Modeling solvation contributions to conformational free energy changes of biomolecules using a potential of mean force expansion

Matteo Pellegrini
Department of Physics, Stanford University, Stanford, California 94305-4090

Sebastian Doniach
Department of Physics and Department of Applied Physics, Stanford University, Stanford, California 94305-4090

(Received 2 February 1995; accepted 7 April 1995)

The standard free energy perturbation (FEP) techniques for the calculation of conformational free energy changes of a solvated biomolecule involve long molecular dynamics (MD) simulations. We have developed a method for performing the same calculations many orders of magnitude faster. We model the average solvent density around a solute as the product of the relevant solute–solvent correlation functions (CF), following the work of García, Hummer, and Soumpasis. We calculate the CF's by running Monte Carlo simulations of a single solute atom in a box of explicit water molecules and also angular dependent CF's for selected pairs of solute atoms. We then build the water shell around a larger solute (e.g., alanine dipeptide) by taking the product of the appropriate CF's. Using FEP techniques we are able to calculate free energy changes as we rotate the dihedral angles of the alanine dipeptide and we find they are in close agreement with the MD results. We also compute the potential of mean force as a function of distance between two solvated methanes and calculate the contribution of the solvent to the free energy change that results from rotating n-butane about its dihedral angle. © *1995 American Institute of Physics.*

I. INTRODUCTION

Recent advances in molecular dynamics (MD) simulations coupled with free energy perturbation (FEP) techniques render the calculation of conformational free energy changes of solvated biomolecules quite feasible.[1] However these simulations typically require placing the solute molecule in a box of hundreds of explicit water molecules, and then running 100 000 MD steps to get accurate statistics. With the computational power of present computers this entails running a simulation for days. It would be useful then to develop techniques that are able to approximate the MD results in a fraction of the time.

Recently it was shown by Hummer and Soumpasis[2] that they could accurately reproduce the density distribution of water at the water–ice boundary using a potential of mean force expansion. They accomplished this by calculating the two- and three-point correlation functions (CF) between all possible combination of oxygens and hydrogens (OO, OH, HH, OOO, OOH, OHH, and HHH) in the liquid state. Although these CF calculations involve lengthy Monte Carlo (MC) runs, once they have been calculated the computation of the density is accomplished in seconds. They have also used this same approach to calculate the water structure around DNA, and find that it successfully reproduces many known experimental features.[3–5]

We have applied a similar approach to build the mean water density around small biomolecules. We then extend the previous work to compute the effect of this mean field water on free energy changes associated with conformational changes in our solute molecules. The free energy changes are computed using FEP techniques.

We calculate the two-point CF's between single solute atoms and water oxygens or hydrogens. We also compile angular dependent CF's, with cylindrical symmetry, for water oxygens or hydrogens relative to pairs of solute atoms. We can then calculate the water density around any simple solute that can be built up out of our two-point and angular dependent CF's. As we perturb the conformation of our solute we compute the change in solvation energy with respect to the mean water density of the unperturbed state. This FEP procedure allows us to build up the component due to solvation of the free energy profiles for any conformational change in the solute.

We demonstrate the usefulness of this technique on three systems: (1) The potential of mean force between two solvated methanes, as a function of the distance between them. (2) The solvent contribution to the free energy change associated with the rotation of n-butane about its dihedral angle. (3) The free energy change necessary to rotate the alanine dipeptide about its dihedral angles along four separate trajectories.

II. METHODS

A. Potentials of mean force expansion

Following the derivation of Hummer and Soumpasis we first develop the expression for the density of water around a solute as a function of the solute–solvent CF's. The conditional density of a simple monoatomic liquid may be expressed as

$$\rho^{(r|r_1\cdots r_n)}=\rho\,\frac{g^{(n+1)}(r,r_1,\ldots,r_n)}{g^{(n)}(r_1,\ldots,r_n)},\tag{1}$$

where $g^{(n)}$ is the n-particle correlation function and ρ is the bulk density. In the case of water one would have two separate expressions for the density of oxygen, ρ_{O}, and hydrogen,

0021-9606/95/103(7)/2696/7/$6.00

ρ_H. In our situation the coordinates $r_1,\ldots,r_n$ would be those of the solute, while r would be the position at which we compute the density of water oxygens and hydrogens. The potential of mean force (PMF) is defined as

$$W^n(r_1,\ldots,r_n) = -k_B T \ln g^{(n)}(r_1,\ldots,r_n). \tag{2}$$

The PMF may then be expressed in terms of few body potentials (see, e.g., Munster[6]) as

$$W^n(r_1,\ldots,r_n) = \sum_i w^{(1)}(r_i) + \sum_{i>j} \sum_j w^{(2)}(r_i,r_j) + \sum_{i>j>k} \sum_{j>k} \sum_k w^{(3)}(r_i,r_j,r_k) + \cdots \tag{3}$$

