PROTEINS: Structure, Function, and Genetics 18:119–132 (1994)

Evaluation of the Conformational Free Energies of Loops in Proteins

Kenneth C. Smith and Barry Honig
Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032

ABSTRACT **In this paper we discuss the problem of including solvation free energies in evaluating the relative stabilities of loops in proteins. A conformational search based on a gas-phase potential function is used to generate a large number of trial conformations. As has been found previously, the energy minimization step in this process tends to pack charged and polar side chains against the protein surface, resulting in conformations which are unstable in the aqueous phase. Various solvation models can easily identify such structures. In order to provide a more severe test of solvation models, gas-phase conformations were generated in which side chains were kept extended so as to maximize their interaction with the solvent. The free energies of these conformations were compared to that calculated for the crystal structure in three loops of the protein *E. coli* RNase H, with lengths of 7, 8, and 9 residues. Free energies were evaluated with a finite difference Poisson–Boltzmann (FDPB) calculation for electrostatics and a surface area-based term for nonpolar contributions. These were added to a gas-phase potential function. A free energy function based on atomic solvation parameters was also tested. Both functions were quite successful in selecting, based on a free energy criterion, conformations quite close to the crystal structure for two of the three loops. For one loop, which is involved in crystal contacts, conformations that are quite different from the crystal structure were also selected. A method to avoid precision problems associated with using the FDPB method to evaluate conformational free energies in proteins is described.**

Key words: electrostatics, protein conformation, DelPhi, hydrophobicity, RNase H

INTRODUCTION

The accurate prediction of conformational and binding free energies of molecules in aqueous solution has been an elusive goal for some time. Recently, considerable effort has been devoted to this problem, driven in large part by advances in the treatment of solvent contributions to free energy. The effects of water on conformation include dielectric and ionic strength screening of the pairwise Coulombic term, solvation forces acting on individual charges, solute–solvent van der Waals interactions, and the forces due to cohesive water–water interactions. Simulations involving an atomic level description of the solvent can in principle account for all of these effects, but at present they are too time consuming to be feasible when a large number of conformations must be considered. In addition, they are not yet able to treat ionic strength effects, which can be quite important for highly charged systems.

Macroscopic solvent models offer a possible alternative if the loss of microscopic detail is not accompanied by a concomitant loss in accuracy. However, evidence has accumulated which suggests that continuum solvent models are, at the least, comparable in accuracy to explicit models in the calculation of solvation free energies.[1–3] This has led to a number of recent attempts to combine a molecular mechanics treatment of the solute with a continuum treatment of the solvent.[4–9] In this paper we consider the implicit assumptions that underlie such hybrid models, and we describe a physically meaningful computational approach which makes it possible to evaluate solvation, conformational, and binding free energies. Although the approach is quite general, a preliminary application is made in this work to the prediction of the conformation of loops in proteins.

Several methods are currently being used to predict loop conformations. One approach is to generate, through random[10–12] or systematic[13–18] searches, a large number of possible structures for the loop, and to use energetic or other criteria to choose the best conformations. Related Monte Carlo search techniques are also under development.[19] Another method is to search a library of X-ray crystal structures[20] to find fragments of specified lengths which could connect the residues adjacent to either end of the loop. A related method, which has so far been restricted primarily to antibodies, is the

Received June 29, 1993; revision accepted October 8, 1993.
Address reprint requests to Dr. Barry Honig, Department of Biochemistry and Molecular Biophysics, Columbia University, 630 W. 168th St., New York, NY 10032.

canonical structure approach developed by Chothia, Lesk, and co-workers.[21,22] Using criteria based on loop length and key residues in the sequence, they have been successful in predicting the conformation of hypervariable loops from the most appropriate canonical structure. In any prediction scheme, it would clearly be highly desirable to have a method that is capable of distinguishing stable from unstable loop conformations, or more generally, correctly folded from "misfolded" structures.[23]

Toward this goal of improving energy based conformation selection, we focus on the integration of an atomic level force field with a continuum solvent model. The two must be coupled so as to ensure that all relevant interactions are accounted for, either implicitly or explicitly. In addition, "double counting" must be avoided. As a consequence, the solvation model that is adopted will be dependent on the molecular mechanics force field used to describe the solute. In the full solvation model we will describe, all electrostatic contributions to solvation free energies are obtained from finite difference solutions to the Poisson–Boltzmann equation (the FDPB method) while nonpolar solvation contributions are accounted for with surface area terms.

After the methodology is introduced, we discuss precision issues associated with the application of the FDPB method to conformational searches. The method is tested by calculating the conformational free energy of three loops in the protein *E. coli* ribonuclease H[24] (RNase H). We compare also the use of an alternative method for calculating solvation energies, based on the atomic solvation parameter techniques of Eisenberg and co-workers.[1,7] Overall, the addition of solvation terms to an in vacuo force field can significantly improve the discrimination of correct from incorrect loop conformations. Directions for the improvement of current conformational search and free energy evaluation procedures are considered in the discussion.

THEORY AND METHODS

Conformational Free Energies

The conformational free energy of a molecule in solution, $\Delta G_{conf}(s)$, can be written as

$$\Delta G_{conf}(s) = \Delta G_{conf}(g) + \Delta G_{solv} \quad (1)$$

where $\Delta G_{conf}(g)$ is the conformational energy in the gas phase and ΔG_{solv} is the gas to water solvation free energy. $\Delta G_{conf}(g)$ can be obtained from a quantum mechanical calculation or, alternatively, from a molecular mechanics force field. These involve sums of terms over all the internal coordinates of the molecule, plus a sum over all pairs of nonbonded interactions between atoms i and j separated by the distance r_{ij} (van der Waals plus Coulombic).

It is well known that the value of the various parameters used in such force fields are interdependent, for example atomic charges and dielectric constants (ε) must be chosen so as to yield good hydrogen bond energies. Most current force fields are not specifically designed to produce accurate gas phase conformational energies. Indeed ε is frequently taken to be larger than 1 (or is assumed proportional to distance), in part to account for the effects of solvent screening. In this case, one obtains a "pseudo"-gas phase force field denoted here by $\Delta G_{conf}(g, \varepsilon > 1)$. The use of such a force field would at least formally invalidate Eq. (1), since if solvent screening is accounted for in $\Delta G_{conf}(g)$, it should not be calculated again in ΔG_{solv}. A further issue concerns the internal dielectric constant of the solute. A value of 2 accounts approximately for electronic polarizability[3] and, when used in a solvation calculation, yields changes in solute dipole moment that accompany phase transfer. For this reason, in this paper we have used a dielectric constant of 2 to calculate $\Delta G_{conf}(g)$ when the electrostatic solvation energy is to be added.

The total solvation free energy of a given conformation can be written as

$$\Delta G_{solv} = \Delta G_{es} + \Delta G_{np}. \quad (2)$$

ΔG_{es} is the difference in electrostatic free energy of the solute in the two phases. It can be obtained directly from FDPB calculations as has been described previously.[2,25,26]

ΔG_{np} is the gas to water free energy of transfer of a hypothetical molecule identical in size and shape to the solute, but which is completely nonpolar. The free energy of inserting a nonpolar solute into water is frequently written in the form

$$\Delta G_{np} = \Delta G_{cav} + \Delta G_{dis} \quad (3)$$

where the first term is the free energy to open a cavity in the solvent and the second is the dispersion term, the free energy to fill the cavity with neutral atoms identical in size and shape to those of the solute. ΔG_{np} is usually calculated from the free energy of transfer of alkanes, which in turn can obtained from the solubilities of alkanes in water. It is generally assumed that ΔG_{np} is proportional to accessible surface area and is given by

$$\Delta G_{np} = \gamma_{vw} A_T. \quad (4)$$

γ_{vw}, the vacuum-to-water transfer free energy coefficient, can be determined empirically as discussed below. A_T is the total accessible area of the solute. Note that A_T corresponds to total area rather than just nonpolar area, since all polar groups are made to be neutral before the phase transfer step. In Eq. (4), ΔG_{np} incorporates van der Waals interactions between solute and solvent through their surface area dependence, while $\Delta G_{conf}(g)$ calculates van der Waals interactions between solute atoms explicitly. Thus the use of Eqs. (1) and (3) accounts for both

solute/solute and solute/solvent van der Waals interactions.

