Protein Science (1992), *1*, 396–400. Cambridge University Press. Printed in the USA.

Inclusion of solvation free energy with molecular mechanics energy: Alanyl dipeptide as a test case

CELIA A. SCHIFFER,[1,2] JAMES W. CALDWELL,[3] ROBERT M. STROUD,[1,2] AND PETER A. KOLLMAN[2,3]

[1] Department of Biochemistry and Biophysics, [2] Graduate Group in Biophysics, and [3] Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California 94143

(RECEIVED August 29, 1991; ACCEPTED October 31, 1991)

Abstract

A combined force field of molecular mechanics and solvation free energy is tested by carrying out energy minimization and molecular dynamics on several conformations of the alanyl dipeptide. Our results are qualitatively consistent with previous experimental and computational studies, in that the addition of solvation energy stabilizes the C_5 conformation of the alanyl dipeptide relative to the C_7.

Keywords: alanyl dipeptide; energy minimization; molecular dynamics; molecular mechanics; solvation free energy

Protein structure is affected by solvent, from the actual folding of the protein to the conformation of external side chains. Therefore the inclusion of solvent when calculating physical properties derived from protein structure is essential. Modeling explicit solvent in calculations, however, is computationally expensive and can be impractical in deriving certain physical properties. Thus, a method of representing the effects of water without having water molecules explicitly in the calculation is potentially very useful.

One such method assigns an atomic solvation parameter to each atom and then multiplies this by the solvent-exposed surface area to calculate a free energy of solvation ($\Delta G = \Delta\sigma * A$) (Eisenberg & McLachlan, 1986; Eisenberg et al., 1989). This solvation free energy (SFE) has been used to distinguish between correctly and incorrectly folded static protein structures (Novotny et al., 1988; Chiche et al., 1989, 1990) as well as side-chain conformations (Schiffer et al., 1990).

In carrying out energy minimization and molecular dynamics (MD) on a molecule it is important that the derivative of the SFE is used as a component of the force to describe a trajectory of the molecule relevant to its solvated state. Such a derivative of the SFE added into a molecular mechanics force field should be able to mimic more correctly the forces in a solvated molecule.

In order to test such a combined force field we have looked at several conformations of alanyl dipeptide (C_{7eq}, C_{7ax}, C_5, α_R). The structure of the alanyl dipeptide has been determined experimentally under several solvent conditions by a variety of techniques. When crystallized in methanol it assumes a twisted β-strand conformation (Harada & Iitaka, 1974). In a nonpolar solvent, the C_{7eq} conformation was confirmed to be the most stable by Raman spectroscopy. This conformation was postulated to retain its stability in water by coordination of a water molecule between its carboxyl oxygen and amino hydrogen (Avignon et al., 1973). More recently, however, by CD and NMR, Madison and Kopple (1980) have shown that although the C_{7eq} conformation dominates the population in nonpolar solvents, as the solvents become more polar the population of peptides in the C_{7eq} conformation decreases and a variety of other conformations exist. In that paper they find ". . . the experimental data suggest to us that the C_7 population of AcAlaNHMe in chloroform goes to α_r and P_{II} forms in water, and that any C_5 population present remains approximately constant" (Madison & Kopple, 1980).

Computationally, the structure of the alanyl dipeptide has also been extensively studied. Early studies looked

Reprint requests to: Peter A. Kollman, Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California 94143-0446.

at the structure of the water around only a few conformations of the alanyl dipeptide by Monte Carlo (Hagler et al., 1980) and MD (Rossky & Karplus, 1979). More recent studies using intramolecular potential of mean force show that only the C_{7eq} and C_{7ax} conformations are accessible under vacuum conditions, whereas under aqueous solvent conditions many conformations are significantly sampled with the C_5 conformation being the lowest energy and most populated conformation (Pettitt & Karplus, 1985; Lau & Pettitt, 1987). In another recent study, the free energy differences between four conformations of the alanyl dipeptide were calculated with MD and explicit water molecules (Anderson & Hermans, 1988). In this case the C_{7eq} and C_5 conformations were clumped together into one lowest-energy β region and the free energy differences were calculated between it and the α_R, α_L, and C_{7ax} conformations.

