ELSEVIER

Biophysical Chemistry 78 (1999) 195–205

Biophysical Chemistry

ES/IS: Estimation of conformational free energy by combining dynamics simulations with explicit solvent with an implicit solvent continuum model

Yury N. Vorobjev*,[1] Jan Hermans

Department of Biochemistry and Biophysics, School of Medicine, University of North Carolina, Chapel Hill, NC 27599-7260, USA

Received 15 October 1998; received in revised form 23 November 1998; accepted 23 November 1998

Abstract

This paper reviews a recently developed method for calculating the total conformational free energy of a solute macromolecule in water solvent. The method consists of a relatively short simulation by molecular dynamics with *explicit solvent* molecules (ES) to produce a set of microstates of the macroscopic conformation. Conformational internal solute energy and entropy are obtained from the simulation, the latter in the quasi-harmonic approximation by analysis of the covariance matrix. The *implicit solvent* (IS) surface energy–dielectric continuum model is used to calculate the average solvation free energy as the sum of the free energies of creating the solute-size hydrophobic cavity, of the van der Waals solute–solvent interactions and of the polarization of water solvent by the solute's charges. We have earlier applied this method to calculate the conformational free energy of native and intentionally misfolded globular conformations of proteins (the EMBL set of deliberately misfolded proteins), and have obtained good discrimination in favor of the native conformations in all instances. These results are summarized and further analyzed to show that, on average, three major component terms of the free energy all contribute in favor of discrimination. We discuss possible improvements of the ES/IS method. It is shown how the force field can be made self-consistent by adapting the parameters for calculation of surface and polarization free energies closely to the molecular mechanics force field used in the dynamics simulation, using established simulation methods to compute free energies for cavity formation and a charging process with the molecular mechanics force field to provide a set of (quasi-experimental) reference data that can be used to refine the parameters of the continuum models. The molecular surface area together with a microscopic surface free energy near 70 cal/(mol Å^2) is found to be a consistent descriptor of the cavity free energy. Preliminary results indicate that a linear-response approximation for the polarization of water solvent reaction near typical polar and charged protein groups is accurate to within approximately 90%.

Keywords: Misfolded proteins; Conformational free energy; Implicit solvation; Continuum dielectric; Molecular dynamics

* Corresponding author.
[1] On leave from the Novosibirsk Institute of Bioorganic Chemistry, Novosibirsk, 630090, Russia.

0301-4622/99/$ - see front matter
PII: S0301-4622(98)00230-0

1. Introduction

The goal of distinguishing between different conformations of macromolecules in terms of a free energy based on a strictly physical potential has been elusive. While free energy perturbation methods, based on microscopic simulation of a macromolecule with explicit solvent, may in principle be suitable, this in practice meets with tremendous difficulties due to the large molecular size, the need to sample adequately over solvent and solute conformations and properly evaluate long-range electrostatic interactions [1,2]. This paper reviews an approach that combines *explicit solvent* (ES) and *implicit solvent* (IS) models to the calculation of the conformational free energy of a macromolecule in aqueous solution that has been recently presented elsewhere [3]. The ES/IS method uses a microscopic simulation with explicit solvent in order to provide a representative sample of microstates of a given macroscopic conformation. The free energy is evaluated with an accurate expression for the free energy of a macromolecular solute in a solvent, in terms of the solute's average solvent-mediated energy and a conformational entropy that accounts for the solute's internal degrees of freedom. We have shown that by this approach we were able to distinguish incorrectly folded conformations of a series of proteins from the corresponding correct folds [3,4]. In this paper, we analyze in more detail the components, which are evaluated as part of the total conformational free energy in the ES/IS method.

In the following section we review the formulation of ES/IS. Following this, we review the results on the incorrectly folded proteins, and present some results regarding the models used to evaluate the hydrophobic solvation free energy and of the implicit solvent continuum dielectric model.

2. Formulation of the ES / IS method for calculating conformational free energy

An accurate calculation of the free energy of a macromolecule in an aqueous solution requires evaluation of the partition function and thus the volume of accessible phase space [1,2,5,6]. While it is possible to estimate the free energy of a structurally highly organized system from the results of a simulation, the same is not true for the solvent. To avoid the difficult problem of properly sampling solvent configuration, we have adopted an implicit description of solvation, and thereby obtained a partition function of the solute in which the interactions with the solvent are represented through a solvation potential (a potential of mean force) that depends explicitly on the solute's coordinates.

The partition function, Z of a solute molecule (coordinates $\mathbf{x}$) in a solvent (coordinates $\mathbf{y}$) can be written as the ratio of the partition functions for solution and pure solvent systems (containing identical numbers of solvent molecules) [7], i.e.

$$Z = \frac{\int d\mathbf{x} \int d\mathbf{y} \exp\{-\beta[U_m(\mathbf{x}) + U_{m,s}(\mathbf{x},\mathbf{y}) + U_{s,s}(\mathbf{y})]\}}{\int d\mathbf{y} \exp[-\beta U_{s,s}(\mathbf{y})]} \tag{1}$$