with

$$w^{(1)}(r_i) = W^{(1)}(r_i), \tag{4}$$

$$w^{(2)}(r_i,r_j) = W^{(2)}(r_i,r_j) - W^{(1)}(r_i) - W^{(1)}(r_j), \tag{5}$$

$$\begin{aligned} w^{(3)}(r_i,r_j,r_k) = {} & W^{(3)}(r_i,r_j,r_k) - W^{(2)}(r_i,r_j) \\ & - W^{(2)}(r_i,r_k) - W^{(2)}(r_j,r_k) \\ & + W^{(1)}(r_i) + W^{(1)}(r_j) + W^{(1)}(r_k). \end{aligned} \tag{6}$$

In the case of a fluid $W^{(1)}r=0$. Retaining only two- and three-body terms in the expansion we then write the conditional density as

$$\rho^{(r|r_1\cdots r_n)} = \rho \prod_{i=1}^{n} g^{(2)} \times (r,r_i) \prod_{j=1}^{n-1} \prod_{k=j+1}^{n} \frac{g^3(r,r_j,r_k)}{g^2(r,r_j) g^2(r_j,r_k) g^2(r,r_k)}. \tag{7}$$

If we apply the Kirkwood approximation[7]

$$g^{(3)}(r_1,r_2,r_3) = g^{(2)}(r_1,r_2) g^{(2)}(r_2,r_3) g^{(2)}(r_1,r_3) \tag{8}$$

to the above formulation of the density then we are simply approximating the density as a product of two-point CF's. In other words, we are treating a many-body PMF as a sum of two-body PMF's. For simplicity we will refer to the use of this approximation as a first order expansion while writing the density as a product of two- and three-point CF's, as in Eq. (7), as a second order expansion.

If we wish to calculate the water density around a single hydrophobic solute atom it is sufficient to use a first order expansion since the density is spherically symmetric. We have also found that to compute the PMF between two methanes this approximation is sufficient to reproduce MD results effectively. However if the solute is more complex then one must include second order terms or three-point CF's to correctly account for the density.

For instance, a first order expansion of Eq. (4) is unable to correctly account for the density of liquid water around a fixed water molecule. Because of the nature of the hydrogen bond, water molecules, even in the liquid state, tend to form tetrahedral structures; in these configurations the water oxygens form an isosceles triangle with two sides of length 2.8 Å and the third 4.5 Å. Since the water oxygen–oxygen two-point correlation function has a peak at approximately 2.8 Å, following the Kirkwood approximation an equilateral triangle of water oxygens separated by 2.8 Å would be strongly favored. However, empirically water is found preferentially in the above isosceles triangle configuration while the equilateral triangle one is strongly disfavored, showing the importance of the three-point CF's. Hummer[2] found that a second order expansion that includes both two- and three-point CF's is sufficient to reproduce this effect and that the density of water around a fixed water molecule computed this way reproduces well the MC results.

TABLE I. OPLS parameters used in simulations.

Atom type	Charge qe	ϵ	σ
CH_4	0.0	3.73	0.294
CH_3	0.0	3.905	0.175
CH_2	0.0	3.905	0.118
C_α	0.2	3.8	0.08
C(O)	0.5	3.75	0.105
O(C)	−0.5	2.96	0.21
N(H)	−0.57	3.25	0.17
H(N)	0.37	0.0	0.0
C_β	0.0	3.91	0.16
(NH)CH_3	0.2	3.8	0.17

In the case of the alanine dipeptide the solute has a complicated charge distribution therefore we are forced to extend our expansion to second order. Since the alanine dipeptide contains 12 atoms (in a united atom model) this involves the computation of many three-point CF's. We further simplify the problem by choosing to calculate only the CF's associated with the strongest dipoles (the CO, C_α C_β, and NH groups and water oxygens or hydrogens) while approximating all other three-point CF's by using the Kirkwood approximation. This may appear to be a rather limiting approximation; however we will show that it works reasonably well in reproducing MD results for certain conformations of the alanine dipeptide but that additional three-point CF's are needed in the region of the (ϕ,ψ) plane where intramolecular hydrogen bonds are involved.

B. Calculation of correlation functions

To compute the correlation functions (CF) we placed our solute (e.g., united atom methane) in a 19 Å box with periodic boundary conditions. We then solvated it with SPC (Ref. 8) water molecules at a density of 1 g/cm^3.

The potential functions and parameters we used were those developed by Jorgensen and Tirado-Rives, since these are optimized for liquid simulations (hence known as OPLS potential functions).[9,10] The form of the potential function is

$$V(r_1,\ldots,r_n) = \sum_{i,j,i>j} \left(\frac{q_i q_j e^2}{r_{ij}} + \frac{A_{ij}}{r_{ij}^{12}} - \frac{C_{ij}}{r_{ij}^6} \right). \tag{9}$$

In all cases the CH_n groups are treated as united atoms, so that only one position is used to describe these groups (see Table I).

