COMMENTS

Comment on "Can a Continuum Solvent Model Reproduce the Free Energy Landscape of a β-Hairpin Folding in Water?" The Poisson−Boltzmann Equation

Ruhong Zhou,[†,‡] Goran Krilov,[‡] and B. J. Berne*[,†,‡]

Computational Biology Center, IBM Thomas J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, New York 10598, and Department of Chemistry, Columbia University, New York, New York 10027

Received: December 11, 2003; In Final Form: February 18, 2004

In a previous paper[1], we found that the folding free energy landscape of a β-hairpin using the widely used generalized Born (GB) implicit solvent model is quite different from that using an explicit solvent model, namely the SPC model, with the same protein OPLSAA force field. It is of great interest to see what the folding free energy landscape will be in a more rigorous continuum solvent Poisson−Boltzmann (PB) model. In the GB model, some non-native states are heavily overweighted and the lowest free energy state is no longer the native state. An overly strong salt-bridge between charged residues (D46, D47, E56, and K50) was found to be responsible for this behavior in the implicit solvent model. On the other hand, the explicit solvent model reproduces the experimental results quite well near the biological temperatures, even though the temperature dependence of the β-hairpin population is not correct at high temperatures. In a follow-up study, we also examined another implementation of the GB model parameterized with AMBER force fields to see whether the problem arises from the specific implementation or the associated parameterization, however, similar erroneous results were found.[2] Thus, it appears that the problem is common to different GB models. The balance between polar electrostatic and nonpolar cavity interactions seems to be shifted toward favoring electrostatic interactions over hydrophobic interactions in some residues, causing overly strong Coulombic interactions between some charged groups and resulting in an over-weighting of non-native structures.

The generalized Born (GB) model is an empirical method that estimates the electrostatic response from the polarization of the dieletric medium. Thus, it is important to ascertain if an implicit solvent model based on the more rigorous Poisson−Boltzmann (PB) solver, such as Delphi,[3] would reproduce experiment as well as the explicit solvent simulations, even though PB models in principle do not provide a complete electrostatic description either, unless one knows the dielectric constant at any position in the protein. In this context the more accurate PB model can be used to examine where the problem arises: does it spring from the polar term, the nonpolar term, or both? As in the GB model, the PB model treats the solute (protein) as some fixed charge distribution contained in a region of low dielectric constant that is surrounded by a high dielectric continuum solvent. Free salt ions may be present in the solvent and are assumed to be Boltzmann-distributed at equilibrium. The problem then reduces to solving the Poisson−Boltzmann equation (which in principle gives a complete electrostatic description of the mixed dielectric system):

$$\nabla(\epsilon\nabla\phi) = -\frac{4\pi\rho}{kT} + 4\pi C \sinh(\phi e) \qquad (1)$$

where ϵ is the dielectric constant, ϕ is the electrostatic potential, and ρ is the solute charge density. $C = 2eI/k_BT$ where I is the ionic concentration of the salt dissolved in the solvent. In the present work, we combined a modified Delphi model[3] (from Honig's group) with a molecular dynamics method devised to calculate the free energy landscape of a small protein. The atomic radii used to define the dielectric boundary surface of the PB model are derived from the van der Waals radii in the OPLSAA force field, and the fixed charge distribution of the protein was constructed from the OPLSAA partial atomic charges. The nonpolar cavity term is measured by the usual solvent accessible surface areas with a surface tension of 5.0 cal/mol/Å^2, which is the same as used in previous GB studies.[1,2]

Our molecular dynamics approach, based on the replica exchange method (REM), is used to sample the conformational space of the β-hairpin, the C-terminus β-hairpin of protein G, as before. This protein subunit contains 16 residues and 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. A total of 18 replicas are simulated with the same temperature range as in the previous studies[1,2] (spanning the interval from 270 to 690 K). A dielectric constant of 2.0 is used for the protein region, while the aqueous solvent region is assigned a dielectric constant of 80.0. It should be noted that assigning a dielectric constant for proteins can be very tricky. As Schutz and Warshel[5] and others[4,3] have pointed out, one can justify using dielectric constants ranging from 1 to 4, and even higher, depending on the details of the model. The assumption that the protein can be treated as a single dielectric continuum is likely to be problematic altogether.[5] Each replica is run for 3.0 ns for data collection. The replica exchanges are attempted every 2.0 ps, and protein configurations are saved every 80 fs, giving a total of 0.675 million configurations.

The free energy landscape is obtained through a histogram analysis on the probability distribution.[1] Figure 1 shows the comparison of the free energy contour maps at 310 K computed using the explicit solvent model (Figure 1a) and the continuum solvent PB model (Figure 1b). The free energy is plotted against the two reaction coordinates used previously,[1,2] i.e., the number of β-strand hydrogen bonds (N_{HB}^{β}) and the radius of gyration of the hydrophobic core (R_g^{core}). N_{HB}^{β} is defined as the number of backbone−backbone hydrogen bonds excluding the two at the turn of the hairpin.[1] R_g^{core} is the radius of gyration of the side-chain atoms on the four hydrophobic residues, W43, Y45, F52, and V54. Interestingly, the free energy contour map from the PB model is still significantly different from that of the

* Corresponding author.
† IBM Thomas J. Watson Research Center.
‡ Columbia University.