Here $U_m(\mathbf{x})$ is the intra-molecular potential energy, $U_{m,s}(\mathbf{x},\mathbf{y})$ is the potential energy of the solute–solvent interactions and $U_{s,s}(\mathbf{y})$ is the potential energy of the solvent–solvent interactions. This can be rewritten as a partition function with solvent-mediated interactions between atoms of the solute molecule [8,9]

$$Z = \int d\mathbf{x} \exp\{-\beta[U_m(\mathbf{x}) + \Delta W(\mathbf{x})]\} \tag{2}$$

where the sum $[U_m(\mathbf{x}) + \Delta W(\mathbf{x})]$ presents an effective configurational energy, and

$$\exp[-\beta \Delta W(x)] = \frac{\int d\mathbf{y} \exp\{-\beta[U_{m,s}(\mathbf{x},\mathbf{y}) + U_{s,s}(\mathbf{y})]\}}{\int d\mathbf{y} \exp[-\beta U_{s,s}(\mathbf{y})]} \tag{3}$$

Here $\Delta W(\mathbf{x})$ is the free energy of solvation of the solute molecule with conformation $\mathbf{x}$. Considering scaled molecule–solvent interaction with

coupling parameter λ, the solvation free energy $\Delta W(\mathbf{x})$ can be written in the framework of the free energy perturbation method

$$\Delta W(\mathbf{x}) = \int_0^1 \mathrm{d}\lambda \frac{\int U_{m,s}(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{y} \exp\{-\beta[\lambda U_{m,s}(\mathbf{x},\mathbf{y}) + U_{s,s}(\mathbf{y})]\}}{\int \mathrm{d}\mathbf{y} \exp\{-\beta[\lambda U_{m,s}(\mathbf{x},\mathbf{y}) + U_{s,s}(y)]\}} \quad (4)$$

and this provides an expression suitable for a microscopic simulation. Considering a three-step sequential 'turning on' of different types of solute–solvent interactions in Eq. (4), one sees that the solvation free energy $\Delta W(\mathbf{x})$ can be decomposed into parts [10–13], namely, G_{cav}, the free energy of creating the empty solute-sized cavity in the solvent, $G_{s,vdw}$, the free energy of inserting the solute molecule into the solute-sized cavity, i.e. turning on the van der Waals interactions between the solute atoms and the solvent molecules, and G_{pol}, the free energy of solvent polarization, i.e. the excess free energy of charging the solute atoms in the solvent from charges set to zero to their regular values, the excess being relative to the same process in vacuo, i.e.

$$\Delta W = G_{cav} + G_{s,vdw} + G_{pol} \quad (5)$$

where the terms in Eq. (5) describe elementary physical processes that can be treated individually.

Finally, presenting all phase space of a solute molecule as a sum of sub-spaces A,B,.., each of which describes a distinct macroscopic solute conformation, it follows from Eq. (2) that the free energy A_A of a solute molecule in a macroscopic conformation A can generally be presented in terms of average configurational energy and entropy over the molecular degrees of freedom $\mathbf{x}$

$$G_A \approx A_A = \langle U_m(\mathbf{x})\rangle_A + \langle \Delta W(\mathbf{x})\rangle_A - TS_{conf,A} \quad (6)$$

where $\langle\rangle_A$ denotes an average over micro-configurations of the conformation A, U_m represents the intra-protein conformational energy and S_A is the entropy of the conformation A.

The solvation free energy $\Delta W(\mathbf{x})$ is written as a sum of terms for cavity formation, solute–water van der Waals interactions and electrostatic polarization of solvent by the polar components of the solute. As a result, Eq. (6) becomes

$$G_A \approx A_A = \langle U_{m,sh}\rangle_A + \langle U_{m,coul}\rangle_A - TS_{conf,A} + \langle G_{cav}\rangle_A + \langle G_{s,vdw}\rangle_A + \langle G_{pol}\rangle_A \quad (7)$$

where the intramolecular potential energy U_m has been represented as a sum of short-range (i.e. angle deformation and van der Waals) terms, $U_{m,sh}$ and electrostatic Coulombic interactions, $U_{m,coul}$.

In the ES/IS method, a representative set of microscopic configurations $\mathbf{x}_{A,i}$ of a solute in a solvent is generated by MD simulation with explicit solvent along a relatively short trajectory (50–100 ps). Of the six terms in Eq. (7), three, namely $\langle U_{m,sh}\rangle$, $\langle U_{m,coul}\rangle$ and $\langle G_{s,vdw}\rangle$ are accumulated as averages during the molecular dynamics simulation. The free energy of van der Waals interactions between solute and solvent, $G_{s,vdw}$ can be accurately approximated by the potential energy of these interactions [14], $U_{s,vdw}$ which can be calculated easily during a molecular dynamics simulation.

$$G_{s,vdw} = U_{s,vdw} \quad (8)$$

(Perturbation theory of liquid structure provides a theoretical justification that liquid structure undergoes negligible changes during the process of 'switching on' the van der Waals forces between solute and solvent [15].)

The entropic term, $TS_{conf,A}$ is estimated in the harmonic approximation from the covariance matrix of the positional fluctuations during the dynamics trajectory [3,16–20]. This term pertains only to the conformational fluctuations of the solute molecule, and is not to be confused with the total entropy.