It has been known for many years that Eq. (4) works quite well for nonpolar area,[27–29] but more recently it has been assumed that polar group solvation is proportional to surface area as well. Eisenberg and McLachlan[1] derived atomic solvation parameters and wrote an expression in the form

$$\Delta G_{solv}=\Delta G_{ASP}(p1/p2)=\sum_i \Delta\sigma_i(p1/p2)A_i \qquad (5)$$

where the $\Delta s_i(p1/p2)$s are atomic solvation parameters (ASPs), obtained from a least squares fit to partition coefficients between two phases (p1 and p2), and A_i is the accessible surface area of atom i. In the original formulation[1] water/octanol partition coefficients were used, but more recently Wesson and Eisenberg obtained solvation parameters for gas/water partitioning, $\Delta\sigma_i(g/w)$, and these values can in principle be used in Eq. (1). However, despite their great utility, the use of ASPs is not strictly valid since electrostatic contributions cannot be strictly proportional to surface area. This must be true since in the ASP model, free energies no longer change once a group is buried, while actual electrostatic free energies are dependent on the shape of the interface and depth below the surface.[30] Moreover, ASPs are essentially "self-energy" terms which are dependent on the square of a charge on an atom, and thus do not account for solvent screening of pairwise interactions, which are proportional to the product of charges on different atoms (see, e.g., discussion in Gilson et al.[30]). The same criticisms are true for any solvation model which uses surface area alone to calculate electrostatic interactions. The magnitude of the error can only be determined empirically, but at least in a formal sense, the approach is incomplete.

It is worth outlining the relationship between the different approaches presented here. If the FDPB method is used to calculate ΔG_{es}, the total solvation free energy of any conformation is given by

$$\Delta G_{solv} = \gamma_{vw}A_T + \Delta G_{es}(FDPB). \qquad (6)$$

In contrast, the ASP model does not separate polar from nonpolar contributions, but rather incorporates them both into a single surface area term. If the two methods are to yield similar results, the relationship

$$\sum \Delta\sigma_i(g/w)A_i \approx \gamma_{vw}\Delta A_T + \Delta G_{es}(FDPB) \qquad (7)$$

would have to be satisfied. Note that for nonpolar carbon atoms there is little or no contribution to the electrostatic term on the right side of the equation, so that $\Delta\sigma_C(g/w)$ should be approximately equal to γ_{vw} and, consequently, be larger than zero since the transfer of nonpolar atoms from the gas phase to water involves an increase in free energy. However for polar atoms $\Delta\sigma_i(g/w)$ is negative whereas γ_{vw} is always positive and is the same for polar and nonpolar atoms. As discussed above, the right side of Eq. (8) is more physically realistic than the left since it accounts directly for solvent screening and does not assume that electrostatic free energy is proportional to surface area. On the other hand, the ASP model involves more limited computational demands, and it is easy to derive parameters from experimental transfer free energies. Results obtained from both methods are reported in this study.

Nonpolar Solvation Contributions

A number of recent papers have questioned traditional methods used to extract microscopic surface tension from solubility and partition experiments.[31–34] Oil/water transfer free energies have been known for some time to be proportional to surface area with a coefficient, γ_{ow}, equal to about 25 cal/mol/Å^2.[27–29] However theories of mixing[35–37] that account for differences between the molar volume of solute and solvent suggest that solute volume as well as surface area may contribute to hydrocarbon solubility. Application of these theories to linear alkanes indicates that the value for microscopic surface tension may be as large as about 50 cal/mol/Å^2.[31,32] This is close to the value of 72 cal/mol/Å^2 which characterizes the surface tension of alkane/water interfaces. The remaining discrepancy between the microscopic and macroscopic values has been attributed to differences in curvature between molecular and planar interfaces.[32,34] On this basis, a new model which defines a curvature-dependent microscopic surface tension was introduced.[34]

Recent experiments on protein stability using site-directed-mutagenesis have found that nonpolar groups contribute about 50 cal/mol/Å^2 [38] in close agreement with predictions based on Flory–Huggins theory. In contrast, it has been argued that this number contains contributions from internal cavities and that the hydrophobic surface tension may be closer to the standard value of 25 cal/mol/Å^2.[39,40] However, recent studies of mutations of surface residues, where cavities should not be present, indicate that the larger value may be more appropriate.[38,41] Unfortunately, this complex issue is not yet resolved.

The uncertainty in γ_{ow} is repeated when one attempts to derive values of γ_{vw} from solubility experiments. The standard value of 7 cal/mol/Å^2 [42] can be obtained from the slope of the line relating the transfer free energy of alkanes to their accessible surface area. However, if Flory–Huggins theory is applied to the alkane data, the slope is 0.72 kcal/mol/CH_2 or about 24 cal/mol/Å^2.[34] As was the case for γ_{ow}, the larger number is more consistent with macroscopic values (B. Honig, unpublished results). This work used the Flory–Huggins-based value corrected for curvature effects using the model of Nicholls et al.,[34] yielding a value for γ_{vw} of 30 cal/

TABLE I. Loops Modeled

Loop	Sequence
3A	Gly-Tyr-Thr-Arg-Thr-Thr-Asn-Asn-Arg 38 39 40 41 42 43 44 45 46
A4	Glu-Ala-Leu-Lys-Glu-His-Cys 57 58 59 60 61 62 63
5E	Lys-Gly-His-Ala-Gly-His-Pro-Glu 122 123 124 125 126 127 128 129

mol/Å^2 (D. Sitkoff, personal communication). However, as will be discussed later, the results of this study were quite insensitive to the value chosen for γ_{vw}.

FDPB Calculations

Calculations were carried out with the program DelPhi.[43] Atomic charges were taken from the X-PLOR[44] toph19x.pro unified atom (only polar hydrogens are explicitly represented) protein topology file. For the initial screen, a cubic grid with 65 grid points per side was used, with a grid spacing of 1.1 grid points/Å. The dielectric inside the protein (ε_{in}) was set to 2.0, while that outside (ε_{out}) was set to 80.0, to approximate that of water. Zero salt concentration and dipolar boundary conditions were used. The water probe radius was 1.4 Å. The atomic radii used were as follows: C 1.9 Å, C_α 1.86 Å, O 1.6 Å, N 1.65 Å, and H 1.0 Å. For those conformations which were in the lowest 40 by total energy, DelPhi was rerun on a cubic grid with 129 grid points per side, at a resolution of 2.2 grid points/Å, to give a more accurate estimate of the electrostatic solvation free energy.

Conformational Search

Calculations were carried out on three loops in *E. coli* RNase H[24] (the structure used in this work was essentially identical to entry 1rnh in the Protein Data Bank, but had been further refined from 2.0 to 1.7 Å) (Wei Yang and Wayne Hendrickson, unpublished results). The loops were loop 3A, connecting β-strand 3 and α-helix A; loop A4, connecting α-helix A and β-strand 4; and loop 5E, connecting β-strand 5 and α-helix E. The sequences and residue numbers are shown in Table I.

To generate loops for free energy evaluation, we used the TWEAK algorithm of Shenkin et al.,[10] incorporated under the name GENLOOP[45] into the torsion-based molecular modeling program PAKGGRAF.[46] TWEAK implements a random conformational search method, producing a large number of backbone conformations for a given loop. It functions by setting the loop backbone torsion angles to random values, and then using an iterative, linearized Lagrange multiplier technique to close the loop to the original end point constraints. All loop residues which are not proline or glycine are treated as alanines, i.e., side chain interactions are ignored. Conformations are rejected if any atoms in the loop are closer to one another or to an atom of the surrounding crystal framework than the sum of their van der Waals radii minus 1.1 Å; 500 (loops A4 and 5E) or 1000 (loop 3A) conformations were generated for each loop. PAKGGRAF was run on a Silicon Graphics Iris 220GTX.

An extensive set of rotamer-based side chain positioning algorithms[12,45] has been incorporated into PAKGGRAF. Two of these, CUSROT and NOVOSIDE, were used to position side chains on each backbone conformation generated. Starting conformations, along with the probability of occurrence of each rotamer, were taken from the rotamer library of Ponder and Richards.[47] The CUSROT algorithm produces a customized rotamer library by performing torsional minimization on each rotamer of selected sidechains. This minimization is done as though the other selected side chains are not present. NOVOSIDE rapidly finds reasonable side chain conformations in the context of the rest of the molecule, using the customized rotamer library. It was used with options that selected rotamers based on their probability of occurrence, and on contact distances. NOVOSIDE first initializes each selected side chain to its most probable rotamer. Then side chains are tested sequentially, and if a bad contact is found, further rotamers are tried in order of probability. The algorithm makes repeated passes through the selected side chains until no further replacements are made. The contact criterion used here was 2.2 Å between atom centers. Only side chains in the loop being treated were selected for placement by CUSROT and NOVOSIDE, but all atoms were used in scoring bad contacts.