As suggested by Levitt,[11] the Coulomb component of the potential is truncated by multiplying it by the function

$S(r)=[1-(r/r_c)]^2$, with $r_c=8.0$ Å. This function eliminates spurious forces at the cutoff distance since its derivative also goes to zero there. We have verified that water–water CF's calculated using $S(r)$ are not significantly different than those computed using an unscreened Coulomb potential and that furthermore they reproduce the experimental (obtained from neutron diffraction) water–water CF's well.[12,13]

To select moves (translations and rotations) of molecules we use the standard Monte Carlo Metropolis algorithm.[14] The moves were selected from a flat distribution with ranges ±0.15 Å for translations and ±15° for rotations. At the chosen simulation temperature (25 °C) these moves yield and acceptance rate of approximately 40%.

In the case where our test solute consists of a pair of atoms with partial charges we approximate the three-point CF between the two solute atoms and water oxygens or hydrogens as cylindrically symmetric angular dependent CF's. The distance between the two solute atoms is held fixed throughout the simulation. We treat the midway point between the pair as the center, and calculate separate CF's for 15 degree intervals, from 0 to 180 deg. At each angular interval we calculate the distance dependence with bin widths of 0.1 Å (this same value is used for the spherically symmetric CF's). Thus we tabulate 12 different CF's for all the selected pairs.

C. Free energy perturbation

To calculate the free energy change associated with a conformational change in the solute we use free energy perturbation (FEP) theory.[15] The change in free energy due to a change in angle $\delta\phi_i$ can be written as

$$\Delta A = A(\phi_i+\delta\phi_i)-A(\phi_i)$$
$$=-\beta^{-1}\ln\langle\exp\{-\beta[U(\phi_i+\delta\phi_i)-U(\phi_i)]\}\rangle_{\phi_i}. \quad (10)$$

Here ϕ_i represents any internal coordinate or coordinates and $\langle\ \rangle$ defines an equilibrium average over all other degrees of freedom with ϕ_i constrained.

To apply this formula we calculate the water shell around the solute that is in its conformation ϕ_i. We construct a cubic grid with 0.5 Å spacings extending 8.0 Å from the solute and compute the density of water oxygens and hydrogens at each grid point. If, for example, there are no internal degrees of freedom in the solute other than ϕ, then we need not average over any other degrees of freedom since our water shell is already an average water distribution. We can therefore immediately calculate the free energy change by computing $U(\phi_i+\delta\phi_i)-U(\phi_i)$ with the water shell fixed. We calculate $U(\phi_i)$ by summing the OPLS interaction energy between the solute atoms and the grid points with their respective densities of water oxygens and hydrogens.

To calculate the free energy profiles we decompose the trajectory into N intervals (for example 15 in the case of the alanine dipeptide). We fix the solute in the conformation ϕ_n, and compute the water shell around it. We then perturb ϕ to its intermediate value between ϕ_n and ϕ_{n+1} and compute the energy between the solute and the water shell of ϕ_n. We also perturb ϕ to the intermediate value between ϕ_n and ϕ_{n-1} and again calculate the energy with respect to the water shell of ϕ_n. We repeat this procedure to obtain $2N$ values of ΔA which we then sum to construct a free energy profile.

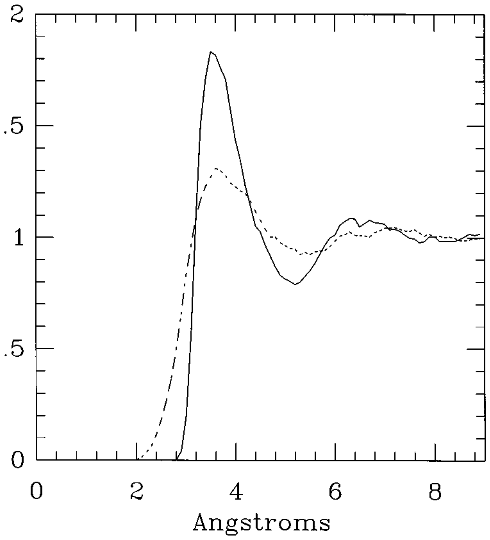

FIG. 1. Two-point correlation function between a united-atom methane and water oxygens (solid) and hydrogens (dashed) obtained from MC simulations.

III. RESULTS AND DISCUSSION

A. Potential of mean force between two methanes

The potential of mean force (PMF) between two methanes has been studied extensively because it is one of the simplest potentials that describes the hydrophobic effect.[16] As expected from previous work[17] the ordering of water around a hydrophobic solute creates an effective attraction between two methanes since reducing their surface area in contact with water is energetically favored. Previous studies had also found that a second minimum in the PMF occurs when the two methanes are separated by about 7 Å, allowing a single water molecule to lie between them.