Finally, the remaining two terms, the free energy for formation of the cavity G_{cav} and the free

energy of solvent polarization, G_{pol}, which are difficult to estimate by microscopic simulations, are found with models in which the solvent is treated *implicitly* with an appropriate physical model, as a continuum. As will be discussed, the free energy for formation of the cavity is well approximated with a term given by the product of the molecular surface and a microscopic surface tension and the free energy of solvent polarization, G_{pol} is found by modeling the solvent as a continuum dielectric, with Poisson's equation. These two terms are calculated for a set of evenly spaced conformations from the trajectory. A molecular surface is computed with the SIMS program [21], which is smooth, invariant to rotation and translation and tessellated more uniformly than the molecular surface produced by Connolly's program. The continuum dielectric model is evaluated with use of an adaptive multigrid boundary element method (FAMBE [22]). [The SIMS and FAMBE programs are available from the authors on request (vorobjev@femto.med.unc.edu) or via the internet (http://femto.med.unc.edu/SIMS and http://femto.med.unc.edu/FAMBE).]

As was already mentioned, we have found the ES/IS method to reliably discriminate the correctly folded protein from deliberately misfolded structures in the EMBL data base [3,4]. Averaging over microstates enhances the accuracy of the free energy estimation for the macroscopic conformation. Our results show the solvation free energy to be a significant part of the total conformational free energy. The success of the ES/IS method is determined by the accuracy and stability of the molecular dynamics model and on the efficiency, accuracy and stability of the methods used to calculate the molecular surface of the solute and for solution of the Poisson equation in the dielectric continuum model.

3. Application to misfolded proteins

3.1. Free energies of the EMBL set of misfolded protein structures

We have applied this new method of free energy calculation to the so-called EMBL set of deliberately misfolded protein structures. These misfolded protein models had been generated by selecting pairs of proteins with known structures with equal number of residues. Using the backbone conformation of the X-ray crystal structures, the sequences were swapped and the side chain structure optimized using a fast Monte Carlo algorithm with a simple energy function. The misfolded models had then been subjected to 500 steps of steepest descent energy minimization. The models were named AAA-on-BBB, where AAA is the Protein Data Bank (PDB) identifier for the sequence and BBB is the PDB identifier of the backbone coordinates [4].

The terms of the total free energy Eq. (7) have been calculated for families of the native and misfolded structures. In all cases the calculated conformational free energy, G_A is lowest for the native fold, i.e. with this formulation of the free energy, each native structure is, correctly, found to be more stable than misfolded structures of the same sequence [3].

Fig. 1 shows how in one case the total conformational free energy varies along the equilibrium trajectories for native and two misfolded structures; it is obvious that the differences in the mean values of the free energy, which for this case are 141 and 138 kcal/mol, are highly significant. The total free energy difference between the native folds and the different misfolded models varies between 11 and 190 kcal/mol.

It is interesting to separate the contributing terms into classes: (i) The electrostatic term, which is the sum of the internal Coulomb energy and the polarization free energy, $U_{m,coul} + G_{pol}$, favors the native conformation in all but two cases. (ii) The cavity term, which is the sum of the cavity free energy and the solvent–solute energy, $G_{cav} + U_{s,vdw}$, favors the native state in all but one case. (iii) The local packing energy, which is the sum of the energy for geometry deformation and the intramolecular van der Waals energy, $U_{geom} + U_{m,vdw}$ is found to always favor the native fold. (iv) The free energy term related to the protein conformational mobility, $-TS_{conf}$, favors the native fold structure in only a single instance, but makes a very small contribution to the free energy difference between native and misfolded structures.

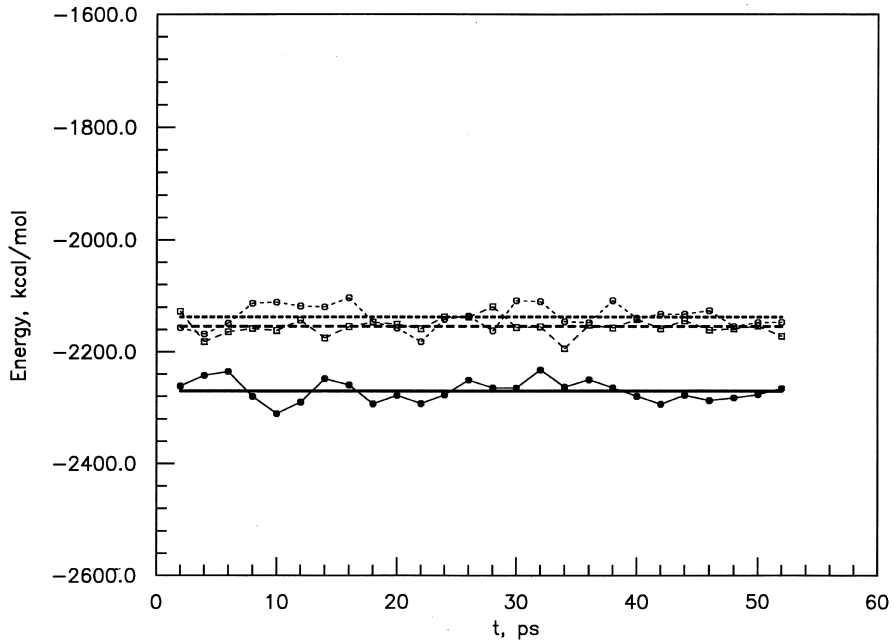

Fig. 1. Total free energy along molecular dynamics trajectories for the protein 2ci2 (native conformation, filled circles and solid line), and misfolded conformations 2ci2-on-2cro (open circles and dashed line) and 2ci2-on-1sn3 (open squares and dotted line).