10.1021/jp037812c CCC: $27.50
Published on Web 05/04/2004

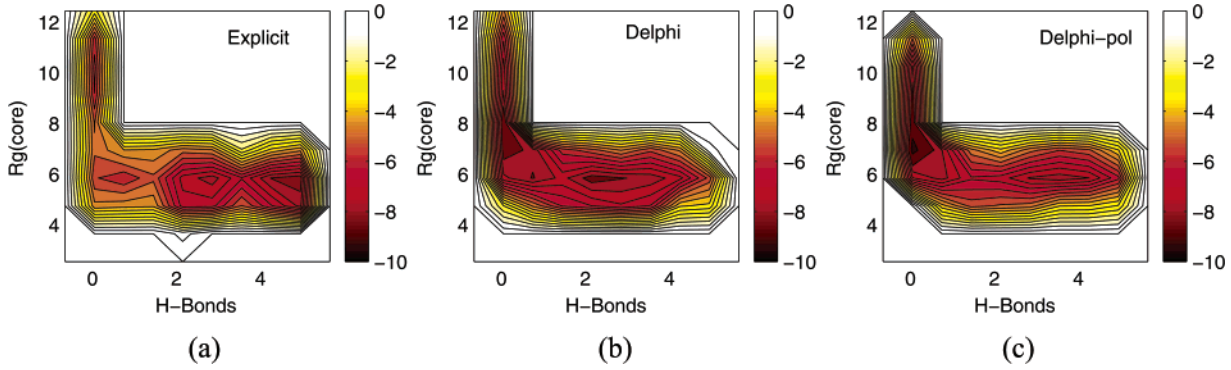

Figure 1. Comparison of the free energy versus the number of β-sheet H-bonds N_{HB}^{β} and the hydrophobic core radius gyration R_g^{core} at 310 K: (a) from the explicit solvent model, (b) from the implicit solvent PB model, and (c) from the electrostatic term only PB model (without the cavity term). 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 120.0 degree. The free energy is in units of RT, and contours are spaced at intervals of 0.5 RT.

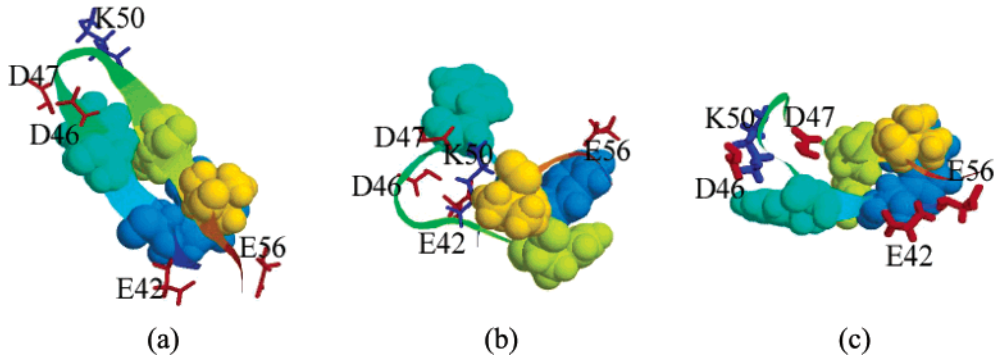

Figure 2. Comparison of the representative structures from free energy landscape: (a) the lowest free energy structure from the explicit solvent, (b) the lowest free energy structure (H state) from the implicit solvent PB model, and (c) the folded state structure from the implicit solvent PB model. The 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 structures from PB model show very different features from the explicit solvent structure, see text for details.

explicit solvent, even though it shows improvement over that of the GB model.[1] A few noticeable differences include: (1) the folded state in the PB model ($N_{HB} \approx 2-3$ and $R_g^{core} \approx 5.8$ Å) has a smaller number of the native beta-strand hydrogen bonds, and it is not the lowest free energy state; (2) the so-called H state (no beta-strand hydrogen bonds, but hydrophobic core formed, at least partially) has the lowest free energy. Its free energy is lower by approximately 0.4 kcal/mol (0.7 RT) than that of the folded state. Its radius of gyration for the hydrophobic core is slightly larger than that of the explicit solvent, $R_g^{core} \approx 7.2$ Å vs $R_g^{core} \approx 5.8$ Å. In our previous simulation with the GB model, however, the H state was found to be 2.9 kcal/mol lower than the native state. Thus, the PB model does seem to perform better than the GB model, even though the improvement is still not sufficient to predict the correct lowest free energy structure.