Using the OPLS parameters for a united-atom methane (see Table I) we first calculate the CF's between methane and water oxygens or hydrogens (see Fig. 1). For a given methane–methane distance we build a water shell around the two methanes by calculating the oxygen density ρ_O and hydrogen density ρ_H, each as a product of the two CF's, as explained above. We compute the density of the water oxygens or hydrogens on a cubic grid with 0.5 Å spacings that extends 8 Å from the methanes. We then construct a series of new water shells as the separation between the methanes grows from 3.5 to 8 Å in 0.1 Å steps.

Using FEP theory in the form

$$\Delta A = A(r+\delta r)-A(r)$$
$$=-\beta^{-1}\ln\langle\exp\{-\beta[U(r+\delta r)-U(r)]\}\rangle_r,$$

where $U(r)$ is the sum of the standard OPLS energy between the solute and ρ_O and ρ_H at each grid point, we were able to calculate the PMF. We find remarkable agreement with previous MD studies,[16] as can be seen in Fig. 2. The position of

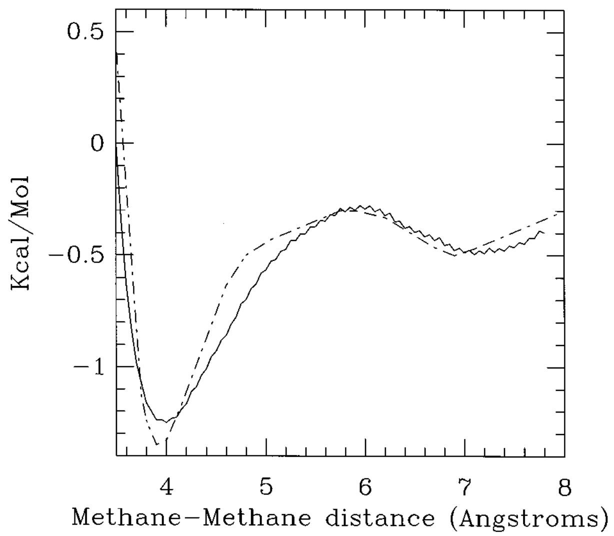

FIG. 2. Potential of mean force as a function of distance between two solvated methanes. The dashed line is from the MD simulation of van Belle (Ref. 16). The solid line is from this work.

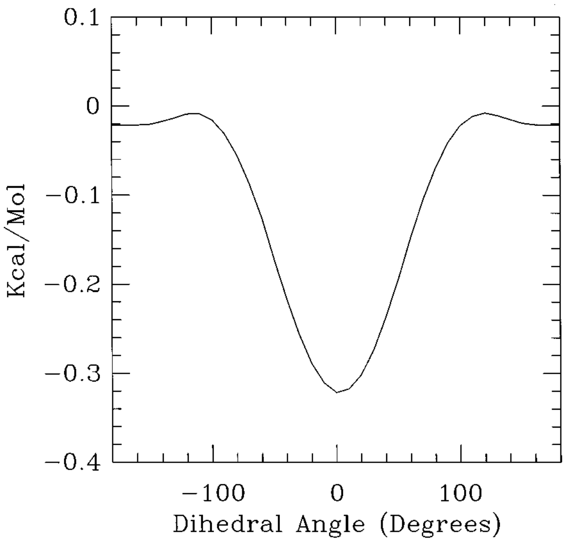

FIG. 3. Solvent contribution to the free energy associated with the rotation of n-butane about its dihedral angle. As can be seen the *trans* and *gauche* conformations are equally favored by water.

the minima, at 4 and 7 Å, are equal to within 0.2 Å in the two curves. The barrier heights are within 0.2 kcal/mol of each other. Thus in this very simple case the methodology is effective in reproducing the results of an extensive MD simulation.

It should also be noted that to construct our PMF requires only a few seconds, while the evaluation of PMF's using MD may require a time many orders of magnitude longer. For instance Belle[16] reports that to construct the curve in Fig. 2 he ran a 670 ps simulation with 2 fs time steps. Thus he evaluated 335 000 time-steps to our 45. To estimate the computational difference of timesteps in the two methods we consider that we have approximately 300 grid points (using 0.5 Å spacings) for every explicit water molecule used in MD simulations. The explicit waters in MD however interact with approximately 60 other waters (using an 8 Å cutoff) while our grid points do not interact with each other but only with the solute. We conclude that the energy calculation of our method at each timestep involves approximately five times as many computations as his. Therefore, given the difference in number of timesteps, we require three orders of magnitude fewer computations to calculate the same curve.