Molecular Dynamics and Energy Minimization

The loop conformations produced by GENLOOP were quite strained. Following side chain placement, molecular dynamics followed by energy minimization[11] using the program X-PLOR was used to improve these conformations. All atoms not in the loop were held fixed in position. X-PLOR V3.0 was run on a Convex C220 computer. The unified atom parameter file param19x.pro, and topology file toph19x.pro supplied with X-PLOR V3.0 were used. The cut-on and cut-off for the nonbonded energy switching function were 10 and 11 Å, respectively.

The goal of the molecular mechanics calculations was to generate energetically reasonable structures to be used as input to a scheme which evaluates their conformational free energies. In principle, solvent effects should be considered during the generation of these structures, but this is not yet feasible if FDPB calculations are to be used to evaluate electrostatic free energies. In order to mimic the effects of solvent as closely as possible, a distant-dependent

dielectric constant was used. In addition, in most simulations the charges on all ionizable side chains were set to zero. Otherwise, the strong coulombic attraction between charged side chains and backbone hydrogen bonding groups resulted in spurious contacts being made, at times distorting the loop conformations (see Results). Thus, we have used a pseudo-gas-phase force field whose goal is to produce as reasonable solution conformations as possible rather than to produce accurate gas-phase conformational energies.

Twenty-five steps of Powell energy minimization were followed by 2 ps of molecular dynamics at 300 K. Even with no charges on the ionizable side chains, van der Waals forces were sufficient to cause long side chains to pack against the surface of the molecule during the molecular dynamics runs. These conformations produced unfavorable electrostatic solvation energies when later evaluated with DelPhi. For this reason, following molecular dynamics, side chains of ionizable residues were reextended through use of the CUSROT and NOVOSIDE algorithms, with the same probability based rotamer options as described above. In the absence of a proper solvation energy evaluation at this stage, this in effect supplies empirical solvation information, resulting in extended conformations for ionizable side chains when possible. Each loop conformation was then subjected to Powell energy minimization in X-PLOR, until the convergence criterion, the norm of the energy gradient, reached a value of 0.1 kcal/mol/Å. In general, the side chains remained extended through the minimization. Some conformations of each of the three loops were further minimized to convergence criteria of 0.01 and 0.001 kcal/mol/Å to test the importance of more complete convergence.

Loop Free Energies

Once a series of loop conformations has been generated, it is necessary to evaluate their relative conformational free energies in solution. In general it would be desirable to use Eq. (1) in conjunction with an accurate gas phase force field, but this is not generally available. In this paper we have tested three different approaches. The first is simply to use a pseudo-gas-phase potential function, as is still the practice in numerous applications. Here we use a distance-dependent dielectric constant which accounts in part for solvent screening of electrostatic interactions. Thus,

$$\Delta G^1_{\text{conf}}(\text{s}) = \Delta G_{\text{conf}}(\text{g},\varepsilon = r). \quad (8)$$

The second method adds the expression used in Eq. (6) for solvation free energies to a gas phase potential function to obtain $\Delta G^2_{\text{conf}}(\text{s})$. As discussed above, available force fields do not include electronic polarizability. In this work we account for electronic polarizability by using a dielectric constant of 2 for the solute in the FDPB calculations.[3] For consistency, we also use here a dielectric constant of 2 to reevaluate the gas phase conformational energy, since the Coulombic potential used in force fields is only valid if the internal dielectric constant of the solute is identical to the dielectric constant of the medium. Thus, we have reevaluated the electrostatic free energy in the gas phase as if it had a dielectric constant of 2. It should be recalled that the atomic charges themselves are not generally parameterized to be used with a true gas phase Coulombic potential. Until these are developed, the dielectric constant used to evaluate gas phase conformational energies will remain arbitrary.

$\Delta G^2_{\text{conf}}(\text{s})$ is defined by

$$\Delta G^2_{\text{conf}}(\text{s}) = \Delta G_{\text{conf}}(\text{g},\varepsilon = 2) + \gamma_{\text{vw}} A_{\text{T}} + \Delta G_{\text{es}}(\text{FDPB};2,80). \quad (9)$$

The electrostatic contribution to solvation in Eq. (9) is evaluated by transferring the solute from a medium of dielectric constant 2 to a medium of dielectric constant 80. In practice, DelPhi was used to calculate both the reaction field and pairwise coulombic energies. The coulombic calculation differed from those in X-PLOR in that DelPhi used all atom pairs, whereas X-PLOR excluded 1–2 and 1–3 pairs. Thus in evaluating $\Delta G^2_{\text{conf}}(\text{s})$, X-PLOR was used only for van der Waals interactions and for the sum over internal coordinates. The surface area term A_{T} was calculated using the program SURFCV,[48] which also calculated the curvature correction term.

The third method to evaluate conformation free energies uses the ASP formalism for solvation free energies. Thus,

$$\Delta G^3_{\text{conf}}(\text{s}) = \Delta G_{\text{conf}}(\text{g},\varepsilon = r) + \Delta G_{\text{ASP}}(\text{g/w}). \quad (10)$$

The rationale for this expression is that the screening of pairwise terms is accounted for in the dielectric constant while self-energy terms are accounted for with ASPs. It is also internally consistent in its treatment of van der Waals interactions since interactions between solute atoms are calculated with the force field, while van der Waals interactions between solute and solvent atoms are accounted for implicitly in the ASPs. [Note that it would not be valid to use water/octanol partition coefficients in Eq. (10) unless the solute force field was somehow modified to describe conformational free energies in octanol.]

RESULTS

Precision Tests of FDPB Calculations

Because DelPhi solves the Poisson–Boltzmann equation on a grid, the accuracy of the calculated electrostatic free energies depends on the resolution

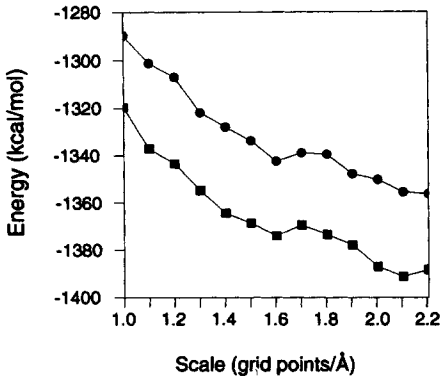

Fig. 1. DelPhi calculated electrostatic solvation energy vs grid scale for two loops. (●) Loop A4 conformation 551. (■) loop 5E conformation 12.

of the grid used. In all work reported in this paper we used a new and more accurate version of the DelPhi program.[49] This version also allowed a fixed coordinate to be placed at the center of the grid. Thus despite the different loop conformations, the bulk of the molecule remained stationary on the grid. To test the scale convergence of the reaction field energy, runs were made for several loop conformations, with the scale varying from 1.0 to 2.2 grid points/Å in steps of 0.1, on a cubic grid of 129 points per side.

Figure 1 shows the decrease in ΔG_{es} as the scale was increased from 1.0 grid points/Å to 2.2 grid points/Å; 2.2 grid points/Å was the maximum scale that could be achieved while keeping the entire molecule within the grid, and providing for a buffer between the grid edge and the molecule (the grid fill was 90%). The relatively smooth approach towards convergence at these low grid resolutions was an improvement over that obtained with earlier versions of DelPhi. Complete convergence was not obtained at 2.2 grid points/Å. For relative energy comparisons, however, this was not necessary, as the following results demonstrate.

A manifestation of the lack of scale convergence in ΔG_{es} is that the exact positioning of the molecule on the grid affects the result, due to the imperfect representation of the surface. The effect of the choice of the coordinates placed at the center of the grid was tested with grids of both the 65 and 129 points per side (1.1 and 2.2 grid points/Å). ΔG_{es} was calculated for the molecule for three conformations of loop A4, using the same coordinates for the center point each time. The molecules were then each displaced 0.1 grid units at a time, along the vector (1,1,1), within the range of ± 0.5 grid units (Fig. 2). It can be seen that for a spacing of 1.1 grid points/Å, the position on the grid has a significant impact on ΔG_{es}. The variation in ΔG_{es} at 2.2 grid points/Å is much smaller, but it is still significant. The variation in ΔG_{es} with grid position gives an estimate of the precision of the calculation for a particular conformation. However, if one compares values of ΔG_{es} only for different loop conformations where the center of the protein is located at the same position on the lattice, the variation is clearly quite small, particularly at 2.2 grid points/Å. This is because much of the precision problem arises from atoms in the protein that are not in the loop of interest. Thus, if the bulk of the protein is kept fixed on the grid, the errors introduced by the loops themselves are quite small.