4. Discussion

We wish to conclude with some comments about the method and its contributing factors, regarding the validity of various assumptions on which the ES/IS approach is based and to the possibility of improving the accuracy of the free energy estimate.

As we have seen above, three of the four different terms contribute to the discrimination between native and misfolded conformations. The small conformational entropy term is the exception, and this may correspond to the fact that our treatment of this term has not been complete (see below). The term representing the internal packing (iii above) discriminates correctly in all cases investigated. However, one should consider this correlation as no more than an incidental result, that may reflect the manner in which the misfolded conformations were prepared. The conformational free energy is a unique measure of stability; the problem in the past has been that this could not be estimated with sufficient accuracy, but this has been overcome with the approach used here. The most excellent correlation can not justify using a component of the free energy in order to replace the whole.

Analysis of known protein structures has provided heuristics that can be used as an alternative to a physics-based potential in order to score alternative folded structures of the same polypeptide chain [23–26]. The use of these scores (we deliberately avoid the word 'potentials' in connection with scores based on statistical analyses of known protein structures) is attractive because they tend to be based on rather simple models, and to be rapidly evaluated. Ultimately, the accuracy of such scores is determined by how well they track the conformational free energy, i.e. a perfect heuristic score should vary monotonically with the latter. This suggests that it will now be possible to evaluate the quality of a heuristic scoring method by computing and comparing the conformational free energy and the heuristic score of a wide series of conformations.

4.1. Model parameters

It will be important to optimize the accuracy of the model used to compute the conformational

free energy. We propose that this can be done by following the principle of self-consistency that was enunciated some time ago by Lifson as a rule to use in the development of empirical or semi-empirical force fields. In our application on misfolded proteins, the terms in Eq. (7) have been computed using independent parametrizations for three models: (i) the molecular mechanics force field used in the dynamics simulations; (ii) the parameters used to calculate the free energy of cavity formation; and (iii) the parameters used in the evaluation of the continuum dielectric model. These three parameter sets have been established independently with reference to different sets of experimental data. A way should be sought to confine the model to a self-consistent set of parameters.

We propose to achieve this using the molecular mechanics model as the reference. This model is complete, in the sense that it can be used to reproduce (with varying accuracy) any physical property not involving bond breaking and making and similar properties that require treatment with quantum mechanics. The molecular mechanics force field includes models of water molecules that allow the simulation of liquid water and aqueous solution, and techniques to compute free energies for cavity formation and solvent polarization with use of this model are well established. Therefore, it is possible to make the optimum choice of parameters for cavity formation and solvent polarization with a continuum model of solvent, by requiring agreement with free energies computed with the all-atom, molecular mechanics model, on the basis of a reference set of molecules. The accuracy of the method will then be determined by what is arguably the best currently available model. Of course, if that model is changed for the sake of greater accuracy, then the parameters used for the continuum models also will have to change.

4.2. Cavity term

Experimental data, microscopic simulations on small systems and scaled particle theory show consistently that the cavity free energy changes linearly with the surface S of the cavity [27–36]

$$G_{\mathrm{cav}} \approx \gamma_{\mathrm{micro}} S \tag{9}$$

The proportionality factor, γ_{micro} is a microscopic surface tension. It remains to determine an optimum choice for the proportionality between surface area and cavity free energy.

Simulations with an explicit water model show the free energy of creating an uncharged 'bubble' in an aqueous solution to be proportional to the macroscopic surface of the cavity, with an interfacial surface tension similar to the experimental gas-solvent surface tension, γ_{macro} for bubbles well exceeding a water molecule in size [37].

The value of the microscopic surface free energy, γ_{micro} used to compute G_{cav} is smaller because, on a molecular scale, the microscopic surface of an interface is much more irregular and somewhat larger than the corresponding macroscopic surface. For planar atomic arrays of densely packed van der Waals atomic spheres either in contact or inter-penetrating up to 30%, one finds the smooth macroscopic surface to be smaller than the irregular microscopic surface by a factor of approximately 0.66. Correspondingly, the microscopic surface free energy should be smaller than the macroscopic surface tension of water by the same factor. With γ_{macro} equal to 102 cal/(mol Å^2), this gives a value of 67 cal/(mol Å^2) for γ_{micro}, in good agreement with the estimate of 70 cal/(mol Å^2) that has been found to optimize the correlation between the results of free energy estimates and experimental data for stability and protein–protein binding of mutant proteins [38,39].

The consistency of the molecular mechanics force field and the continuum model in reproducing the cavity term is demonstrated by the agreement between the distance dependence of ΔG_{cav} for a pair of methane molecules in continuum solvent and of potentials of mean force from Monte Carlo and MD simulations in explicit solvent, Fig. 2 [40,41].