Results

Initially, we calculated the relative energies of the four conformations using only the Weiner et al. (1986) potential with a dielectric constant of 1.0 (Table 1A) for the alanyl dipeptide. Recent quantum mechanical calculations have shown, however, that even under vacuum conditions the C_5 conformation is energetically closer in stability to the C_{7eq} than this potential predicts (Head-Gordon et al., 1991). Therefore, in order to make the potential more accurate, additional dihedral energy terms were added to stabilize the C_5 conformation (Table 1B). These added terms led to an overall increase in stabilization of 2.79 kcal of the C_5 conformation relative to the C_{7eq} conformation.

The SFE was then added into the standard molecular mechanics model with added dihedral terms, and the relative energies were once again calculated (Table 2). The C_5 conformation has the lowest solvation energy (−1.62 kcal) and the lowest relative energy (−0.26 kcal). The α_R conformation has the second lowest solvation energy (−1.31 kcal) and its relative energy decreased compared to the C_{7eq} by 0.97 kcal with the addition of solvation. Thus, the SFE helped stabilize both the C_5 and the α_R conformations relative to the C_{7eq}.

The SFE is able to predict the trends of whether or not a certain conformation is more likely to be solvated. The SFE does not, however, act like explicit water, in that explicit water will screen a charge-charge Coulomb interaction and the SFE will not. Coulombic interactions can be very strong, especially when a dielectric constant of 1.0 is used. Therefore other dielectric constants can also be used to mimic the effect of the "screening" by water. Using a distance-dependent dielectric helped to stabilize the α_R conformation as well as the C_5 conformation (Table 3A,B).

As another test of how the addition of the SFE changed the molecular mechanics force field, four MD calculations were run. These runs were performed to see if the potential energy surface would flatten with the addition of the SFE. Each run used the extra dihedral potential term, and coordinates were saved every two picoseconds and then minimized to the nearest local minima. In the first two of the calculations the dielectric constant was set to one. Four 100-ps MD runs were performed

Table 1. *Minimized structures with the Weiner et al. (1986) potential*[a]

Structure	Φ	Ψ	Total E	ΔE
A. $\epsilon = 1$				
C_{7eq}	−76.5	67.4	−41.95	0.00
C_{7ax}	68.0	−65.2	−40.69	1.26
C_5	−160.2	168.9	−37.16	4.79
α_R	−58.9	−39.9	−37.27	4.68
B. Another Φ-Ψ term added; $\epsilon = 1$; $V = 1.5$ kcal; $n = 1$; $\delta = 0°$ for the Φ-Ψ dihedrals				
C_{7eq}	−77.2	72.8	−39.21	0.00
C_{7ax}	68.9	−70.1	−37.72	1.49
C_5	−160.1	170.6	−37.11	2.10
α_R	−61.9	−36.1	−33.57	5.64

[a] Angles are in degrees; energy in kcal/mol.

Table 2. *Modified Weiner et al. (1986) potential with solvation energy*[a]

Structure	Φ	Ψ	Total E	Solvation E	ΔE
C_{7eq}	−76.4	78.3	−39.47	−0.37	0.00
C_{7ax}	68.3	−72.5	−38.32	−0.65	1.15
C_5	−157.2	161.9	−39.73	−1.62	−0.26
α_R	−63.8	−38.3	−34.80	−1.30	4.67

[a] $\epsilon = 1$; $V = 1.5$ kcal; $n = 1$; $\delta = 0°$ for the Φ-Ψ dihedrals. Angles are in degrees; energy in kcal/mol.

Table 3. *Minimized structures with modified Weiner et al. (1986) potential and a distance-dependent dielectric constant*[a]

Structure	Φ	Ψ	Total E	Solvation E	ΔE
A. $\epsilon = r$; $V = 1.5$ kcal; $n = 1$; $\delta = 0°$ for the Φ-Ψ dihedrals					
C_{7eq}	−77.3	70.0	−10.49		0.00
C_{7ax}	69.8	−67.1	−9.38		1.11
C_5	−163.3	171.0	−10.74		−0.25
α_R	−65.4	−44.2	−6.06		4.43
B. Same as A but with solvation energy					
C_{7eq}	−76.8	73.7	−10.64	−0.24	0.00
C_{7ax}	69.6	−67.9	−9.95	−0.61	0.69
C_5	−161.0	163.8	−12.25	−1.60	−1.61
α_R	−68.1	−42.0	−7.39	−1.38	3.25

