fox T, Chipot C, Pohorille A. The development/application of a 'mini-malist' organic/biochemical molecular mechanic force field using a combination of ab initio calculations are experimental data in computer simulation of biomolecular systems. Computer simulation of Biomolecular Systems, Eds. A. Wilkinson, P. Weiner and W.F. van Gunsteren 1997;3:83.
19. Wang J, Cieplak P, Kollman PA. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J Comput Chem 2000;21:1049.
20. Pande VS, Rokhsar DS. Molecular dynamics simulations of unfolding and refolding of a β-hairpin fragment of protein g. Proc Natl Acad Sci USA 1999;96:9062.
21. Zagrovic B, Sorin EJ, Pande VS. β-hairpin folding simulation in atomistic detail. J Mol Biol 2001;313:151.
22. Dinner AR, Lazaridis T, Karplus M. Understanding β-hairpin formation. Proc Natl Acad Sci USA 1999;96:9068.
23. Garcia AE, Sanbonmatsu KY. Exploring the energy landscape of a β hairpin in explicit solvent. Proteins 2001;42:345.
24. Roccatano D, Amadei A, Nola A Di, Berendsen HJ. A molecular dynamics study of the 41–56 β-hairpin from b1 domain of protein g. Protein Sci 1999;10:2130.
25. Kolinski A, Ilkowski B, Skolnick J. Dynamics and thermodynamics of β-hairpin assembly: insights from various simulation techniques. Biophys J 1999;77:2942.
26. Klimov DK, Thirumalai D. Mechanism and kinetics of β-hairpin formation. Proc Natl Acad Sci USA 2000;97:2544.
27. Zhou R, Berne BJ, Germain R. The free energy landscape for β-hairpin folding in explicit water. Proc Natl Acad Sci USA 2001;98:14931.
28. Zhou R, Harder E, Xu H, Berne BJ. Efficient multiple time step method for use with ewald and particle mesh ewald for large biomolecular systems. J Chem Phys 2001;115:2348.
29. Onufriev A, Bashford D, Case DA. A modification of the generalized born model suitable for macromolecules. J Phys Chem B 2000;104:3712.
30. Gallicchio E, Zhang LY, Levy RM. The sgb/np hydration free energy model based on the surface generalized born solvent reaction field and novel non-polar hydration free energy estimators. J Comp Chem 2002;23:517.
31. Zhou R, Friesner RA, Ghosh Avijit, Rizzo RC, Jorgensen WL, Levy RM. New linear interaction method for binding affinity calculations using a continuum solvent model. J Phys Chem B 2001;105: 10388.
32. Zhou R, Berne BJ. Smart walking: A new method for boltzmann sampling of protein conformations. J Chem Phys 1997;107:9185.
33. Duane S, Kennedy AD, Pendleton BJ, Roweth D. Hybrid Monte Carlo. Phys Lett B 1987;195:216.
34. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett 1999;314:141.
35. Weiser J, Shenkin PS, Still WC. Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). J Comp Chem 1999;20:217.
36. Bashford D, Case DA. Generalized born models of macromolecular solvation effects. Annu Rev Phys Chem 2000;51:129.
37. Schutz CN, Warshel A. What are the dielectric 'constants' of proteins and how to validate electrostatic models. Proteins 2001; 44:400.
38. Garcia-Moreno E, Dwyer JJ, Gittis AG, Lattman EE, Spenser DS,

Stites WE. Experimental measurement of the effective dielectric in the hydrophobic core of a protein. Biophys Chem 1997;64:211.
39. Garcia AE. Large-amplitude nolinear motions in proteins. Phys Rev Lett 1992;68:2696.
40. Frishman D, Argos P. Knowledge-based secondary structure assignment. Proteins 1995;23:566.
41. Beachy M, Chasman D, Murphy R, Halgren T, Friesner R. Accurate ab initio quantum chemical determination of the relative energetics of peptide conformations and assessment of empirical force fields. J Am Chem Soc 1997;119:5908.
42. Connolly ML. Analytical molecular surface calculation. J Appl Cryst 1983;16:548.
43. Pitera J, Swope W. Getting to the folded state: parallel-replica simulations of TRP-cage. Nat Struc Biol 2003, Submitted for publication.
44. Neidigh JW, Fesinmeyer RM, Andersen NH. Designing a 20-residue protein. Nat Struc Biol 2002;9:424.
45. Simmerling C, Strockbine B, Roitberg AE. All-atom structure prediction and folding simulations of a stable protein. J. Am Chem Soc 2002;124:11258.
46. Snow CD, Zargovic B, Pande VS. The TRP cage: Folding kinetics and unfolded state topology via molecular dynamics simulations. J Am Chem Soc 2002;124:14548.
47. Garcia AE, Sanbonmatsu KY. Alpha-helical stabilization by side chain shielding of backbone hydrogen bonds. Proc Natl Acad Sci USA 2002;99:2782.
48. Fersht AR. On the simulation of protein folding by short time scale molecular dynamics and distributed computing. Proc Natl Acad Sci USA 2002;99:14122.
49. Chen SW, Honig B. Monovalent and divalent salt effects on electrostatic free energies defined by the nonlinear poisson-boltzmann equation: Application to DNA reactions. J Phys Chem B 1997;101:9113.
50. Burykin A, Schutz CN, Villa J, Warshel A. Simulations of ion current in realistic models of ion channels: the kcsa potassium channel. Proteins 2002;47:265.
51. Masunov A, Lazaridis T. Potentials of mean force between ionizable amino acid side chains in water. J Am Chem Soc 2003;125:1722.
52. Cheung MS, Garcia AE. Onuchic, Protein folding mediated by solvation: water expulsion and formation of the hydrophobic core occur after the structural collapse. Proc Natl Acad Sci USA 2002;99:685.
53. Kaya H, Chan HS. Solvation effects and driving forces for protein thermodynamics and kinetic cooperativity: how adequate is native-centric topological modeling? J Mol Biol 2003;326:911.