4.3. Solvent polarization

The term for the solvent polarization, G_{pol} and the intra-protein electrostatic energy, $U_{\mathrm{m,coul}}$ make very large contributions to the total conformatio-

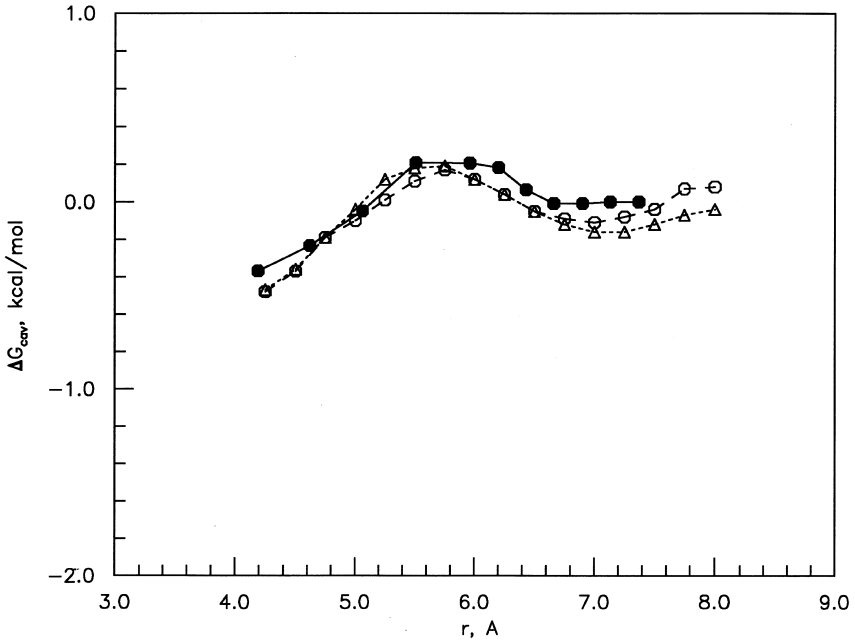

Fig. 2. The solvent-induced potential of mean force between two methane molecules in water. ●, the ES/IS method; ○, △, microscopic simulations via Monte Carlo simulation [40], and via molecular dynamics [41].

nal free energy. In addition, we have found a large anti-correlation for all studied proteins and conformations along their MD trajectories, similar to what was observed in a simulation of a pentapeptide in explicit water solvent [42]. Thus, consistency of the models used to compute these two terms is highly desirable.

The polarization free energy can be identified with the work done in a charging process in which the charges, q of the protein (superscript 1) are gradually 'turned on' by multiplication with a factor λ (lying between 0 and 1), as a result of which the solvent (superscript 0) is polarized. This can be written in terms of a polarization-related potential, V_{pol}, or in terms of summation over explicit solute–solvent Coulomb terms

$$G_{\mathrm{pol}} = \int_0^1 \mathrm{d}\lambda \sum_i \lambda q_i^1 \langle V_{\mathrm{pol}}(\mathbf{r}_i)\rangle\lambda = \int_0^1 \mathrm{d}\lambda \sum_i \lambda q_i^1 \langle \sum_j \frac{q_j^0}{r_{ij}}\rangle_\lambda \quad (10)$$

This expression is used to compute G_{pol} in simulations with thermodynamic integration or perturbation, and has been applied to numerous small molecules [6,30,43–48]. For a small solute molecule simulation times of 100 ps are needed to attain a sufficiently precise answer [47], and this is more for larger solutes. Fig. 3 shows convergence of the polarization free energy of a dipeptide from simulations in aqueous solution. Similar calculations for a molecule the size of a protein clearly require a large computational effort.

Faster calculation of the polarization free energy can be made with one or both of two approximations, the linear response model and the macroscopic electrostatic model. With a linear response approximation [15], V_{pol} varies linearly with the extent to which the charges are turned on,

$$V_{\mathrm{pol}}(\lambda) = \lambda V_{\mathrm{pol}}(\lambda = 1) \quad (11)$$

It then follows with Eq. (10) that

$$G_{\mathrm{pol}}(\lambda) = \tfrac{1}{2}\lambda^2 \sum_i q_i^1 \langle V_{\mathrm{pol}}(\mathbf{r}_i)\rangle_{\lambda=1} \quad (12a)$$

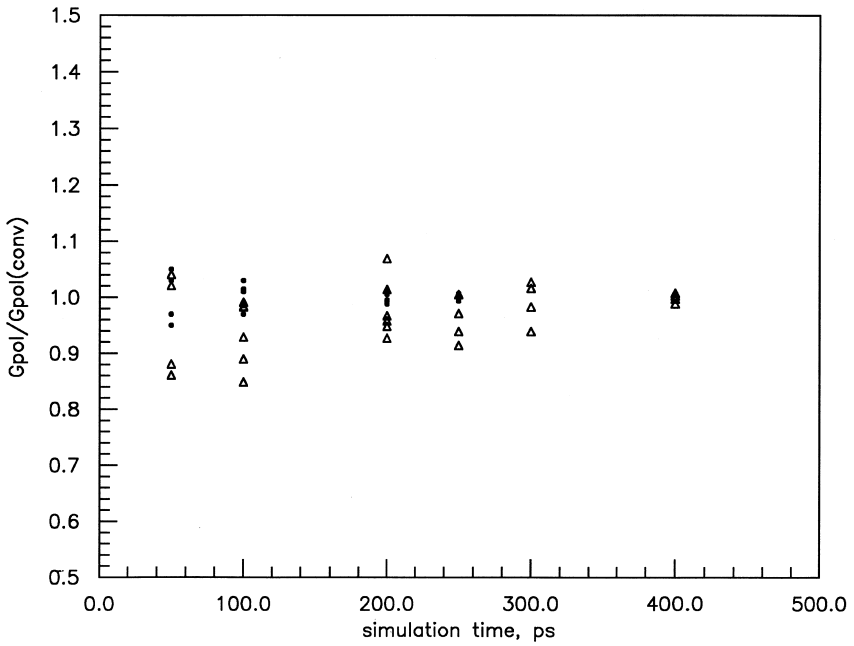

Fig. 3. Reproducibility of values of the polarization free energy computed via slow charging as a function of simulation time for dipeptides Ace-X-Nme in explicit SPC-water (X = Gly, Ala, Val, Glu^-, Lys^+, Arg^+). △, Polar groups; and ●, charged side chain groups.