Generating Loop Conformations

The major criterion used to evaluate the three free energy functions [Eqs. (8)–(10)] was the extent to which the crystal conformation was predicted to be lowest in free energy. However, the energies of the loops whose conformations are defined by the crystal structure were found to be quite high relative to the randomly generated loop conformations which were energy minimized as described above. For this reason, each of the three crystallographically determined loops was subjected to an energy minimization to a convergence criterion of 0.01 kcal/mol/Å, with the rest of the molecule held fixed. The resulting energies were used in comparisons between the crystal and generated loops. All root mean square (rms) deviation measurements, however, were against the original crystal structure. The rms deviation between each loop conformation and the crystal loop was calculated using the backbone atoms C, C_α, and N.

Initially, conformation optimization with molecular dynamics was performed with full charges on ionizable side chains. As shown in Figure 3a for loop A4, conformations so generated were much more stable than the minimized crystal loop structure when the energy was evaluated with $\Delta G^1_{conf}(s)$. When $\Delta G^2_{conf}(s)$ was used, the crystal structure was properly selected as the most stable (Fig. 3b). While it is encouraging that our solvation model is capable of identifying the correct conformation, the conformations generated in the gas phase with ionized side chains do not really provide a fair test of the methodology. This is because the charged side chains form hydrogen bonds with main chain atoms. In the gas phase there is a strong Coulombic force driving this process, while there is no compensating solvation term. In order to avoid generating physically unrealistic conformations, all further results reported here are for ionizable side chain charges set to zero during molecular dynamics and energy minimization.

In preliminary studies of loop generation, up to 5 psec of molecular dynamics were used for the optimization process. It was found that much of the energetic improvement and backbone atom movement

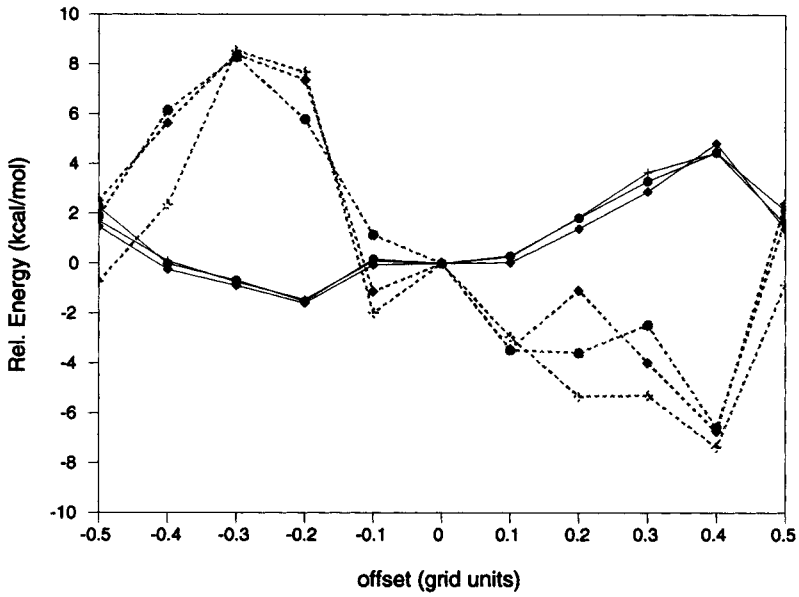

Fig. 2. Electrostatic solvation energies vs grid offset along displacement vector (1,1,1), calculated with DelPhi for three conformations of loop A4. Solid lines: scale 2.2 grid points/Å; broken lines: scale 1.1 grid points/Å. Conformation (●) 414, (◆) 479, (+) 551. Energies are relative to the value at zero offset. At zero offset, each molecule has the same coordinates at the center of the grid.

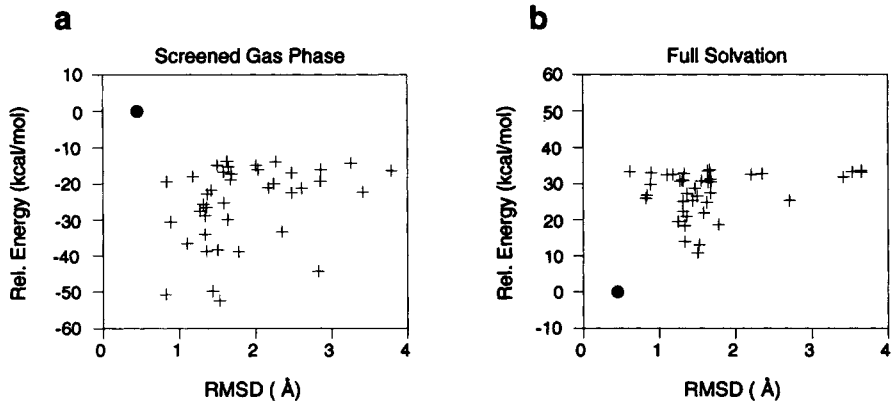

Fig. 3. Total energies, evaluated with two energy functions, for a set of loop A4 conformations (+) optimized with full side chain charges present during molecular dynamics and energy minimization, and no side chain extension steps, plotted against backbone rms deviation (RMSD) from the crystal structure. (**a**) Screened gas phase [$\Delta G^1_{conf}(s)$]; (**b**) full solvation model using DelPhi [$\Delta G^2_{conf}(s)$]. The energies are relative to the energy minimized crystal loop (●). Each graph shows the 40 lowest energy conformations by each energy evaluation method.

occurred during the first two picoseconds. Tests of increasing the convergence criterion of the energy minimization from 0.1 to 0.01 kcal/mol/Å caused some conformations to change slightly and drop in energy. However, there was a large cost in CPU time. Continuing to 0.001 kcal/mol/Å brought almost no further changes, but required another large increment of CPU time. 0.1 kcal/mol/Å was thus chosen as the best convergence criterion for energy minimization.

Analysis of Loop Conformations

Loop 3A

The boundaries of this loop were chosen to include one residue of β-sheet 3 and three residues of α-helix A as a test of the ability of GENLOOP to model such regions. This loop was highly constrained by surrounding residues. In initial runs of the search procedure, most conformations generated could not reach reasonable (negative) energies even after 5

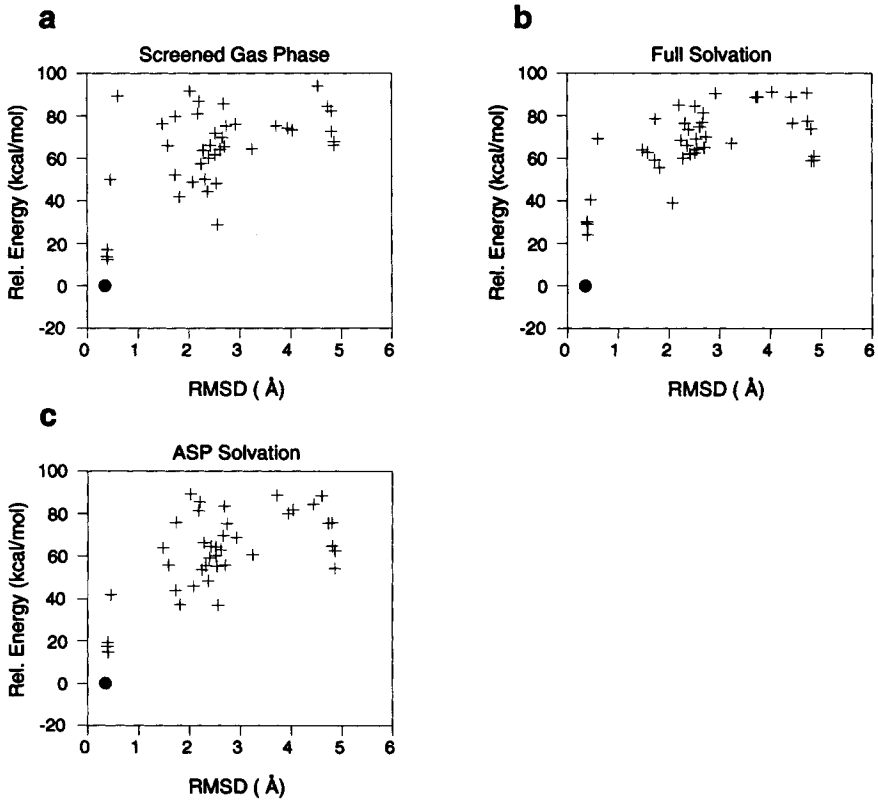

Fig. 4. Total energies for a set of loop 3A conformations (+), plotted against backbone rms deviation (RMSD) from the crystal structure. (**a**) Screened gas phase [$\Delta G^1_{conf}(s)$]; (**b**) full solvation model incorporating DelPhi [$\Delta G^2_{conf}(s)$]; (**c**) solvation portion calculated with atomic solvation parameters [$\Delta G^3_{conf}(s)$]. The energies are relative to the energy minimized crystal loop (•). Each graph shows the 40 lowest energy conformations by each energy evaluation method.

psec of molecular dynamics. For this reason, a preliminary energy minimization screen was performed on 1000 conformations. Conformations which could not reach a negative energy with energy minimization to a convergence criterion of 0.1 kcal/mol/Å were rejected. Of 1000 conformations 127 passed this screen.