[a] Angles are in degrees; energy in kcal/mol.

starting with the alanyl dipeptide at each of the four minima (C_5, α_r, C_{7ax}, C_{7eq}). The first set of calculations was with the in vacuo molecular mechanics force field. In the three 100-ps MD runs starting from the C_5, α_r, and C_{7eq} conformations of the alanyl dipeptide, the molecule settled into the C_{7eq} conformation within 20 ps. The MD run that started the conformation of the alanyl dipeptide in C_{7ax} remained in that local minima for the entire 100 ps. A plot of the conformations sampled over these four runs is shown in Figure 1A. In the second calculation the SFE was added to the force field. Once again the three 100-ps MD runs starting from the C_5, α_r, and C_{7eq} conformations of the alanyl dipeptide of the molecule settled into the same local minima. This time, however, it was the C_5 conformation. The MD run that started the conformation of the alanyl dipeptide in C_{7ax} once again remained in that local minima for the entire 100 ps. Figure 1B shows that several local minima with intermediary conformations exist during the course of the run. The molecule remained over 50% of the time in the C_5 conformation,

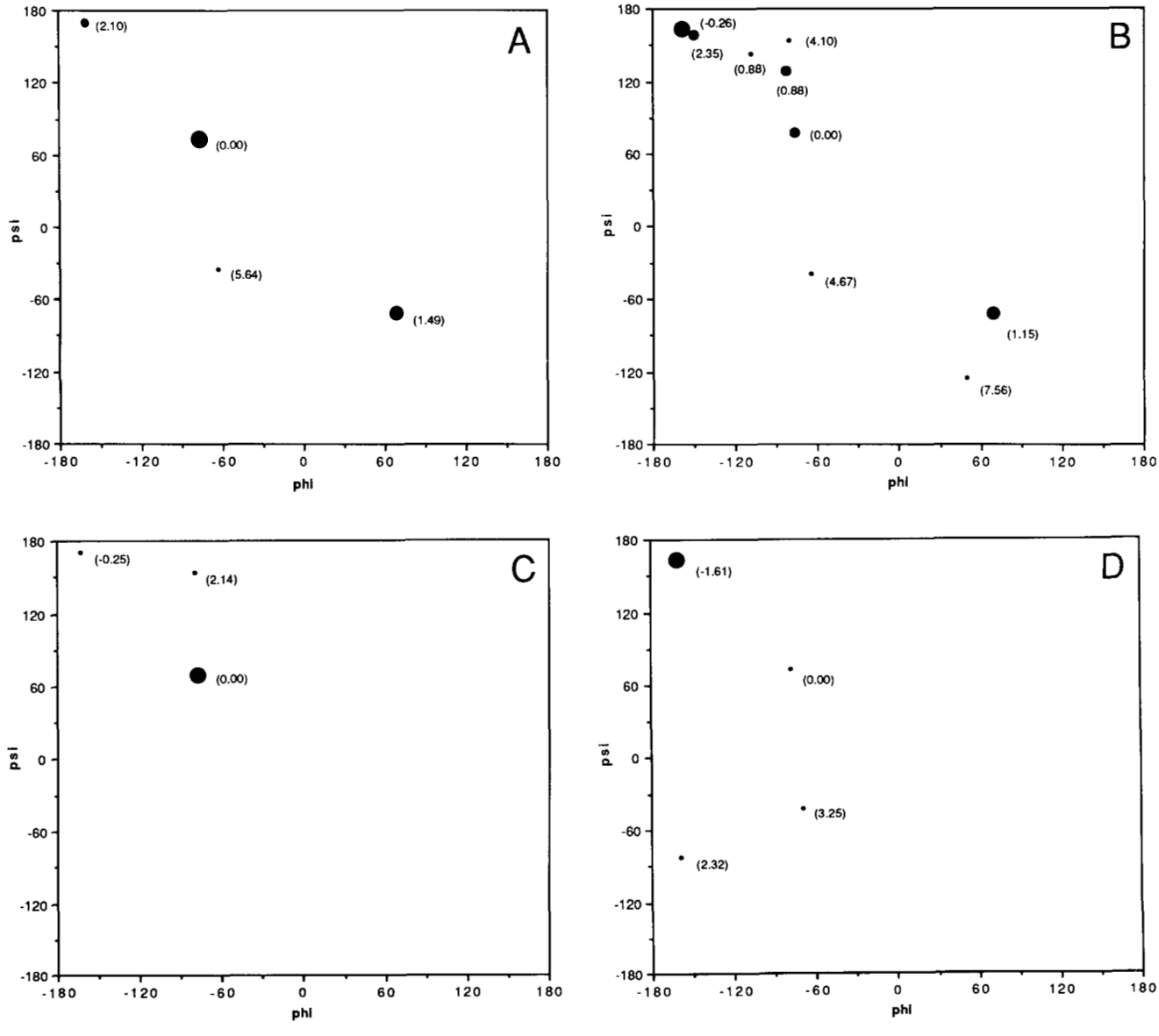

Fig. 1. Ramachandran plots of MD calculations on alanyl dipeptide. Structures were saved every 2 ps, and energy was minimized to the nearest local minima. The energies, relative to the C_{7eq}, of the minima are listed in kcal. The sizes of the spots are approximately proportional to the log of the number of structures that fall in that particular minima. **A:** Molecular mechanics with a dielectric of one: four 100-ps runs with initial conformations C_5, C_{7eq}, C_{7ax}, α_r. **B:** Molecular mechanics and SFE with a dielectric of one: four 100-ps runs with initial conformations C_5, C_{7eq}, C_{7ax}, α_r. **C:** Molecular mechanics with a distance-dependent dielectric: a 100-ps run with an initial conformation of α_r. **D:** Molecular mechanics and SFE with a distance-dependent dielectric: a 100-ps run with an initial conformation of α_r.

which had the lowest energy and occasionally flipped to C_{7eq} and to a second higher energy C_5 conformation where the carboxyl methyl and oxygen switched positions.