$$G_{\mathrm{pol}} = \tfrac{1}{2}\sum_i q_i^1 \langle V_{\mathrm{pol}}(\mathbf{r}_i)\rangle_{\lambda=1}$$

$$= \tfrac{1}{2}\sum_i q_i^1 \langle \sum_j \frac{q_j^0}{r_{ij}} \rangle_{\lambda=1} = \tfrac{1}{2} U_C^{0,1} \qquad (12b)$$

i.e. when a linear response model holds, the polarization free energy is one-half the solute–solvent Coulomb energy. When this is assumed to hold, the polarization free energy can be estimated from simulations for a single solute molecule charge distribution, and there is no need to explicitly perform the charging process [48–50]. Even then, the solvent potential V_{pol} in a simulation of lysozyme in explicit solvent showed slow convergence [51], and low precision achieved in practice may outweigh the higher accuracy achievable in principle, relative to use of a continuum dielectric model.

In a majority of molecular dynamics simulations of polar and charged molecules a linear response was observed [30,43–48]. Apparently, it is sufficient for the solute's partial charges to not exceed 1 e, and for the electrostatic field near the solute surface not to exceed 50 kT/(e Å) [46]. Fig. 4 shows the results of simulations of dipeptides with charged and polar groups held in fixed conformations. The solute–solvent electrostatic energy is found to be reasonably closely proportional to λ^2, as is expected on the basis of a linear response model, Eq. (12a), and these results provide a basis for use of models in which the solvent is treated as a continuum dielectric with linear polarization.

According to results of these simulations (data not shown), equating one-half the solute–solvent Coulomb energy computed in simulations at full charge with the polarization free energy, which is correct for the linear response model [Eq. (12b)], overestimates the polarization free energy by 20% for a polar group and underestimates it by 10% for charged groups. These relative errors translate into absolute errors of approximately 1 kcal/mol for polar groups and 10 kcal/mol for charged groups, and hence, the use of the linear

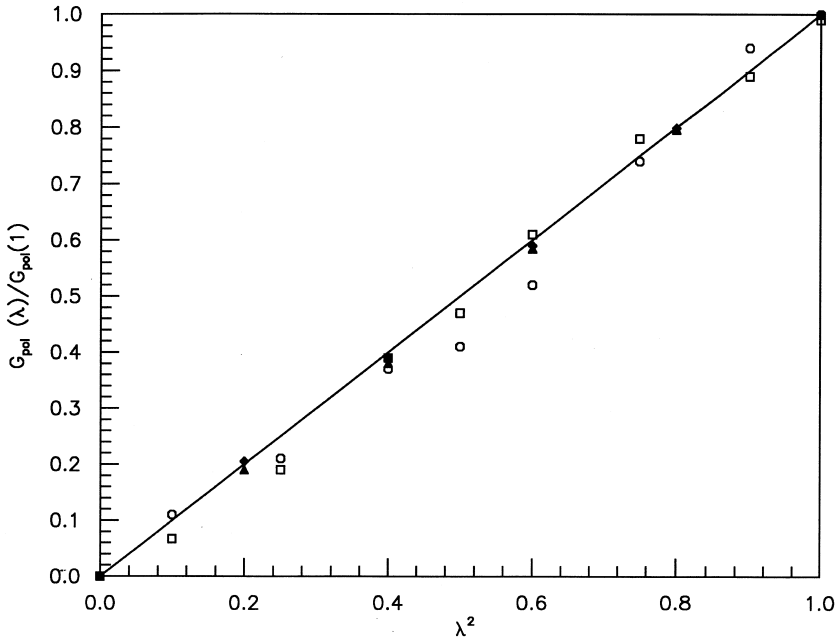

Fig. 4. The polarization free energy computed via the slow charging method [Eq. (10)] as a function of the square of the fraction of charging, λ, for dipeptides Ace-X-Nme, where X = Gly(○), Ala(□), Lys^+(▲) and Glu^-(■).

response model can easily introduce relatively large errors.

It remains to be seen if these errors can be smaller with use of a continuum dielectric model. The polarization free energy computed with the continuum dielectric model that we have used depends on parameters, namely, the atomic partial charges, and the atomic radii and the radius of the solvent molecule which together define the boundary between the interior of the solute (where the dielectric constant is low) and the solvent (where the dielectric constant is high). In accord with the principle of self-consistency of the force field, the partial charges have been chosen equal to those used in the molecular mechanics model, which leaves only the atomic radii as adjustable parameters. These can be chosen to optimize the consistency of the explicit and implicit models, with use of a test set of molecules (including some macromolecules) in a number of conformations, for which precise polarization free energies can be computed via a charging process [cf. Eqs. (12a) and (12b)]. This test set then replaces the sets of experimental solvation free energies that have traditionally served as reference for optimization of the parameters of continuum dielectric models [11,43].