Following molecular dynamics and energy minimization, the three low energy conformations evaluated using $\Delta G^1_{conf}(s)$ had backbone rms deviations from the crystal very similar to that of the minimized crystal loop (Fig. 4a). The lowest energy nonnative backbone conformations first appeared 29 kcal/mol above the minimized crystal loop energy, and with backbone rms deviation of greater than 2 Å. $\Delta G^2_{conf}(s)$ (Fig. 4b) gave a broad relative energy increase of all conformations, with the low energy, low rms deviation conformation at 24 kcal/mol, and with the first nonnative backbone conformation now 40 kcal/mol above the minimized crystal loop energy. However, in each case the three lowest energy conformations still had the near native backbone, with rms deviation of about 0.40 Å. Since the rms calculation is only for backbone atoms, it appears that the relatively high energy of the low rms structures is due to incorrect placement of the side chains.

$\Delta G^3_{conf}(s)$ also provides an improved discrimination of the crystal conformation relative to $\Delta G^1_{conf}(s)$ (Fig. 4c). It does not produce as large a desolvation penalty as is seen with $\Delta G^2_{conf}(s)$, resulting in a somewhat less effective identification of conformations with buried charges. On the other hand, the results are clearly quite satisfactory.

Loop A4

Molecular dynamics followed by energy minimization, when evaluated with $\Delta G^1_{conf}(s)$, yielded two conformations close to the crystal conformation in rms, but which had lower energies (Fig. 5a). $\Delta G^2_{conf}(s)$ resulted in a dramatic restriction of the conformations selected, all of which were within a backbone atom rms deviation of 1.7 Å from the crystal (Fig. 5b). The two lowest energy conformations both had energy almost identical to the minimized crystal loop. $\Delta G^3_{conf}(s)$ somewhat improved the correlation between energy and rms deviation when com-

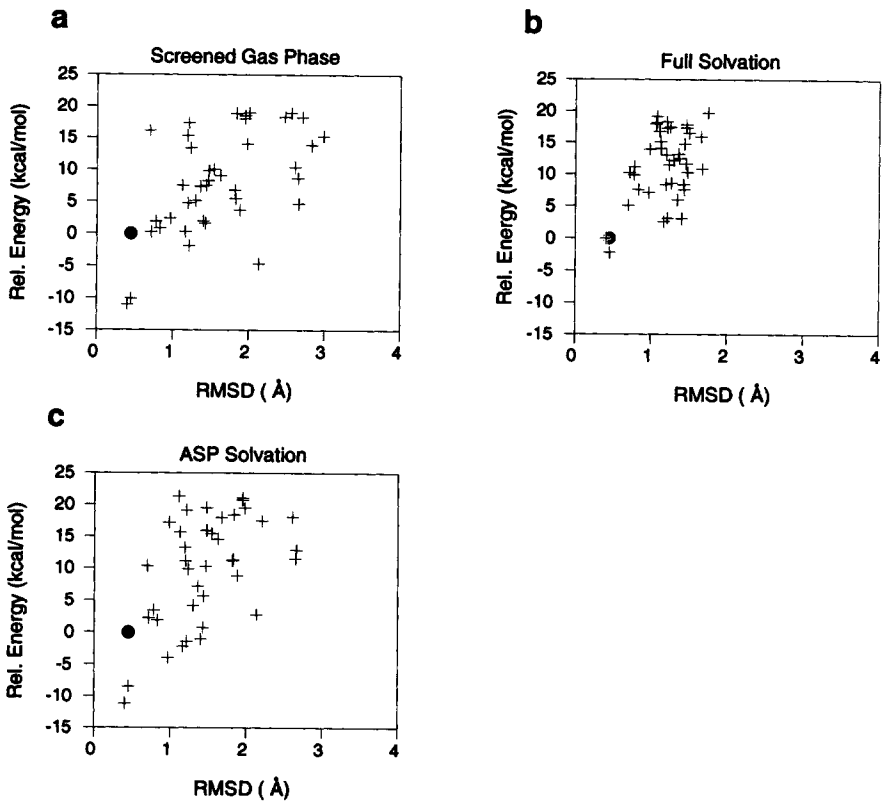

Fig. 5. Same as Figure 4, for loop A4.

pared to $\Delta G^1_{conf}(s)$ (Fig. 5c). However, for this loop $\Delta G^3_{conf}(s)$ was clearly much less effective than $\Delta G^2_{conf}(s)$ in identifying the crystal conformation.

Loop 5E

The minimized crystal loop energy was higher than the energy of about half of the 40 lowest energy conformations produced by molecular dynamics followed by energy minimization, when evaluated with $\Delta G^1_{conf}(s)$ (Fig. 6a). This is most likely due to the fact that this loop is in a crystal contact, and is thus somewhat distorted. $\Delta G^2_{conf}(s)$ radically changed the set of conformations comprising the lowest energy 40 (Fig. 6b). A conformation 3.14 Å rms deviation from the crystal became favored by 4 kcal/mol over the next two lowest energy conformations, one of which was also the conformation with lowest backbone rms deviation, and by 25 kcal/mol over the minimized crystal loop. $\Delta G^3_{conf}(s)$ improved the energy of the low rms deviation structure over $\Delta G^1_{conf}(s)$, however as for $\Delta G^2_{conf}(s)$, there was little correlation between energy and rms deviation from the crystal structure (Fig. 6c).

Comparison of Solvation Methods

The solvation components of $\Delta G^2_{conf}(s)$ and $\Delta G^3_{conf}(s)$ are compared directly in Figure 7 for the three loops. For loop 3A the agreement is relatively good, but shows considerable scatter. For loop A4, the agreement is fair, with a large amount of scatter. For loop 5E, there is little or no correlation. These results could be related to the fact that loop 3A only has two, relatively well-exposed charged side chains, while loop 5E has four charged side chains, several of which interact with other charged residues, thus forming a much more complex electrostatic environment.

The Nonpolar Solvation Term and Its Use

The range of $\Delta\Delta G_{np}$, for the low energy 40 conformations, relative to the minimized crystal loop for each of the three loops, was 8 kcal/mol, which was the smallest contribution of any of the energy terms. The calculations were repeated without the use of curvature correction,[34] using a value for γ_{vw} of 24 cal/mol/Å^2 instead of 30 cal/mol/Å^2. This reduced the range of $\Delta\Delta G_{np}$ to 6 kcal/mol for each of the three loops, and the changes in the energy plots of $\Delta G^2_{conf}(s)$ were almost imperceptible (results not shown). In a large scale screening of conformations, this term can probably be omitted in the initial stages. The curvature computation, which is currently quite time consuming, probably only needs to