B. Effect of solvation on *n*-butane

n-Butane is a small hydrophobic molecule with only one major degree of freedom, its dihedral angle. In the extended-H or united atom model it can be described by only four united carbon atoms. This is the model we chose to study using the OPLS parametrization[10] (see Table I).

A recent study of the effects of solvation on the conformational states of n-butane by Tobias and Brooks[18] found that using this description the relative populations of the *gauche* and *trans* conformers were equal in vacuum and water; they report an equilibrium constant $K=(x_g/x_t)=0.54$ in both the gas and water phase. However it should be noted that these results are strongly model dependent. For instance Brooks[18] reports that in an all atom model of n-butane (as opposed to a united atom model) the water tends to favor the more compact *gauche* over the more extended *trans* conformation; in this case $K=0.59$ in the gas phase and $K=0.85$ in the water phase.

We computed angular dependent three-point CF's for the CH_3, CH_2 groups of n-butane described by the OPLS parameters. We construct an n-butane using ideal bond angles and bond lengths. During the simulation we freeze all degrees of freedom (bond angle, bond length, etc.) other than the dihedral angle. We then calculate the density of water oxygens and hydrogens on a 0.25 Å cubic grid around the n-butane as we rotate the dihedral angle from −180 to 180 deg in 10 deg steps. Using FEP we compute the corresponding free energy changes that result from this rotation. For clarity we only consider the free energy change due to solvation and neglect internal dihedral potentials of n-butane.

In Fig. 3 we show the solvent contribution to the free energy profile about the dihedral angle of n-butane. We have symmetrized the profile about the angle $\psi=0.0$. We find that the free energy has two minima; one in the *cis* and one in the *trans* conformation. The *trans* is approximately 0.3 kcal/mol higher than the *cis*, and the barrier between them is approximately 0.3 kcal/mol (this is within 25% of the barrier found by Roux[19]).

The *gauche* conformer occurs when the dihedral angle is approximately 60.0°. Thus in our profile the *trans* conformer is higher than the *gauche* by 0.1 kcal/mol. Our results agree to within 0.1 kcal/mol to those of Brooks[18] who finds that the relative populations of the *gauche* and *trans* conformers are unaffected when going from the vacuum to the water phase.

Thus in our calculations the hydrophobic effect favors the more compact *gauche* conformation over the more extended trans by 0.1 kcal/mol. Experiments are not yet able to

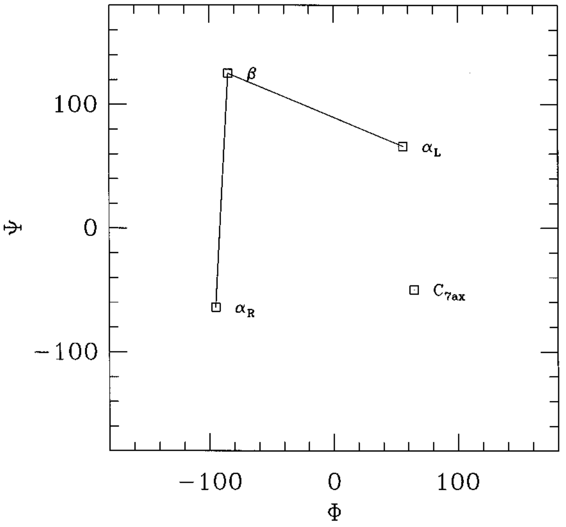

FIG. 4. The $\phi-\psi$ plot of the alanine dipeptide conformations that correspond to equilibria. The lines denote the constrained trajectories of Figs. 5 and 6. The exact values of the minima are $\alpha_R=(-95,-65)$, $\beta=(-85,125)$, $\alpha_L=(55,65)$, and $C_{7ax}=(65,-50)$.

measure the relative populations of these conformations in water, thus we are not able to verify this prediction.

C. Solvation of the alanine dipeptide

The alanine dipeptide, *N*-acetylalanyl-*N*-Methylamide, has been extensively studied because of its simplicity and its relevance to protein chemistry.[20–23] In the extended atom representation it consists of only 12 atoms whose major degrees of freedom are rotations around the dihedral ϕ and ψ angles. There are four primary minima in the $\phi-\psi$ space; these conformations are labeled as β, α_R, α_L, and C_{7ax} (the β region however has several minima). Various groups have reported on the effect of solvation on the relative free energy of these states. Recently Tobias and Brooks have reported on the free energy change along the four pathways (see Fig. 4) connecting the minima using standard MD techniques.[20] Using the above procedure we have calculated the free energy change along two of these pathways which we then compared to the previous work.

