

Atomic solvation parameters applied to molecular dynamics of proteins in solution



LAURA WESSON AND DAVID EISENBERG

Molecular Biology Institute and Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90024-1570

(RECEIVED August 27, 1991; ACCEPTED October 10, 1991)

Abstract

A solvation energy function for use in the molecular simulation of proteins is proposed. It is based on the accessible surface areas of atoms in the protein and on atomic solvation parameters derived from empirical vapor-to-water free energies of transfer of amino acid side-chain analogs. The energy function and its derivatives were added to the CHARMM molecular simulation program (Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., & Karplus, M., 1983, *J. Comput. Chem.* 4(2), 187–217). The effect of the added energy term was evaluated by 110 ps of molecular dynamics on the 26-residue protein melittin. The melittin monomer and tetramer were studied both with and without the added term. With the added energy term the monomer partially unfolded, while the secondary structure of the tetramer was preserved, in agreement with reported experiments (Brown, L.R., Lauterwein, J., & Wuethrich, K., 1980, *Biochim. Biophys. Acta* 622(2), 231–244; Lauterwein, J., Brown, L.R., & Wuethrich, K., 1980, *Biochim. Biophys. Acta* 622(2), 219–230).

Keywords: atomic solvation parameters; melittin; molecular dynamics

A common approach to protein energetics uses theoretically based potential functions with terms such as electrostatic energy, van der Waals energy, and bond and angle distortion energies. Some molecular simulation programs that use this approach are CHARMM (Brooks et al., 1983), GROMOS (Van Gunsteren & Berendsen, 1987), AMBER (Weiner & Kollman, 1981), ECEPP/2 (Momany et al., 1975; Nemethy et al., 1983; Sippl et al., 1984), and the OPLS modification of AMBER (Jorgensen & Tirado-Rives, 1988). However, the explicit simulation of water presents additional challenges, in part because the structure of water is less well understood than the structure of proteins and in part because water interacts with proteins through hydrophobic forces (Kauzmann, 1959) as well as electrostatic and van der Waals forces. The hydrophobic interaction, which results from the ordering of water molecules around apolar side chains of a protein, is more difficult than the other forces to represent effectively with a simple atom–atom energy potential.

We introduce here a solvent interaction potential function that describes protein–water interactions based solely

on the positions of protein atoms, as an alternative to including explicit water molecules in a simulation. The potential function uses atomic solvation parameters (Eisenberg & McLachlan, 1986; Eisenberg et al., 1989), derived from free energies of transfer, which express the hydrophobicity of each protein atom type i . It describes protein–water interaction energies based on the atomic solvation parameters of protein atoms, $\Delta\sigma_i$, and their solvent-accessible areas, A_i . To combine atomic solvation parameters with traditional molecular dynamics, we have derived new atomic solvation parameters to express the energy required to transfer a protein molecule from vapor to water. When the resulting potential function is added to a simulation of a protein in a vacuum, the total potential describes a protein solvated in water. Our solvation potential was also added to the AMBER force field and tested on an alanyl dipeptide (Schiffer et al., 1992).

In our approach we follow Lee and Richards (Lee & Richards, 1971; Richards, 1977) in defining the solvent-accessible area of a protein atom as the area swept out by the center of a spherical solvent probe tangent to the atom, as it rolls over the surface of the protein. Richmond (1984) derived analytical expressions for the solvent-exposed areas and their positional derivatives and

Reprint requests to: David Eisenberg, Molecular Biology Institute and Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90024-1570.

wrote the program ANAREA to calculate them. In the Results and discussion section we describe our modification of ANAREA to calculate solvent-protein interaction energy and forces, which we have incorporated into CHARMM as a subroutine.

A practical limitation of the Richmond method for calculating exact solvent accessible areas is that it is computationally expensive, using roughly one central processing unit (CPU) minute on a VAX 4000 to calculate accessible areas and derivatives for the atoms in a 100-residue protein. At least two suggestions have been made for rapid analytical approximations to solvent-accessible areas. Wodak and Janin (1980) approximated accessible surface areas analytically on a per-residue basis, with a statistical model that treated residues as spheres. W. Clark Still (Hasel et al., 1988) adapted the formulation of Wodak and Janin to calculate surface areas on an atomic basis, an approximation that is accurate to within about 20% and is faster than the exact calculation.