The third and fourth calculations were of 100 ps each, starting the alanyl dipeptide in the α_R conformation. A distance-dependent dielectric constant was used. The third calculation was with only the standard molecular mechanics force field. By the second picosecond, the conformation of the peptide was in the C_5 conformation. In the fourth picosecond, it was in an intermediary position between C_5 and C_{7eq}. For the remainder of the calculation, the peptide remained in the C_{7eq} position, even though this position was 0.25 kcal/mol higher in energy than the C_5 conformation (Fig. 1C). During the final calculation with both a distance-dependent dielectric constant and SFE, the peptide moved from the α_R to the C_5 conformation by the 10th picosecond. C_5 was the lowest energy conformation and the molecule spent the remaining 90 ps in the conformation (Fig. 1D).

The goal here is not to redetermine the Φ–Ψ map; the basic features of such a map with force field calculations have been presented by Weiner et al. (1986) and Pettitt and Karplus (1985), among others. It is rather to evaluate the relative solvation energies of key conformations of the alanyl dipeptide. By running 100-ps MD simulations starting with the four conformations presented in Table 1, any new low-energy conformations should appear if they exist. These results suggest that there are none and that the key effect of the SFE calculation is to stabilize C_5 and α_r relative to C_{7eq} and C_{7ax}.

Discussion

In this study of the alanyl dipeptide we have shown that the addition of SFE to a molecular mechanics force field can mimic some aspects of the effect of explicit solvent. This seems to be especially true when a distance-dependent dielectric constant is used to help further screen the charge–charge interactions. We have shown that the addition of the force from the SFE helps to sample conformational space during an MD run. This sampling is more in the manner of how an alanyl dipeptide would sample conformational space in an experimental aqueous system in that the relative stability of C_{7eq} is decreased. The dominance of the C_5 conformation rather than the increase in the α_r or P_{II} populations seen by Madison and Kopple (1980) is likely an inherent defect in SFE models such as the one presented here. In particular, reaction field effects due to dipolar alignment of solute atoms are not represented in this model; such effects would be expected to stabilize α_r over C_5 conformations.