4.4. Macromolecular entropy

In the calculations reported here, the free energy term related to the protein conformational mobility, $-TS_{conf}$ makes only a small contribution [3] to the total conformational free energy of Eq. (7). This is not unexpected, as only a small fraction of the modes correspond to global motions, and only these can be expected to depend significantly on the overall molecular fold [3,16,17,52]. Because of computational limitations, the mobility of only the C^{α} atoms was used in the analysis. It remains to be determined if this is sufficient, or if the mobility of other atoms should be included. In particular, the freedom of side chains to assume multiple conformations may have to be represented. As a first step, this may be done explicitly by assessing the freedom of

surface-exposed side chains to assume different conformers in terms of an absence of bad contacts with other parts of the protein.

Improvements of the accuracy of the conformational free energy calculated with the ES/IS model, and application to other problems of biological structure are the subject of ongoing work.

Acknowledgements

This work was supported by a grant from the National Center for Research Resources of the National Institutes of Health (RR08102) and by a grant from the National Science Foundation (MCB-9314854).

References

[1] P. Kollman, Free energy calculations: applications to chemical and biochemical phenomena. Chem. Rev. 93 (1993) 2395–2417.

[2] R.J. Radmer, P.A. Kollman, Free energy calculation methods: a theoretical and empirical comparison of numerical errors and a new method for qualitative estimates of free energy changes. J. Comp. Chem. 18 (1997) 902–919.

[3] Y.N. Vorobjev, J.C. Almagro, J. Hermans, Discrimination between native and intentionally misfolded conformations of proteins: ES/IS, a new method for calculating conformational free energy that uses both dynamic simulations with explicit solvent, and an implicit solvent continuum model. Proteins: Struct. Funct. Genet. 32 (1998) 399–413.

[4] J.A. Moult, Pro-Star: the protein potential site. A resource, exchange and discussion forum for development and application of potentials for protein structure evaluation and prediction. Website http://prostar.carb.nist.gov/, 1997.

[5] D. Frenkel, Molecular dynamic simulations of statistical mechanical systems. Proceedings of the Enrico Fermi Summer School, Varenna, 1985. North-Holland, Amsterdam, 1986, pp. 151–164.

[6] V. Helms, R.C. Wade, Free energies of hydration from thermodynamic integration: comparison of molecular mechanics force fields and evaluation of calculation accuracy. J. Comp. Chem. 18 (1997) 449–462.

[7] T.L. Hill, An Introduction to Statistical Thermodynamics, Dover, New York, 1986.

[8] D. Chandler, L.R. Pratt, Statistical mechanics of chemical equilibria and intramolecular structures of non-rigid molecules in condensed phases. J. Chem. Phys. 65 (1976) 2925–2940.

[9] A. Ben-Naim, Solvent effects on protein association and protein folding. Biopolymers 29 (1990) 567–596.

[10] B. Honig, K. Sharp, A.S. Yang, Macroscopic models of aqueous solutions: biological and chemical applications. J. Phys. Chem. 97 (1993) 1101–1109.

[11] D. Sitkoff, K.A. Sharp, B. Honig, Accurate calculation of hydration free energies using macroscopic solvent models. J. Phys. Chem. 98 (1994) 1978–1988.

[12] T. Simonson, A. Brünger, Solvation free energies estimated from macroscopic continuum theory: an accuracy assessment. J. Phys. Chem. 98 (1994) 4683–4694.

[13] J. Langlet, P. Claverie, J. Caillet, A. Pullman, Improvements of the continuum model. 1. Application to the calculation of the vaporization thermodynamic quantities of non-associated liquids. J. Phys. Chem. 92 (1988) 617–1631.

[14] J. Tomasi, M. Persico, Molecular interactions in solution: an overview of methods based on continuum distribution of the solvent. Chem. Rev. 94 (1994) 2027–2094.

[15] J.P. Hansen, I.R. McDonald, Theory of Simple Liquids, Academic Press, New York, 1986, pp. 245–246.

[16] A. Amadei, A.B.M. Linsen, H.J.C. Berendsen, Essential dynamics of proteins. Proteins: Struct. Funct. Genet. 17 (1993) 412–425.

[17] A.E. Garcia, Large-amplitude nonlinear motions in proteins. Phys. Rev. Lett. 68 (1992) 2696–2699.

[18] B.R. Brooks, D. Janezic, M. Karplus, Harmonic analysis of large systems. I. Methodology. J. Comp. Chem. 16 (1995) 1522–1542.

[19] D. Janezic, B.R. Brooks, Harmonic analysis of large systems. II. Comparison of different protein models. J. Comp. Chem. 16 (1995) 1543–1553.

[20] D. Janezic, R.M. Venable, B.R. Brooks, Harmonic analysis of large systems. III. Comparison with molecular dynamics. J. Comp. Chem. 16 (1995) 1554–1566.

[21] Y.N. Vorobjev, J. Hermans, SIMS, Computation of a smooth invariant molecular surface. Biophys. J. 73 (1997) 722–732.

[22] Y.N. Vorobjev, H.A. Scheraga, A fast adaptive multigrid boundary element method for macromolecular electrostatics in a solvent. J. Comp. Chem. 18 (1997) 569–583.

[23] C.M.R. Lemer, M.J. Rooman, S.J. Wodak, Protein structure prediction by threading methods: evaluation of current techniques. Proteins: Struct. Funct. Genet. 23 (1995) 337–355.