In a similar vein to the present investigation, Vila et al. (1991) studied several solvation models based on atomic solvation parameters (Eisenberg & McLachlan, 1986; Ooi et al., 1987). The areas of the protein atoms were calculated using the Connolly molecular surface, which is defined as the closest approach of the surface of a spherical probe to the van der Waals surface of the protein (Connolly, 1983). Vila et al. added a solvation potential function based on the Connolly surface areas and derivatives to the ECEPP/2 program and tested several solvation models using the modified ECEPP/2 potential. The solvation models were evaluated by the concordance between solvation free energy and root mean square (RMS) deviation from the crystal structure in 39 near-native conformations of bovine pancreatic trypsin inhibitor.

Results and discussion

Fundamental relationships

The thermodynamic cycle shown in Figure 1 demonstrates how an energy function that computes the vacuum-to-water free energy of transfer for a protein can be used to convert a simulation in vacuum to one in water. The protein in the figure has two conformational states, with the domain hinge either closed or open. A molecular dynamics program such as CHARMM can in principle compute the free energy difference between the final and initial states in vacuum, $\Delta G_{\text{vacuum}}^{f-i}$. Our energy function can be used to find the vacuum-to-water free energies of transfer of the protein in the initial and final states, ΔG_s^i and ΔG_s^f . Then the energy difference between the two states in water is

$$\Delta G_{\text{water}}^{f-i} = \Delta G_{\text{vacuum}}^{f-i} + \Delta G_s^f - \Delta G_s^i. \quad (1)$$

The energy required to transfer a protein from vacuum to water without conformational change can be approximated by (Eisenberg & McLachlan, 1986)

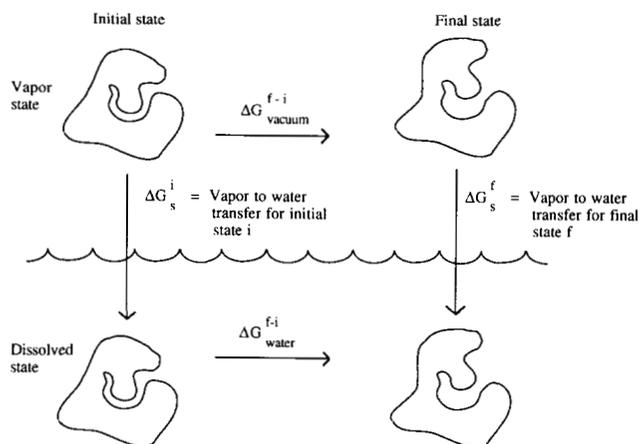


Fig. 1. Thermodynamic cycle that shows how vapor-to-water free energies of transfer may be used to compute protein free energy changes in solution. $\Delta G_{\text{vacuum}}^{f-i}$ is the free energy difference between two conformational states in vacuum, available in principle by molecular simulation. ΔG_s^i and ΔG_s^f are the vapor-to-water free energies of transfer for the protein in the initial and final conformational states. The free energy difference between the two conformational states in water, $\Delta G_{\text{water}}^{f-i}$, may be calculated from the thermodynamic cycle.

$$\Delta G_s = \sum_{\text{atoms } i} \Delta \sigma_i A_i, \quad (2)$$

where $\Delta \sigma_i$ is the atomic solvation parameter for atom i , and A_i is the Lee and Richards solvent-accessible surface area for the atom. In our area calculations, we have used a probe radius of 1.4 Å to simulate a water molecule. The atomic solvation parameter $\Delta \sigma_i$, in units of $\text{cal mol}^{-1} \text{Å}^{-2}$, is an estimate of the free energy required to transfer the atom from vacuum to water divided by exposed surface area. The atomic solvation parameter depends on the atom type. We classify atoms into five types: C, uncharged O or N, S, O⁻, and N⁺.

Derivation of hydrophobic interaction forces

We derived forces from the vacuum-to-water free energy of transfer ΔG_s , for addition to the standard CHARMM potentials for energy minimization and molecular dynamics. This was done as follows. The derivative of ΔG_s with respect to the position of atom i is

$$\frac{\delta(\Delta G_s)}{\delta \mathbf{x}_i} = \sum_{\text{atoms } j} \Delta \sigma_j \frac{\delta A_j}{\delta \mathbf{x}_i}. \quad (3)$$

This expands to

$$\frac{\delta(\Delta G_s)}{\delta \mathbf{x}_i} = \Delta \sigma_i \frac{\delta A_i}{\delta \mathbf{x}_i} + \sum_{\text{atoms } j \neq i} \Delta \sigma_j \frac{\delta A_j}{\delta \mathbf{x}_i}. \quad (4)$$