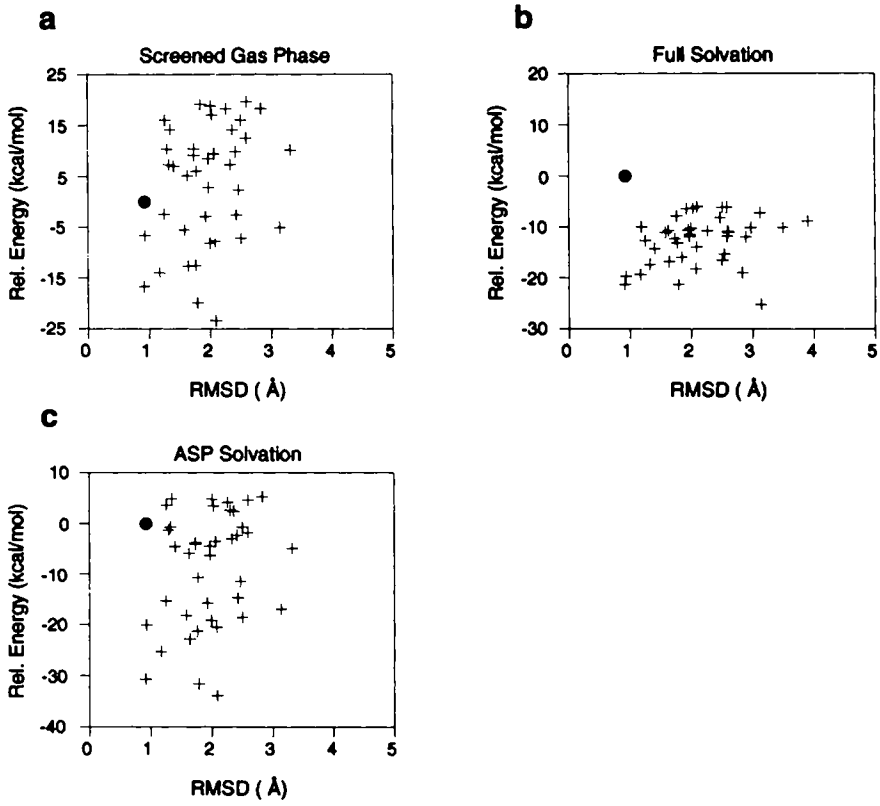

Fig. 6. Same as Figure 4, for loop 5E.

be done in a final comparison of low energy structures, if at all.

CPU Time Requirements

The evaluation of ΔG_{es} for a conformation using DelPhi at 1.1 grid points/Å required 29 sec on the Convex C220. The high resolution (2.2 grid points/Å) DelPhi runs required 304 sec per conformation, but due to the high quality of the relative solvation energy obtained at the lower resolution, only a small fraction of conformations need be so evaluated. A high resolution run of the program SURFCV to evaluate surface area and curvature required 55 sec per conformation on the Iris 220GTX.

The loop conformation generation process used here proved rather slow. While GENLOOP required only 3 sec per accepted conformation, and sidechain placement took 12 sec (both on the Iris), 2 psec of molecular dynamics and energy minimization required 376 sec per conformation on the Convex C220 for loop A4, the shortest modeled.

DISCUSSION

In this paper we describe a method for the inclusion of electrostatic and nonpolar solvation energies into a loop conformational search procedure. Loop backbones were first generated randomly and side chains were added using conformations taken from a rotamer library. Pseudo-gas-phase molecular dynamics calculations were then carried out on these structures. In the last stage of loop generation, side chains were placed in extended conformations again, with the help of rotamer libraries. This procedure was successful in generating conformations in which the side chains were well solvated. When gas-phase molecular dynamics alone was used to generate loop conformations, side chains packed against the backbone, even if they were charged, leading to high energy conformations in which the side chains are not properly solvated. Such conformations are easily detectable if a reasonable electrostatic model is used to calculate solvation free energies. While this may be viewed as a success of solvation free energy calculations, it does not test their true utility since it is clear, a priori, that most of the conformations being evaluated are incorrect.

There have been several previous efforts to include the effect of solvent in conformational search techniques. Moult and James[13] used an image charge method for including the effects of solvent screening, and also used exposed hydrophobic area as a final filter in selecting conformations. For the

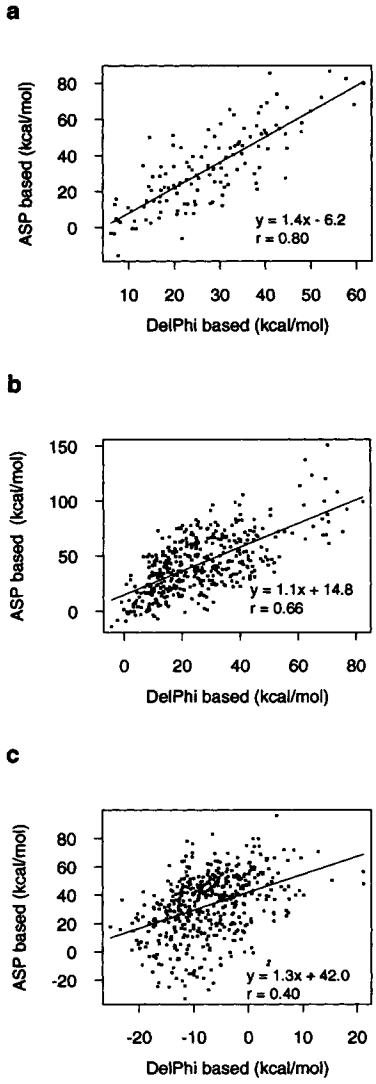

Fig. 7. Comparison of the DelPhi vs ASP based solvation energy functions, with coulombic terms also included. As used here, the ASP-based method equals $\Delta G_{ASP} + \Delta G_{coul}(\varepsilon = r)$. The DelPhi-based method equals $\Delta G_{np} + \Delta G_{coul}(\varepsilon = 2) + \Delta G_{es}(\varepsilon_{in} = 2, \varepsilon_{out} = 80)$. The best fit line is shown along with its defining equation and correlation coefficient (r). (**a**) Loop 3A; (**b**) loop A4; (**c**) loop 5E.

two loops they constructed, these terms were sufficient to choose low rms deviation conformations from among those which had passed prior filters. Martin et al.,[16] using the conformational search program CONGEN in combination with database search techniques, reported that CONGEN's in vacuo potential rarely selected low rms conformations by low energy. Instead, it picked conformations which optimized the van der Waals interactions, and hence tended to be too compact. Following work of several others, they dropped the attractive term of the van der Waals potential for solvent exposed atoms. This was to compensate for the missing dispersion interactions between water and protein. Note that this is the opposite of the approach we have adopted, where solute/solvent van der Waals interactions are included implicitly in the surface area term [Eq. (6)]. Martin accounted for solvent screening by simply dropping the electrostatic energy completely from the final energy evaluation.

A number of recent papers report the use of ASP's to evaluate solvation free energies. Mas et al.[45] used a joint screening procedure based on X-PLOR and ASP's to select among conformations for H3 on an antibody against carcinogenic embryonic antigen. Schiffer et al.[50] developed a homology modeling method which energy minimized various conformers of each targeted residue along with its nearest neighbors, using the program AMBER with a distance-dependent dielectric constant. ASPs were used to estimate solvation energies for each targeted residue. It was found that for hydrophilic surface residues, the minimization process often resulted in a less favorable solvation energy than that of the equivalent residue in the crystal. As we found in this work, this was due to the loss of well solvated extended states by surface residues, which instead tended to form hydrogen bonds to other protein atoms.

While this work is similar in approach to earlier studies, it is based on a more complete solvation model through its use of the FDPB method. As such, it provides a means of testing simpler but less time-consuming computational models. One goal of this work was to determine the extent to which the inherent precision problems in the FDPB method could be overcome in the evaluation of relative conformational free energies. The results reported above provide a prescription for obtaining reliable results even given the uncertainties associated with evaluating energies on a lattice.