We next examined the representative structures from the lowest free energy state in the PB model in order to understand why the continuum solvent model favors non-native structures. Using clustering, we found representative structures in each free energy basin. Figure 2 shows the comparison of the lowest free energy structures from both the explicit solvent (Figure 2a, folded state) and implicit (Figure 2b, H state) models. The lowest free energy structure from the explicit solvent model mimics the native structure closely. Two interesting observations emerge from the comparison of the two structures: (1) the hydrophobic residue Y45 is expelled from the hydrophobic core in the PB model, while it is well packed with the other three hydrophobic residues (W43, F52, V54) in the native structure or the structure from the explicit solvent model. In other words, the four hydrophobic residues form a well-packed core in the explicit solvent but not in the continuum solvent; (2) in explicit solvent the side chains of charged residues are fully solvated, while in the continuum solvent PB 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 N-terminal end residue E42 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 N-terminal end (E42), into closer contact with residue K50 thereby expelling the hydrophobic residue Y45 from the hydrophobic core. These structural features are surprisingly similar to those of the GB model, except that in the GB model the hydrophobic residue F52 was expelled in order to make a closer contact between the residues E56 and K50.[1] Figure 2c also shows a representative structure in the folded state of the PB model. Again, this structure shows a tendency for charged groups to get closer; for example, the loop of the β-hairpin is bent more significantly in order for the residue K50 to get closer to the other four negatively charged residues. These results suggest that the balance between the electrostatic interactions and the hydrophobic interactions is shifted toward favoring the electrostatic interactions. The strong electrostatic interactions between the charged residues (salt bridges) overwhelm the local hydrophobic interactions.

Does the problem with the continuum model for the β-hairpin spring from the polar electrostatic term, the nonpolar cavity term, or both? To answer this question, and to explore possible routes to fixing the problem, we have calculated the free energy contour

map for the PB model using only the electrostatic interactions (i.e., the surface tension set to zero), as shown in Figure 1c. Clearly, without the cavity term the result is worse, since the H state will be even more populated and will have a free energy 0.9 kcal/mol (1.5 RT) lower than that of the folded state. Interestingly, the PB model with only electrostatic contributions is still better than the full GB model (see Figure 1b in ref 1). Thus, it appears that the GB model (at least in the surface GB implementation[1]) overestimates the electrostatic interactions between charged groups more than the PB model does. On the other hand, if we increase the cavity term by a factor of 10 (i.e., set surface tension to be 50 cal/mol/Å^2, data available upon request), the free energy of the H state still remains about 0.4−0.5 kcal/mol lower than that of the folded state, while the unfolded structures now assume a more "globular" shape because the stronger cavity term favors minimizing the surface area. Thus, the problem cannot be easily fixed by uniformly increasing the cavity term. In the previous study with the GB model, we tried increasing the dielectric constant of the β-hairpin, from 1.0, 2.0, to 4.0, to screen the electrostatic interactions, however this approach could not fix the problem either.[1] It follows that for both the PB and GB models the problem cannot be easily fixed by uniformly increasing the nonpolar term or decreasing the polar term. To fix the problem, it may suffice to devise a more accurate nonpolar cavity term, such as the one recently proposed by Levy's group,[6] combined with a stronger dielectric screening (a much larger dielectric constant) of the Coulomb interaction between charged residues as was suggested previously in another context as well.[7,8] In our opinion, the assignment of a uniform dielectric constant inside proteins is an unjustified assumption, particularly for proteins with heavy charged residues, such as the β-hairpin. Even if one can fix the problem in this β-hairpin system or any other systems by fitting more adjustable parameters, such as atomic radii and/or residue or atom based dielectric constants, it might still be an open question as to how portable these parameters are. Also, as pointed out recently,[9,10] the continuum solvent model, either GB or PB, has no structural features (no solvent discretness), thus, it will not be able to predict solvent discretness effects,[9] and it will not be able to predict dewetting phenomena either.[10] However, given the large computational savings of implicit solvent models for protein folding simulations, it is important to develop better continuum solvent models.

Acknowledgment. We thank Barry Honig for sending us the Delphi code for the Poisson−Boltzmann continuum solvent model, and Emilio Gallicchio and Ray Luo for helpful discussions. This work was supported in part by NIH Grant (GM4330) to B.J.B.

References and Notes

(1) Zhou, R.; Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12777.
(2) Zhou, R. *Proteins* **2003**, *53*, 148.
(3) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144.
(4) Simonson, T. *Curr. Opin. Struct. Biol.* **2001**, *11*, 243.
(5) Schutz, C. N.; Warshel, A. *Proteins* **2001**, *44*, 400.
(6) Gallicchio, E.; Zhang, L. Y.; Levy, R. M. *J. Comput. Chem.* **2002**, *23*, 517.
(7) Burykin, A.; Schutz, C. N.; Villa, J.; Warshel, A. *Proteins* **2002**, *47*, 265.
(8) Garcia-Moreno, E.; Dwyer, J. J.; Gittis, A. G.; Lattman, E. E.; Spenser, D. S.; Stites, W. E. *Biophys. Chem.* **1997**, *64*, 211.
(9) Papazyan, A.; Warshel, A. *J. Phys. Chem. B* **1998**, *102*, 5348.
(10) Huang, X.; Margulis, C. J.; Berne, B. J. *Proc. Nat. Acad. Sci. U.S.A.* **2003**, *100*, 11953.