Moving atom i in the \mathbf{x}_i direction is equivalent to moving all the other atoms in the opposite direction, so that

$$\frac{\delta A_i}{\delta \mathbf{x}_i} = - \sum_{\text{atoms } j \neq i} \frac{\delta A_j}{\delta \mathbf{x}_j}. \quad (5)$$

This means that the force F_i on atom i can be expressed as

$$F_i = -\frac{\delta(\Delta G_s)}{\delta \mathbf{x}_i} = \sum_{\text{atoms } j \neq i} \left(\Delta \sigma_i \frac{\delta A_i}{\delta \mathbf{x}_j} - \Delta \sigma_j \frac{\delta A_j}{\delta \mathbf{x}_i} \right). \quad (6)$$

Equation 6 expresses solvation forces in terms of derivatives of accessible areas with respect to atomic positions. Richmond (1984) provided expressions for the area derivatives. We can adopt his expressions if we replace Equation 46 of his paper with

$$\frac{\delta t_{\lambda}^+}{\delta b_k} = \pm \frac{c_k}{a_k \sqrt{a_k^2 + c_k^2}}. \quad (7)$$

Similarly, we replace his Equation 43 with

$$\frac{\delta t_{\lambda+1}^-}{\delta b_k} = \pm \frac{1}{a_k}. \quad (8)$$

Here, as defined by Richmond (1984), t_{λ} and $t_{\lambda+1}$ are angles that parameterize the circular arcs C_{λ} and $C_{\lambda+1}$, which are on the boundary of a region of exposed surface area on sphere i (Fig. 2). The variable t_{λ}^+ is the maximum value of t_{λ} on the arc C_{λ} , at the intersection of C_{λ} and $C_{\lambda+1}$. Similarly, $t_{\lambda+1}^-$ is the minimum value of $t_{\lambda+1}$ on the arc $C_{\lambda+1}$. The vector (a_k, b_k, c_k) is from the center of sphere i to the center of sphere k .

Richmond's program ANAREA was modified by inserting Equations 7 and 8 to calculate the area derivatives and Equations 2 and 6 to calculate the solvent interaction

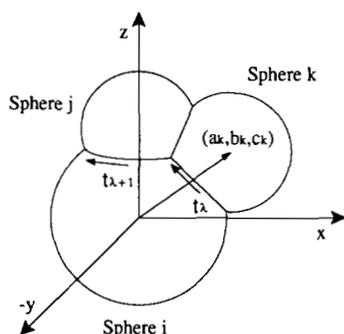


Fig. 2. Illustration of two components of the derivative of the accessible surface area, as defined by Richmond (1984), and as given by Equations 7 and 8. The coordinate frame has been chosen so that the center of sphere i is at the origin, the vector between the centers of sphere i and sphere j coincides with the z axis, and the vector (a_k, b_k, c_k) between the i th and k th sphere centers lies in the xz plane.

The accessible surface area of sphere i is bounded by the circular arcs C_{λ} and $C_{\lambda+1}$, formed by the intersection of sphere i with spheres k and j . The arcs C_{λ} and $C_{\lambda+1}$ are parameterized by the angles t_{λ} and $t_{\lambda+1}$. The variables t_{λ}^+ and $t_{\lambda+1}^-$ are the values of t_{λ} and $t_{\lambda+1}$ at the intersection of the arcs C_{λ} and $C_{\lambda+1}$. We derived expressions for the derivatives $\delta t_{\lambda}^+ / \delta b_k$ and $\delta t_{\lambda+1}^- / \delta b_k$, and incorporated them into the program ANAREA (Richmond, 1984), to calculate accessible surface area derivatives.

energy and forces. The program was then included as a subroutine in CHARMM. The ambiguous \pm signs in Equations 7 and 8 were resolved empirically by testing the solvent interaction forces by the finite differences method. In other words, if atom i is moved a small distance in a random direction $\Delta \mathbf{x}_i$, then

$$\Delta G_s(\mathbf{x}_i + \Delta \mathbf{x}_i) \cong \Delta G_s(\mathbf{x}_i) + \Delta \mathbf{x}_i \cdot \frac{\delta(\Delta G_s(\mathbf{x}_i))}{\delta \mathbf{x}_i} \quad (9)$$

where \mathbf{x}_i is the initial position of the atom. We used the solvent interaction energies with atom i at positions \mathbf{x}_i and $\mathbf{x}_i + \Delta \mathbf{x}_i$ to test the correctness of the force $-\delta(\Delta G_s(\mathbf{x}_i)) / \delta \mathbf{x}_i$.