The results presented here are also consistent with the relative energies of the conformations of the alanyl dipeptide as calculated by potential mean force (Pettitt & Karplus, 1985; Lau & Pettitt, 1987), in that in the vacuum state there exists only two very deep minima at C_{7ax} and C_{7eq} and the addition of solvation makes the potential surface have many more accessible minima. Recently, an approximate numerical method for minimizing the SFE has been presented (Hasel et al., 1988; Still et al., 1990). That method, however incorrectly, estimates the solvent-accessible surface area by over 4 Å^2 per atom. It is as yet unclear how these inaccuracies in the calculations will affect the low-energy conformation of a peptide. On the other hand, that method, or others that estimate electrostatic SFEs by simple continuum or Born-type models (Gilson & Honig, 1990), can provide more stabilization for conformations such as α_R of the alanyl dipeptide, compared to approaches such as ours where solvation is only a function of solvent exposure. This method has recently been applied with MD to the alanyl dipeptide (Sharp, 1991), and the addition of solvation to the system also destabilizes the C_{7eq} relative to C_5 and α_r. However, the C_{7eq} still remains the lowest energy conformation by several kilocalories using this Born-type model. A better solvation model may ultimately be developed that combines both the solvent-exposed surface area considered here and the electrostatic effects analyzed by Sharp.

The addition of SFE in the MD calculation increases the length of time of the calculation over plain MD, but it is still faster than using a periodic box of explicit solvent, although not faster than a small shell of explicit solvent (Guenot & Kollman, 1991, unpubl.). The speed of the calculation is machine dependent, as the determination of a solvation energy is the most scalar component of the calculation. On a VAX 8650 100 ps of MD on the alanyl dipeptide took 31 min of central processing unit (CPU) time, and the addition of the SFE increased the time of the calculation to almost 4 h. On an IBM RISC 6000/530 the same calculation (100 ps of MD with SFE) took 45 min of CPU time. Further study is necessary to determine exactly which atomic solvation parameters are the most relevant for use in protein systems. In any case, using approximate methods for representing the effects of solvent should continue to be useful in computing energies of large macromolecule systems. This is particularly relevant in situations when it is impractical to use explicit solvent in the calculations, such as when making a homology model where extensive conformational searching is required (Schiffer et al., 1990). This code is available upon request with the rest of the AMBER software (Singh et al., 1986).

Materials and methods

The molecular mechanics force field we used was the standard AMBER force field with the exception of the addition of a dihedral potential term that stabilizes the C_5 conformation relative to C_{7eq}. $E_{\text{dihedral}} = \frac{1}{2}V[1 + \cos(n\theta - \delta]$ and $V = 1.5$ kcal, $n = 1$ and $\delta = 0°$ for θ equal Φ–Ψ dihedral angles. No scaling of the 1–4 nonbonded interactions was used. This was done to reduce the in vacuo

difference of C_{7eq} and C_5 to approximately 2 kcal/mol to make this consistent with recent quantum mechanical calculations (Head-Gordon et al., 1991).

In order to add the SFE we needed to incorporate the derivatives of the surface area into AMBER. These derivatives were taken from the work of Richmond (1984) with some adjustments (Wesson & Eisenberg, 1992). We checked the analytical derivatives with those determined numerically and found them to be consistent. We also did not encounter any difficulties with respect to surface area discontinuities during any minimization or dynamics runs.

The SFE and its derivatives were added directly into the derivatives calculated for the force in AMBER. The SFE is then a driving force along with all the other terms in the molecular mechanics potential. We used a set of atomic solvation parameters calculated from the free energy of transfer between water and vapor of small organic molecules (Wesson & Eisenberg, 1992).

For all the energy minimization calculations the dipeptide was started near one of four known low-energy conformations. Each conformation was minimized to a gradient of 0.02 kcal * mol/Å or 10,000 cycles, whichever came first. For the molecular dynamics calculation, 0.5-fs time steps and a temperature of 300 °K were used for a total of 100 ps starting from the α_R conformation. Coordinates were saved every 2 ps, and then energy was minimized as described above.

Acknowledgments

This research was supported by NIH grants GM-29072 (P.A.K.) and GM-24485 (R.M.S.).

References

Anderson, A.G. & Hermans, J. (1988). Microfolding: Conformational probability map for the alanine dipeptide in water from molecular dynamic simulations. *Proteins Struct. Funct. Genet. 3*, 262–265.

Avignon, M., Garrigou-Lagrange, C., & Bothorel, P. (1973). Conformational analysis of dipeptides in aqueous solution. II. Molecular structure of glycine and alanine dipeptides by depolarized rayleigh scattering and laser raman spectroscopy. *Biopolymers 12*, 1651–1669.