Overall, the results clearly demonstrate that a good solvation model is capable of discriminating stable from unstable conformations. Gas phase conformations generated from molecular dynamics calculations alone are clearly unstable when conformational free energies in solution are evaluated with $\Delta G^2_{conf}(s)$, particularly if ionizable side chains are charged in the simulations. This can be seen quite clearly in Figure 3. When more reasonable structures are generated, $\Delta G^2_{conf}(s)$ is quite successful in selecting conformations of low rms deviation from the crystal structure for loop 3A and 4A, but fails for loop 5E. However, loop 5E forms an intermolecular contact in the crystal, and consequently does not constitute a fair test of the methodology. Figure 8 compares the lowest energy conformation for each loop to that found in the crystal. The backbone conformation for loops 3A (Fig. 8a) and A4 (Figure 8b) are clearly quite close to the crystal conformations, but this is clearly not the case for loop 5E (Fig. 8c). On the other hand, the second lowest energy confor-

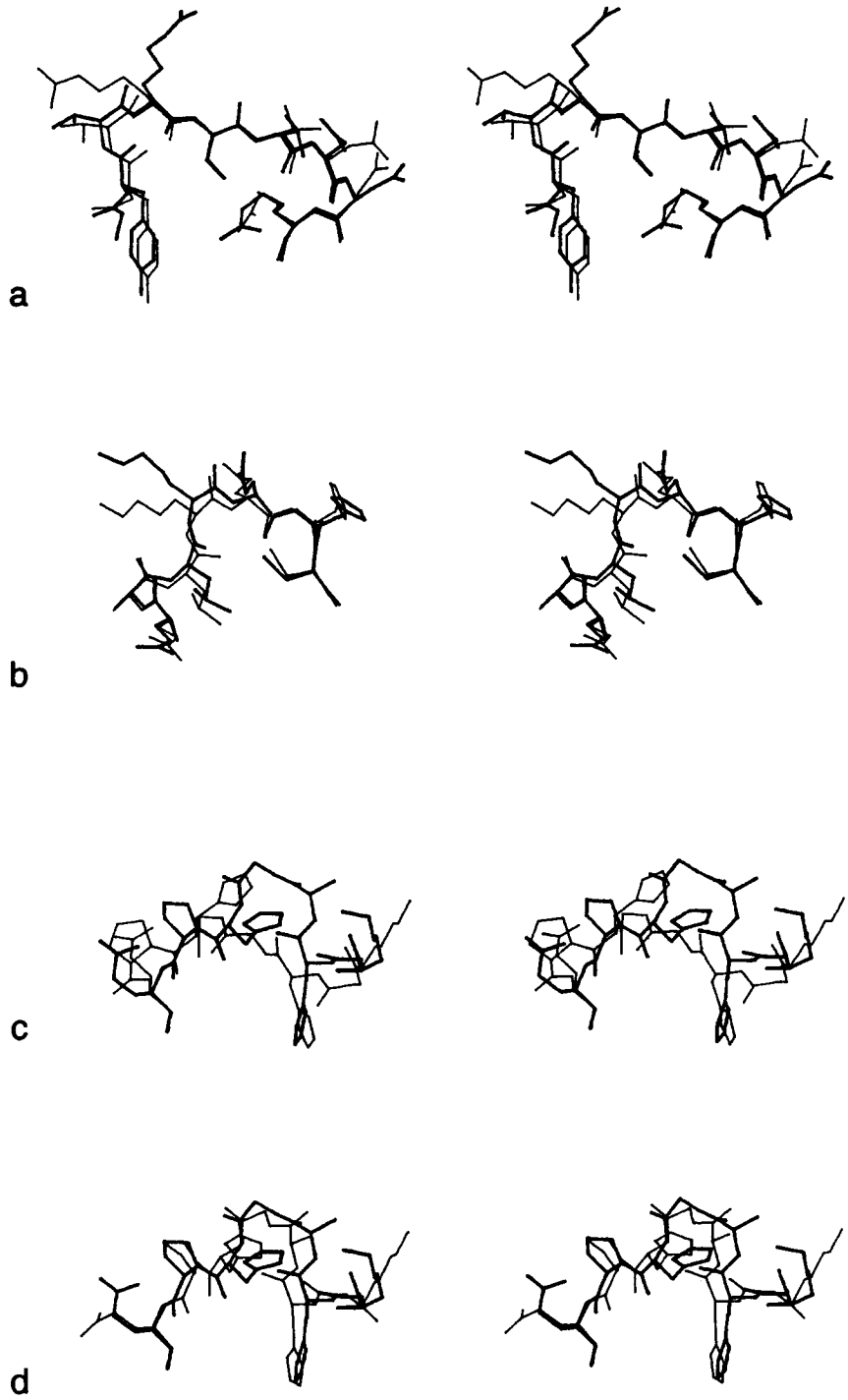

Fig. 8. The low energy loop conformations by $\Delta G^2_{conf}(s)$ (light lines) compared to the crystal structure (heavy lines). (**a**) Loop 3A, conformation 731; (**b**) loop A4, conformation 551; (**c, d**) for loop 5E, the lowest energy structure, conformation 388, has a high rms deviation (**c**), but the next lowest energy structure, conformation 12, has the lowest rms deviation (**d**) (see Fig. 6b).

mation for loop 5E has a backbone conformation quite close to that observed in the crystal (Fig. 8d). $\Delta G^3_{conf}(s)$, which utilizes ASP-based solvation energies, also succeeds in identifying the crystal structure as a low energy conformation in two of the three loops. The low correlation found between the solva-

tion components of $\Delta G^2_{conf}(s)$ and $\Delta G^3_{conf}(s)$ (Fig. 7) is somewhat surprising, and may result in part from the fact that the atomic charges and radii used in DelPhi have not been parameterized to reproduce the solvation energies used to derive the ASPs. This work used the X-PLOR charge set for the continuum calculation so as to maintain some consistency between X-PLOR and DelPhi calculated electrostatic energies. In future work it would seem more reasonable to treat solvation as a separate problem from the calculation of gas-phase conformational energies and to allow the use of different parameters in each set of calculations. A set of parameters specifically designed for continuum solvation free energy calculations is currently under development.[51]

The method used in this work to generate loop conformations does not constitute an optimal approach to the problem. For example, the preference given for extended side chain conformations is not justified in the general case, although in this work it did help avoid conformations in which polar groups are packed against the protein surface. Such conformations are easily detected as unstable once solvation is taken into account, which may be viewed as a success of the two solvation models that have been used. The problem with the strategy used in this and much similar work is that a great deal of effort is wasted generating conformations that one knows will have unrealistic features since they are based on gas-phase potential functions. Thus it is essential to develop search methods in which solvation free energies are accounted for during the stage of generating trial structures.

Efforts in this direction have already begun. For example, several groups have begun work on incorporating solvation energy into molecular mechanics programs. Williams et al.[9] used empirical parameterization methods to estimate solvation free energy from solvent accessible surface area. This energy term was added to the conformational energy calculated by the program ECEPP/2, and used in energy minimization. Wesson and Eisenberg[7] describe the addition of a solvation energy function, based on ASPs, to CHARMM. Schiffer et al.[8] incorporated a similar ASP-based function into the molecular mechanics program AMBER. Another approach is that of Gilson and Honig,[5] who simulated charge-solvent interactions with a $1/r^4$-based term for pairwise atomic interactions. More detailed electrostatic models that are not based on surface area alone have reported by Sharp,[4] Zauher,[6] and Gilson et al.[5]

The work reported here provides a basis for evaluating such approaches because the solvation model used in $\Delta G^2_{conf}(s)$ constitutes a complete macroscopic treatment. It is too time consuming to be used in a conformational search, but it is fast enough (29 secs per conformation) to allow the evaluation of a large number of conformations. Thus, it might be fruitful to consider approaches in which simpler solvation models were used to generate trial conformations which would then be evaluated with a function similar to $\Delta G^2_{conf}(s)$. The fact that it has been possible to overcome the precision problems inherent in the FDPB method suggests that such calculations can be useful in the evaluation of conformational free energies.

ACKNOWLEDGMENTS

We thank Wei Yang and Wayne Hendrickson for supplying the crystal structure of RNase H refined to 1.7 Å, and for numerous helpful discussions throughout this project. We also thank Richard Fine for many helpful discussions. We thank Doree Sitkoff for her calculation of the alkane vacuum-water transfer free energy, David Yarmush for assistance with PAKGGRAF, and Anthony Nicholls and Kim Sharp for their help and advice. Financial support from NIH (GM 30518) and the Office of Naval Research (N00014-90-J-1713) is gratefully acknowledged.