In the CHARMM implementation of the program, shared-charge pairs, such as the carboxy oxygens of glutamate, are handled dynamically during energy minimization or molecular dynamics by assigning the atomic solvation parameter for the charged atom to the atom with greater exposed surface area. Hydrogen atoms and nonprotein atoms such as metal ions are ignored. The solvation energy subroutine is written for the VAX/VMS implementation of CHARMM and is available on the Diskette Appendix or on request to Laura@uclaue.mbi.ucla.edu.

Determination of $\Delta \sigma$ values for vacuum-to-water transfer

Atomic solvation parameters for the transfer of atoms from vacuum to water were obtained by a least-squares fit to empirical amino acid side-chain transfer energies, using Equation 2. The procedure is the same as that used earlier (Eisenberg & McLachlan, 1986) to obtain octanol-to-water atomic solvation parameters, but here we use Wolfenden's measured free energies of transfer of amino acid side-chain analogs from vapor to water (Wolfenden et al., 1981). Wolfenden's measurements have been adjusted for the entropy of mixing both by Kyte and Doolittle (1982) and by Sharp et al. (1991). We obtained atomic solvation parameters from both sets of adjusted free energies of transfer. The side-chain analogs chosen by Wolfenden are the side chain plus a hydrogen atom. The side chain analog for glycine, a hydrogen molecule, was excluded from our calculation, as there is no atomic solvation parameter for hydrogen. Proline was excluded as it has no side-chain analog. The arginine side chain was included in the Sharp et al. free energies of transfer but not in the Kyte and Doolittle transfer energies. The Wolfenden free energies of transfer and the two sets of adjusted free energies of transfer are shown in Table 1.

The areas A_i of the amino acid side-chain analogs were estimated from amino acid side-chain coordinates in the Brookhaven Data Bank (Bernstein et al., 1977) entries 1CRN (Hendrickson & Teeter, 1981), 2CCY (Weber et al., 1981), 2CYP (Poulos et al., 1980), and 6CHA (Blevins & Tulinsky, 1985). For each amino acid, coor-

Table 1. Vapor-to-water free energies of transfer for amino acid side-chain analogs, based on measurements of Wolfenden et al. (1981), and adjustments for the entropy of mixing by Kyte and Doolittle (1982) and by Sharp et al. (1991), in units of kcal mol⁻¹

Amino acid	Side-chain analog	ΔG_{obs}^0 , Wolfenden	ΔG_{obs}^0 , Kyte and Doolittle adjustment	ΔG_{obs}^0 , Sharp et al. adjustment
Ala	Methane	1.94	2.34	2.63
Arg	<i>n</i> -Propylguanidine	-19.92		-17.46
Asn	Acetamide	-9.68	-9.04	-8.31
Asp	Acetic acid	-10.95	-10.04	-9.64
Cys	Methanethiol	-1.24	-0.63	0.01
Gln	Propionamide	-9.38	-8.59	-7.35
Glu	Propionic acid	-10.24	-9.09	-8.36
His	4-Methylimidazole	-10.27	-9.57	-8.25
Ile	1-Butane	2.15	3.06	4.89
Leu	Isobutane	2.28	3.20	5.20
Lys	<i>n</i> -Butylamine	-9.52	-8.08	-6.84
Met	Methylethyl sulfide	-1.48	-0.58	0.93
Phe	Toluene	-0.76	0.23	2.18
Ser	Methanol	-5.06	-4.63	-4.31
Thr	Ethanol	-4.88	-4.23	-3.54
Trp	3-Methylindole	-5.88	-4.77	-2.40
Tyr	4-Cresol	-6.11	-5.10	-3.17
Val	Propane	1.99	2.78	4.07

dinates for the first 20 side chains were extracted, except for methionine, for which only 15 side chains were found, and for histidine, for which only 12 side chains were found. To obtain the accessible surface areas for the cysteine side chain, the β -carbon and the γ -sulfur atoms were extracted from both cysteine and cystine residues. The Wolfenden data do not include a side chain equivalent for a disulfide-bonded cystine residue.