1. Correlation functions

As mentioned above, we extend the potential of mean field expansion to second order for the case of the alanine dipeptide. However since this would entail the calculation of many three-point CF's we further simplify the problem by approximating most of the three-point CF's as the product of the two-point CF's as prescribed by the Kirkwood approximation. Therefore we only calculate the three-point CF's explicitly for the cases where we expect the Kirkwood approximation to break down. We expect this to occur for the three-point CF's between pairs of solute atoms with a strong dipole moments and water oxygens or hydrogens. This choice is consistent with the work Hummer who found that the liquid water density around ice, which has a strong dipole moment, was not well approximated by Kirkwood and that the inclusion of three-point CF's greatly improved the results. Furthermore, we have already shown that when dealing with uncharged atoms, such as *n*-butane, two-point CF's are sufficient.

Applying the above criterion, we find that using OPLS parameters the largest dipoles in most conformations of the alanine dipeptide arise from the CO, $C_\alpha C_\beta$, and NH pairs (see Table I). We compute angular dependent CF's, which are approximations of three-point CF's, for these three pairs. For the remaining two single atoms (the two methyl groups in a united atom description) we compute spherically symmetric two-point CF's.

In one conformation, which is labeled C_{7ax}, there is an intramolecular hydrogen bond between a carboxyl oxygen and an amide hydrogen. These two atoms form a strong dipole and hence to get reasonable accuracy we would need to compute the three-point CF between these two atoms (as a function of their spacing) and water oxygens or hydrogens. In the present work however we choose not to explore the free energy profiles around this conformation, and will leave this for future work.

A consequence of our use of only a limited number of explicit three-point CF's is that at certain positions around the alanine dipeptide in any conformation our procedure computes unreasonably large water oxygen and hydrogen densities (more than 50 times the bulk water density). To obtain reasonable answers in our free energy calculations we therefore truncate all densities above 10 times the bulk density to 10. This value is arrived at empirically by trying different values and comparing the resulting free energy profiles of our two trajectories to the MD results. We should however point out that these profiles are not very sensitive to the value of this parameter; changing the value from 10 to 15 changes the profiles by less than 1 kcal/mol. A further motivation for the use of the value 10 is that it corresponds to the largest density, or peak in observed CF, that we find in our MC simulations with explicit water that we run to compute the CF's.

2. Free energy profiles

We wish to compare the free energy curves along the two pathways in vacuum and in water. To perform these calculations we must include a potential for bonded interactions. We have chosen the standard CHARMM potential that includes terms for bond stretching, angle bending, torsional potentials, and improper torsional potential.[24]

To compute the two free energy profiles we first have to locate the free energy minima in phi–psi space in the presence of solvent, that correspond to the α_R, β, and α_L conformers. We therefore perform an unconstrained MC simulation for 100 000 steps starting near the expected positions of these minima. At each step we construct a water shell on a cubic grid with 0.5 Å spacings and calculate the resulting interaction with the solute. We then bin up the phi–psi space into 10×10 deg bins and take the peaks of the probability distributions to be our free energy minima. We use these minima as the start and end points for the profiles in both vacuum and in solution.

To perform the calculation of the free energy profile in vacuum we impose holonomic internal coordinate con-

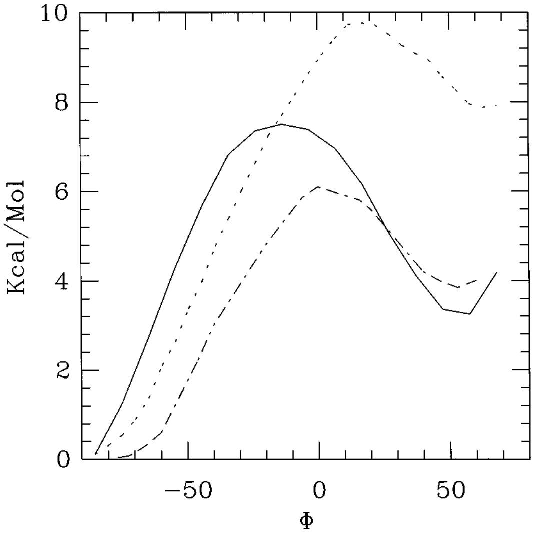

FIG. 5. The free energy profile connecting the β and α_L conformers of alanine dipeptide. In the β conformation $\phi=-85.0$ and $\psi=125$. In the α_L conformation $\phi=55.0$ and $\psi=65.0$ (all angles in degrees). The dashed line represents the vacuum calculation. The solid line represents the solvated calculation using the techniques of this work. The dash–dot line is the data from the work of Tobias and Brooks using MD to compute the free energy profile in the water phase.