Chiche, L., Gaboriaud, C., Heitz, A., Mornon, J.P., Castro, B., & Kollman, P.A. (1989). Use of restrained molecular dynamics in water to determine three-dimensional protein structure: Prediction of the three-dimensional structure of *Ecballium elaterium* trypsin inhibitor II. *Proteins Struct. Funct. Genet. 6*(4), 405–417.

Chiche, L., Gregoret, L.M., Cohen, F.E., & Kollman, P.A. (1990). Protein model structure evaluation using the solvation free energy of folding. *Proc. Natl. Acad. Sci. USA 87*(8), 3240–3243.

Eisenberg, D. & McLachlan, A.D. (1986). Solvation energy in protein folding and binding. *Nature 319*, 199–203.

Eisenberg, D., Wesson, M., & Yamashita, M. (1989). Interpretation of protein folding and binding with atomic solvation parameters. *Chem. Scripta 29A*, 217–221.

Gilson, M.K. & Honig, B. (1990). Calculations of the total electrostatic energy of a macromolecular system: Solvation energies, binding energies, and conformational analysis. *Proteins Struct. Funct. Genet. 4*(1), 7–18.

Hagler, A.T., Osguthorpe, D.J., & Robson, B. (1980). Monte Carlo simulation of water behavior around the dipeptide *N*-acetylalanyl-*N*-methylamide. *Science 208*, 599–601.

Harada, Y. & Iitaka, Y. (1974). The crystal and molecular structures of *N*-acetyl-DL-alanine-*N*-methylamide and *N*-acetyl-L-alanine-*N*-methylamide. *Acta Crystallogr. B30*, 1452–1459.

Hasel, W., Hendrickson, T.F., & Still, W.C. (1988). A rapid approximation to the solvent accessible surface areas of atoms. *Tetrahedron Comput. Methodol. 1*(2), 103–116.

Head-Gordon, T., Head-Gordon, M., Frisch, M.J., Brooks, C.L., III, & Pople, J.A. (1991). Theoretical study of blocked glycine and alanine peptide analogues. *J. Am. Chem. Soc. 113*(16), 5989–5997.

Lau, W.F. & Pettitt, B.M. (1987). Conformations of the glycine dipeptide. *Biopolymers 26*, 1817–1831.

Madison, V. & Kopple, K.D. (1980). Solvent-dependent conformational distributions of some peptides. *J. Am. Chem. Soc. 102*(15), 4855–4863.

Novotny, J., Rashin, A.A., & Bruccoleri, R.E. (1988). Criteria that discriminate between native proteins and incorrectly folded models. *Proteins Struct. Funct. Genet. 4*(1), 19–30.

Pettitt, B.M. & Karplus, M. (1985). The potential of mean force surface for the alanine dipeptide in aqueous solution: A theoretical approach. *Chem. Phys. Lett. 121*(3), 194–201.

Richmond, T.J. (1984). Solvent accessible surface area and excluded volume in proteins: Analytical equations for overlapping spheres and implications for the hydrophobic effect. *J. Mol. Biol. 178*, 63–89.

Rossky, P.J. & Karplus, M. (1979). Solvation: A molecular dynamics study of a dipeptide in water. *J. Am. Chem. Soc. 101*(8), 1913–1937.

Schiffer, C.A., Caldwell, J.W., Kollman, P.A., & Stroud, R.M. (1990). Prediction of homologous protein structures based on conformational searches and energetics. *Proteins Struct. Funct. Genet. 8*(1), 30–43.

Sharp, K. (1991). Incorporating solvent and ion screening into molecular dynamics using the finite-difference Poisson-Bolzmann method. *J. Comp. Chem. 12*(4), 454–468.

Singh, U.C., Weiner, P., Caldwell, J., & Kollman, P. (1986). AMBER 3.0. University of California, San Francisco.

Still, W.C., Tempczyk, A., Hawley, R.C., & Hendrickson, T. (1990). Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc. 112*(16), 6127–6129.

Weiner, S.J., Kollman, P.A., Nguyen, D.T., & Case, D.A. (1986). An all atom force field for simulations of proteins and nucleic acids. *J. Comp. Chem. 7*(2), 230–252.

Wesson, L. & Eisenberg, D. (1992). Atomic solvation parameters applied to molecular dynamics of proteins in solution. *Protein Sci. 1*, 227–235.