Accessible surface areas were calculated for the atoms in each side-chain coordinate set, with the rest of the protein deleted. Then, average areas and standard deviations were calculated for each side-chain atom in each amino acid. These areas and standard deviations are shown in Table 2. The atoms were then assigned an atom type (C, O/N, O⁻, N⁺, or S). For the shared-charge pairs in Glu, Asp, and Arg, the charge was assigned to the atom with greater accessible surface area. Average areas for each atom type were calculated for each amino acid. These areas were fit by least squares to the two sets of adjusted Wolfenden transfer energies, to obtain atomic solvation parameters. The atomic solvation parameters and their standard deviations are shown in Table 3. The two sets of atomic solvation parameters are similar, but the parameters derived from the Sharp et al. adjusted free energies are slightly more negative for nitrogen, oxygen, and sulfur.

Table 2. Average solvent-accessible areas in Å² and their standard deviations (in parentheses) of amino acid side chains from four protein structures in the Brookhaven Protein Database (Bernstein et al., 1977)^a

Ala		Arg		Asn		Asp		Cys		Gln	
CB	137 (0)	CB	74 (3)	CB	79 (1)	CB	82 (1)	CB	91 (0)	CB	74 (2)
		CG	36 (5)	CG	19 (1)	CG	29 (1)	SG	79 (0)	CG	40 (2)
		CD	32 (4)	OD1	37 (1)	OD1	37 (1)			CD	16 (2)
		CZ	12 (2)	ND2	59 (1)	OD2	38 (1)			OE1	35 (2)
		NE	15 (4)							NE2	56 (3)
		NH1	49 (3)								
		NH2	59 (1)								
Glu		His		Ile		Leu		Lys		Met	
CB	75 (1)	CB	77 (0)	CB	41 (1)	CB	68 (1)	CB	76 (2)	CB	73 (4)
CG	41 (2)	CG	8 (0)	CG1	41 (3)	CD1	68 (1)	CG	36 (4)	CG	42 (3)
CD	27 (1)	ND1	25 (1)	CD1	72 (2)	CD2	68 (1)	CD	31 (2)	SD	40 (3)
OE1	32 (2)	CE1	53 (1)	CG2	72 (2)	CG	17 (2)	CE	42 (3)	CE	76 (3)
OE2	37 (1)	NE2	28 (1)					NZ	60 (3)		
		CD2	44 (1)								
Phe		Ser		Thr		Trp		Tyr		Val	
CB	73 (0)	CB	105 (0)	CB	58 (1)	CB	70 (1)	CB	73 (1)	CB	44 (2)
CG	5 (0)	OG	43 (1)	OG1	40 (1)	CG	6 (0)	CG	6 (0)	CG1	76 (1)
CD1	32 (1)			CG2	82 (1)	CD1	43 (1)	CD1	32 (1)	CG2	76 (1)
CE1	39 (1)					CD2	6 (0)	CD2	33 (1)		
CZ	39 (1)					CE2	8 (0)	CE1	38 (1)		
CE2	39 (1)					CE3	31 (1)	CE2	38 (1)		
CD2	33 (1)					CH2	39 (2)	CZ	13 (1)		
						CZ2	38 (1)	OH	38 (1)		
						CZ3	38 (1)				
						NE1	26 (1)				

^a There were 12 occurrences of His, 15 occurrences of Met, and 20 occurrences of each of the other residue types.

Table 3. Atomic solvation parameters, $\Delta\sigma$, and their standard deviations derived from Wolfenden et al. (1981) free energies of transfer (Table 2), as modified by Kyte and Doolittle (1982) and by Sharp et al. (1991), in units of $\text{cal mol}^{-1} \text{\AA}^{-2}$

Parameter	Kyte and Doolittle adjustment	Sharp et al. adjustment
$\Delta\sigma(\text{C})$	4 ± 3	12 ± 3
$\Delta\sigma(\text{O/N})$	-113 ± 14	-116 ± 13
$\Delta\sigma(\text{S})$	-17 ± 22	-18 ± 21
$\Delta\sigma(\text{O}^-)$	-166 ± 38	-175 ± 36
$\Delta\sigma(\text{N}^+)$	-169 ± 31	-186 ± 22

Notice that all the atomic solvation parameters except for carbon are negative. This means that hydration is favorable for nitrogen, oxygen, and sulfur atoms, but unfavorable for carbon atoms. Thus, the $-\delta(\Delta G_s)/\delta x_i$ forces favor the exposure of hydrophilic atoms to solvent and compete with attractions between hydrophilic atoms, such as salt bridges, hydrogen bonds, and van der Waals forces. Because CHARMM simulations of proteins in vacuum tend to stabilize folded states (see below), the addition of the ΔG_s energy term to CHARMM could improve simulations of the dissolving of crystals and the unfolding of proteins in water.