REFERENCES

1. Eisenberg, D., McLachlan, A.D. Solvation energy in protein folding and binding. Nature (London) 319:199–203, 1986.
2. Jean-Charles, J., Nicholls, A., Sharp, K., Honig, B., Tempczyk, A., Hendrickson, T., Still, C. Electrostatic contributions to solvation energies: Comparison of free energy perturbation and continuum calculations. J. Am. Chem. Soc. 113:1454–1455, 1990.
3. Sharp, K.A., Honig, B. Electrostatic interactions in macromolecules: theory and applications. Annu. Rev. Biophys. Biophys. Chem. 19:301–332, 1990.
4. Sharp, K.A. Incorporating solvent and ion screening into molecular dynamics using the finite-difference Poisson-Boltzmann method. J. Comp. Chem. 12:454–468, 1991.
5. Gilson, M., Honig, B. The inclusion of electrostatic hydration energies in molecular mechanics calculations. J. Comp. Aided Mol. Design 5:5–20, 1991.
6. Zauher, R.J. The incorporation of hydration forces determined by continuum electrostatics into molecular mechanics simulations. J. Comp. Chem. 12:575–583, 1991.
7. Wesson, L., Eisenberg, D. Atomic solvation parameters applied to molecular dynamics of proteins in solution. Prot. Sci. 1:227–235, 1992.
8. Schiffer, C.A., Caldwell, J., Stroud, R., Kollman, P. Inclusion of solvation free energy with molecular mechanics energy: Alanine dipeptide as a test case. Prot. Sci. 1:396–400, 1992.
9. Williams, R.L., Vila, J., Perrot, G., Scheraga, H.A. Empirical solvation models in the context of conformational energy searches: Application to bovine pancreatic trypsin inhibitor. Proteins 14:110–119, 1992.
10. Shenkin, P.S., Yarmush, D.L., Fine, R.M., Wang, H., Levinthal, C. Predicting antibody hypervariable loop conformation. I. Ensembles of random conformations for ring-like structures. Biopolymers 26:2053–2085, 1987.
11. Fine, R.M., Wang, H., Shenkin, P.S., Yarmush, D.L., Levinthal, C. Predicting antibody hypervariable loop conformations II. Minimization and molecular dynamics studies of MCPC603 from many randomly generated loop conformatioons. Proteins 1:342–362, 1986.
12. Bajörath, J., Fine, R.M. On the use of minimization from many randomly generated loop structures in modeling antibody combining sites. Immunomethods 1:137–146, 1992.
13. Moult, J., James, M.N.G. An algorithm for determining the conformation of polypeptide fragments in proteins by systematic search. Proteins 1:146–168, 1986.
14. Bruccoleri, R.E., Karplus, M. Prediction of the folding of short polypeptide segments by uniform conformational sampling. Biopolymers 26:137–168, 1987.
15. Bruccoleri, R.E., Haber, E., Novotny, J. Structure of anti-

body hypervariable loops reproduced by a conformational search algorithm. Nature (London) 335:564–568, 1988.
16. Martin, A.C.R., Cheetham, J.C., Rees, A.R. Modeling antibody hypervariable loops: A combined algorithm. Proc. Natl. Acad. Sci. U.S.A. 86:9268–9272, 1989.
17. Dudek, M.J., Scheraga, H.A. Protein structure prediction using a combination of sequence homology and global energy minimization I. Global energy minimization of surface loops. J. Comp. Chem. 11:121–151, 1990.
18. Palmer, K.A., Scheraga, J.A. Standard-geometry chains fitted to X-ray derived structures: Validation of the rigid-geometry approximation. II. Systematic searces for short loops in proteins: Applications to bovine pancreatic ribonuclease A and human lysozyme. J. Comp. Chem. 13:329–350, 1992.
19. Blinn, J.R., Chou, K., Howe, W.J., Maggiora, G.M., Mao, B., Moon, J.B. Computer modelling of constrained peptide systems. NATO advanced research workshop: "The role of computational models and theories in biotechnology," 1991.
20. Jones, T.A., Thirup, S. Using known substructures in protein model building and crystallography. EMBO J. 5:819–822, 1986.
21. Chothia, C., Lesk, A.M. Canonical structures for the hypervariable regions of immunoglobulins. JMB 196:901–917, 1987.
22. Chothia, C., Lesk, A.M., Tramontano, A., Levitt, M., Smith-Gill, S.J., Air, G., Sheriff, S., Padlan, A.A., Davies, D., Tulip, W.R., Colman, P.M., Spinelli, S., Alzari, P.M., Poljak, R.J. Conformations of immunoglobulin hypervariable regions. Nature (London) 342:877–883, 1989.
23. Novotny, J., Bruccoleri, R.E., Karplus, M. An analysis of incorrectly folded protein models. Implications for structure predictions. J Mol. Biol. 177:787, 1984.
24. Yang, W., Hendrickson, W.A., Crouch, R.J., Satow, Y. Structure of ribonuclease H phased at 2Å resolution by MAD analysis of the selenomethionyl protein. Science 249: 1398–1990, 1990.
25. Gilson, M.K., Honig, B. Calculation of the total electrostatic energy of a macromolecular system: Solvation energies, binding energies, and conformational analysis. Proteins 4:7–18, 1988.
26. Sharp, K., Honig, B. Electrostatic interactions in macromolecules: Theory and applications. Annu. Rev. Biophys. Biophys. Chem. 19:301–332, 1990.
27. Hermann, R.B. Use of solvent cavity area in regard to hydrocarbon solubilities and hydrophobic interactions. Proc. Natl. Acad. Sci. U.S.A. 74:4144, 1977.
28. Chothia, C. The nature of the accessible and buried surface in proteins. J. Mol. Biol. 105:1–14, 1976.
29. Tanford, C.H. "The Hydrophobic Effect." John Wiley & Sons. New York: 1980.
30. Gilson, M.K., Rashin, A., Fine, R., Honig, B. On the calculation of electrostatic interactions in proteins. J. Mol. Biol. 183:503–516, 1985.
31. DeYoung, L.R., Dill, K.A. Partitioning of nonpolar solutes into bilayers and amorphous n-alkanes. J. Phys. Chem. 94:801–809, 1990.
32. Sharp, K.A., Nicholls, A., Fine, R.M., Honig, B. Reconciling the magnitude of the microscopic and macroscopic hydrophobic effects. Science 252:106–109, 1991.
33. Sharp, K.A., Nicholls, A., Friedman, R., Honig, B. Extracting hydrophobic free energies from experimental data: Relationship to protein folding and theoretical models. Biochemistry 30:9686–9697, 1991.
34. Nicholls, A., Sharp, K.A., Honig, H. Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins 11:271–280, 1991.
35. Flory, P.J. Thermodynamics of high polymer solutions. J. Chem. Phys. 9:660–671, 1941.
36. Huggins, M.L. Solutions of long chain polymers. J. Chem. Phys. 9:440–449, 1941.
37. Hildebrand, J.H. The entropy of solution of molecules of different size. J. Chem. Phys. 15:225–228, 1947.
38. Serrano, L., Sancho, J., Hirshberg, M., Fersht, A.R. Alpha-helix stability in proteins I. Empirical correlations concerning substitution of side-chains at the N and C-caps and the replacement of alanine by glycine or serine at solvent-exposed surfaces. J. Mol. Biol. 227:544–559, 1992.
39. Eriksson, A.E., Baase, W.A., Zhang, X.-J., Heinz, D.W., Blaber, M., Baldwin, E.P., Matthews, B.W. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. Science 255:178–183, 1992.
40. Lee, B. Estimation of the maximum change in stability of globular proteins upon mutation of a hydrophobic residue to another of smaller size. Prot. Sci. 2:733–738, 1993.
41. Serrano, L., Neira, J., Sancho, J., Fersht, A. Effect of alanine versus glycine in α-helices on protein stability. Nature (London) 356:453–455, 1992.
42. Still, W.C., Tempczyk, A., Hawley, R.C., Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. J. Am. Chem. Soc. 112:6127–6129, 1990.
43. Nicholls, A., Honig, B. A rapid finite difference algorithm utilizing successive over-relaxation to solve the Poisson-Boltzmann equation. J. Comp. Chem. 12:435–445, 1991.
44. Brünger, A.T. X-PLOR Manual, Version 3.0. New Haven: Yale University, 1992.
45. Mas, M.T., Smith, K.C., Yarmush, D.L., Aisaka, K., Fine, R.M. Modeling the anti-CEA antibody combining site by homology and conformational earch. Proteins 14:483–498, 1992.
46. Katz, L., Levinthal, C. Interactive computer graphics and representation of complex biological structures. Annu. Rev. Biophys. Bioeng. 1:465–504, 1972.
47. Ponder, J.W., Richards, F.M. Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes. J. Mol. Biol. 193:775–791, 1987.
48. Sridharan, S., Nicholls, A., Honig, B. A new vertex algorithm to calculate solvent accessible surface areas. Biophys. J. 61:A174, 1992.
49. Nicholls, A. In preparation.
50. Schiffer, C.A., Caldwell, J.W., Kollman, P.A., Stroud, R.M. Prediction of homologous protein structures based on conformational searches and energetics. Proteins 8:30–43, 1990.
51. Sitkoff, D., Sharp, K.A., Honig, B. Accurate calculation of hydration free energies using macroscopic solvent models. J. Phys. Chem. (in press).