[24] J.P. Kocher, M.J. Rooman, S.J. Wodak, Factors influencing the ability of knowledge-based potentials to identify native sequence-structure matches. J. Mol. Biol. 235 (1994) 1598–1613.

[25] P.D. Thomas, K. Dill, Statistical potentials extracted from protein structures: what is wrong with them?. J. Mol. Biol. 257 (1996) 457–469.

[26] R.L. Jernigan, I. Bahar, Structure derived potentials and protein simulations. Curr. Opin. Struct. Biol. 6 (1996) 195–209.

[27] A. Ben-Naim, Y. Marcus, Solvation thermodynamics of nonionic solutes. J. Chem. Phys. 81 (1984) 2016–2027.

[28] C.H. Chothia, Hydrophobic bonding and accessible area in proteins. Nature 248 (1974) 338–339.
[29] R.B. Hermann, Theory of hydrophobic bonding. II. The correlation of hydrocarbon solubility in water with solvent cavity surface area. J. Phys. Chem. 76 (1972) 2754–2759.
[30] G. Hummer, L.R. Pratt, A.E. Garcia, Hydration free energy of water. J. Phys. Chem. 99 (1995) 14188–14194.
[31] G. Hummer, L.R. Pratt, A.E. Garcia, Free energy of ionic hydration. J. Phys. Chem. 100 (1996) 1206–1215.
[32] W. Wallqvist, B.J. Berne, Computer simulation of hydrophobic hydration forces on stacked plates at short range. J. Phys. Chem. 99 (1995) 2893–2899.
[33] W. Wallqvist, B.J. Berne, Molecular dynamics study of the dependence of water solvation free energy on solute curvature and surface area. J. Phys. Chem. 99 (1995) 2885–2892.
[34] H. Reiss, H.L. Frisch, E. Helfand, J.L. Lebovitz, Aspects of the statistical thermodynamic of real fluids. J. Chem. Phys. 32 (1960) 119–124.
[35] R.A. Pierotti, A Scaled particle theory of aqueous and non-aqueous solutions. Chem. Rev. 76 (1976) 717–726.
[36] R.M. Jackson, J.E. Sternberg, Application of scaled particle theory to model the hydrophobic effect: implications for molecular association and protein stability. Protein Eng. 7 (1994) 371–383.
[37] J.P.M. Postma, H.J.C. Berendsen, J.R. Haak, Thermodynamics of cavity formation in water. Faraday Symp. Chem. Soc. 17 (1982) 55–67.
[38] R.M. Jackson, J.E. Sternberg, A continuum model for protein–protein interactions: applications to the docking problem. J. Mol. Biol. 250 (1995) 258–275.
[39] J. Novotny, R.E. Brucooleri, M. Davis, K.A. Sharp, Empirical free energy calculations: a blind test and further improvements of the method. J. Mol. Biol. 268 (1997) 401–411.
[40] C. Pangali, M. Rao, B.J. Berne, A Monte Carlo simulation of the hydrophobic interaction. J. Chem. Phys. 71 (1979) 2975–2981.
[41] D.E. Smith, A.D.J. Haymet, Free energy, entropy and internal energy of hydrophobic interactions: computer simulations. J. Chem. Phys. 98 (1993) 6445–6454.
[42] E. Demchuk, D. Bashford, G.P. Gippert, Thermodynamics of reverse turn motif. Solvent effects and side-chain packing. J. Mol. Biol. 270 (1997) 305–317.
[43] M. Nina, D. Beglov, B. Roux, Atomic radii for continuum electrostatic calculations based on molecular dynamics free energy simulations. J. Phys. Chem. 101 (1997) 5239–5248.
[44] B. Roux, H.A. Yu, M. Karplus, Molecular basis for the Born model of ion solvation. J. Phys. Chem. 94 (1990) 4683–4688.
[45] R.M. Levy, M. Belhadj, D.B. Kitchen, Gaussian fluctuation formula for electrostatic free energy changes in solution. J. Chem. Phys. 95 (1991) 3627–3633.
[46] B. Jayaram, R. Fine, K. Sharp, B. Honig, Free energy calculations of ion hydration: an analysis of the Born model in terms of microscopic simulations. J. Phys. Chem. 93 (1989) 4320–4327.
[47] S.W. Rick, B.J. Berne, The aqueous solvation of water: a comparison of continuum methods with molecular dynamics. J. Am. Chem. Soc. 116 (1994) 3949–3954.
[48] J. Aqvist, T. Hansson, On the validity of electrostatic linear response in polar solvent. J. Phys. Chem. 100 (1996) 9512–9521.
[49] J. Aqvist, C. Medina, J.E. Samuelsson, A new method for predicting binding affinity in computer-aided drug design. Protein Eng. 7 (1994) 385–391.
[50] H.A. Carlson, W.L. Jorgensen, An extended linear response method for determining free energies of hydration. J. Phys. Chem. 99 (1995) 10667–10673.
[51] F. Figueirido, G.S. Del Buono, R.M. Levy, Dielectric behavior in the interior of proteins: result from a large-scale molecular dynamic simulation of solvated lysozyme. Biophys. J. 72 (1997) A216.
[52] D.M.F. van Aalten, B.L. de Groot, J.B.C. Findlay, H.J.C. Berendsen, A. Amadei, A comparison of techniques for calculating protein essential dynamics. J. Comp. Chem. 18 (1997) 169–181.