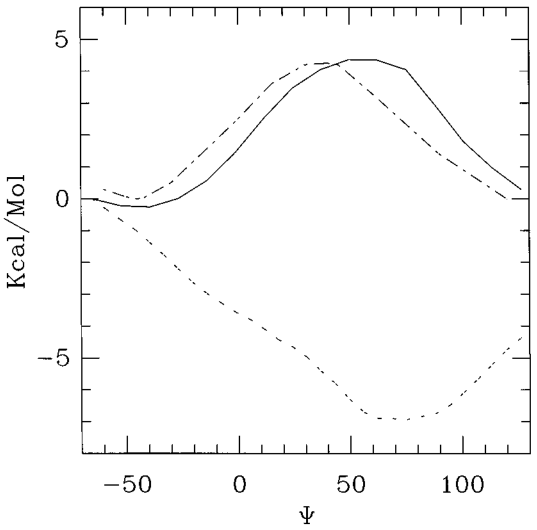

FIG. 6. The free energy profile connecting the α_R and β conformers of alanine dipeptide. In the α_R conformation $\phi=-95.0$ and $\psi=-65$. In the β conformation $\phi=-85.0$ and $\psi=125.0$ (all angles in degrees). The dashed line represents the vacuum calculation. The solid line represents the solvated calculation using the techniques of this work. The dash–dot line is the data from the work of Tobias and Brooks using MD to compute the free energy profile in the water phase.

straints on the dihedral angles and then use FEP theory. This is accomplished by placing a strong harmonic constraint on the dihedral angles at 10 deg intervals along the pathway. At each interval we allow the molecule to equilibrate for 10 000 Monte Carlo steps and then we compute $-\beta^{-1}\ln\langle\exp\{-\beta[U(\phi+\delta\phi,\psi+\delta\psi)-U(\phi,\psi)]\}\rangle_{\phi,\psi}$ for another 10 000 steps.

The solvation contribution to the free energy profiles along the two trajectories are calculated in the same manner as in the case of n-butane. We first generate a structure for the alanine dipeptide using the program QUANTA (Molecular Simulations Inc.). We rotate the dihedral angles of this structure through 15 points equally spaced along the trajectory. At each point we keep the bond lengths and angles fixed. We then proceed to build a shell of water on a cubic grid with 0.5 Å spacings and calculate $-\beta^{-1}\ln\exp\{-\beta[U(\phi+\delta\phi,\psi+\delta\psi)-U(\phi,\psi)]\}_{\phi,\psi}$ with respect to the fixed water shell.

We find that we do not need to allow the alanine dipeptide bond angles and lengths to relax during this calculation since we obtain good agreement with the MD free energy profiles by summing our vacuum profile to our solvation profile with the frozen dipeptide (see Figs. 5 and 6). We believe that this is a consequence of the empirical fact that these solvation profiles are relatively independent of minor (0.2 Å rms) deviations in the atomic positions that occur during a MC run with constrained dihedral angles. Therefore it is a good approximation to neglect the averaging over constrained solute thermal motion in the FEP formula altogether.

Given the nature of our approximations (using a limited set of explicit three-point CF's, frozen alanine dipeptide model for the FEP solvation calculations) and the different parametrization of the molecule (OPLS nonbonded interactions instead of CHARMM) it is remarkable that our free energy profiles are so similar to those computed by Brooks. The relative free energy changes between the three conformers are summarized in Table II. They agree with Brooks' MD calculations within 0.5 kcal/mol. The Barrier heights are also accurate to within 1.5 kcal/mol. Our results are more similar to those of Brooks than several other studies of the alanine dipeptide done by either MD or MC (see Table II). In particular they are much more effective at reproducing the MD results than RISM calculations that also attempt to construct PMF's out of analytically computed CF's.[25,26]

IV. CONCLUSION

We have demonstrated that a restricted set of CF's can be used to accurately calculate the water oxygen and hydrogen density around small molecules. Furthermore we have shown that this water shell can be used in conjunction with FEP theory to compute the solvation contribution to conformational free energy changes of the molecules. The potential of mean force between two solvated methanes calculated in this manner reproduces the MD results very closely even when only two-point CF's are used. Similarly three-point

TABLE II. Comparison of theoretical results for the relative free energies of alanine dipeptide conformations in water.[a]

Conformation	Anderson and Hermans (MD)	Pettitt and Karplus (RISM)	Tobias and Brooks (MD)	This work
α_R	0.0	0.0	0.0	0.0
β	−1.4		−0.2	0.3
α_L	1.1	−0.7	3.8	4.2

[a] Free energies (kcal/mol) with respect to the α_R conformation.

CF's allow us to correctly compute the effect of solvation on dihedral rotations of n-butane; our results agree, within 0.1 kcal/mol, with MD calculations of the relative populations of the *gauche* and *trans* conformations and the barrier between them. Finally, we have shown that even in a more complex system such as the alanine dipeptide, evaluating the explicit three-point CF's for only three dipole pairs in the solute, our resulting free energy profiles along paths joining free energy minima are in good agreement with those obtained by Brooks using MD.[20] In all cases this procedure has the distinct advantage over the standard MD techniques that it can obtain the free energy changes three or more orders of magnitude faster.

This methodology could potentially be extended to efficiently construct oxygen and hydrogen water densities around entire proteins in any conformation. Hummer *et al.* have already shown how a similar procedure reproduces the water density around DNA,[3–5] and Hummer and García have demonstrated that it also reproduces the water density at the interface of an antibody–antigen complex (personal communication). In all cases this requires the computation of many more two and three-point CF's than used in this work. If we estimate about one or two three-point CF's per different amino acid side chain this might entail 30 three-point CF's. However once the CF's are computed the water density can be built up in seconds.

It should also be noted that since the oxygen and hydrogen water densities computed on a grid do not interact with each other, this method could easily and effectively be implemented on a parallel machine.

At present a fully solvated small (50 residue) protein can only be simulated for 1 ns using MD, while much of the interesting behavior (e.g., protein folding) occurs on much longer time scales up to 1 s. Thus this more computationally efficient technique for computing solvation effects, when used in conjunction with novel MD algorithms,[27] that have been used to simulate polypeptide motions for microseconds, could prove extremely useful for extending the simulation time scales of these large biomolecules.

ACKNOWLEDGMENTS

We are indebted to A. García and G. Hummer, who first demonstrated the usefulness of using potential of mean force expansions for solvating biomolecules, for many helpful discussions. We acknowledge partial support of this research by the U.S. Department of Energy through the Stanford Synchrotron Radiation Laboratory.

[1] W. L. Jorgensen, Acc. Chem. Res. **22**, 184 (1989).
[2] G. Hummer and D. M. Soumpasis, Phys. Rev. E **49**, 591 (1994).
[3] G. Hummer and D. M. Soumpasis, *Structural Biology: The State of the Art. Proceedings of the Eight Conversation*, State University of New York, Albany, New York, 1993, edited by H. S. Ramaswamy and M. H. Sarma (Adenine, New York, 1984), p. 273.
[4] G. Hummer, A. García, and D. Soumpasis, Biophys. J. (in press).
[5] G. Hummer, D. Soumpasis, and A. García, in *Nonlinear Excitations in Biomolecules*, edited by M. Peyard (Springer, Berlin, 1995).
[6] A. Munster, *Statistical Thermodynamics* (Springer, Berlin, 1969), Vol. 1, p. 338.
[7] J. G. Kirkwood, J. Chem. Phys. **3**, 300 (1935).
[8] H. J. C. Berendsen, J. P. M. Postma, W. F. Von Gunsteren, and J. Hermans, *Intermolecular Forces*, edited by B. Pullman (Reidel, Dordrecht, 1981), p. 331.
[9] W. L. Jorgensen and J. Tirado-Rives, J. Am. Chem. Soc. **110**, 1657 (1988).
[10] W. L. Jorgensen, J. D. Madura, and C. J. Swenson, J. Am. Chem. Soc. **106**, 6638 (1984).
[11] M. Levitt, M. Hirshberg, R. Sharon, and V. Daggett (unpublished).
[12] A. K. Soper and M. G. Phillips, Chem. Phys. **107**, 47 (1987).
[13] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[14] N. Metropolis, M. N. Rosenbluth, A. Rosenbluth, H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[15] R. W. Zwanzig, J. Chem. Phys. **22**, 1420 (1954).
[16] D. van Belle and S. J. Wodak, J. Am. Chem. Soc. **115**, 647 (1993).
[17] H. S. Frank and W. J. Evans, Chem. Phys. **13**, 507 (1945)
[18] D. J. Tobias and C. L. Brooks, J. Chem. Phys. **92**, 2582 (1990).
[19] D. Beglov and B. Roux, J. Chem. Phys. **100**, 9050 (1994).
[20] D. J. Tobias and C. L. Brooks, J. Phys. Chem. **96**, 3864 (1992).
[21] C. A. Schiffer, J. W. Caldwell, R. M. Stroud, and P. A. Kollman, Prot. Sci. **1**, 396 (1992).
[22] A. G. Anderson and J. Hermans, Proteins **3**, 262 (1988).
[23] G. Nemethy, Z. Hodes, and H. A. Scheraga, Proc. Natl. Acad. Sci. U. S. A. **75**, 5760 (1978).
[24] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, J. S. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **2**, 187 (1983).
[25] B. M. Pettitt and M. Karplus, Chem. Phys. Lett. **121**, 194 (1985).
[26] P. E. Smith, B. M. Pettitt, and M. Karplus, J. Phys. Chem. **97**, 6907 (1993).
[27] N. Grønbech-Jensen and S. Doniach, J. Comput. Chem. **15**, 997 (1994).