A test of molecular dynamics with atomic solvation parameters

We tested CHARMM with the ΔG_s energy term on a melittin tetramer and monomer. Melittin, the principal toxic component of bee venom, is a 26-residue peptide that exists in an equilibrium in solution between a monomeric and tetrameric form (Habermann, 1972). The monomer is a coil in solution, and an α -helix in the tetramer (Lauterwein et al., 1980; Terwilliger & Eisenberg, 1982). The helices are highly amphiphilic, and their hydrophobic faces form the binding interface between the two dimers in the tetramer. The sequence of melittin is GIGAVLKVLTTGLPALISWIKRKRQQ (Habermann & Jentsch, 1967). Note that in addition to forming an amphiphilic helix, melittin is also hydrophobic at the N-terminus and hydrophilic at the C-terminus.

We carried out a molecular dynamics study of the melittin tetramer and monomer, using both the unmodified version of CHARMM and our version with solvation energy added. We used the atomic solvation parameters derived from the Kyte and Doolittle modification of the Wolfenden transfer energies. We obtained coordinates for a melittin tetramer from the crystal structure (Terwilliger & Eisenberg, 1982), Brookhaven Data Bank (Bernstein et al., 1977) entry 2MLT, ignoring sulfate ions and waters. We first relaxed the melittin tetramer crystal struc-

ture with 200 cycles of Adopted Basis Newton Raphson energy minimization (ABNR) in CHARMM, using a step size of 0.02 \AA . The energy parameters were obtained from the PARAM19 parameter set with some additional parameters, modified to support explicit hydrogen bonds. We used a dielectric constant of 40, based on a study by Rees (1980). The minimized tetramer was the starting coordinate set for the dynamics runs with the tetramer, and we used the coordinates of monomer 1 from the minimized tetramer to start the dynamics runs with the monomer. We then carried out 110 ps of Verlet dynamics at 300 $^\circ\text{K}$ on the monomer and tetramer, using both the original version of CHARMM and CHARMM plus vapor-to-water transfer energy. We used 1-fs time steps. At the end of the dynamics runs, the coordinate sets were energy minimized for 200 cycles using the ABNR method.

The changes to the structure of the melittin monomer after dynamics both with and without the ΔG_s energy term are shown in Figure 3. In the crystal structure (Fig. 3A), the monomer is helical, with a kink at residue 12. In the monomer after dynamics with the unmodified version of CHARMM (Fig. 3B), the kink at residue 12 has increased, and residues 12 and 13 are no longer helical. The helix has also unwound slightly between residues 16 and 26. Residues 1–11 remain helical. The RMS deviation from the crystal structure is 4.51 \AA . The structure after dynamics with CHARMM plus the ΔG_s energy term (Fig. 3C) is much more unfolded, with an RMS deviation of 6.93 \AA from the crystal structure. Residues 1–12 are now in random coil conformation. This is the more hydrophobic end of the monomer; residues 1–5 are folded back against residues 8 and 11–13, partially shielding hydrophobic side chains from water. The hydrophilic C-terminal end of the monomer was less affected by the dynamics: residues 16–22 remained helical and residues 23–26 were slightly unwound. The results of this simulation with the ΔG_s energy term agree remarkably well with experimental results. Lauterwein et al. (1980) studied monomeric melittin in aqueous solution using NMR methods and found that although monomeric melittin in solution is predominantly in a flexible extended form, residues 5–9 and 14–20 are more structured than the rest of the chain.

The ΔG_s energy term also produces more realistic conformations for the charged Lys and Arg side chains; these side chains, shown in blue in Figure 3C, extend into the solvent, as they do in experimental protein structures. In contrast, in the model that was optimized by dynamics with the unmodified version of CHARMM, the Lys and Arg side chains in the monomer bend toward the surface of the molecule. This is presumably due to favorable electrostatic interactions and hydrogen bonding of the charged groups with partial charges in the rest of the protein molecule, and van der Waals attraction between the side chains and the molecular surface.

The tetramer crystal structure and the tetramer struc-

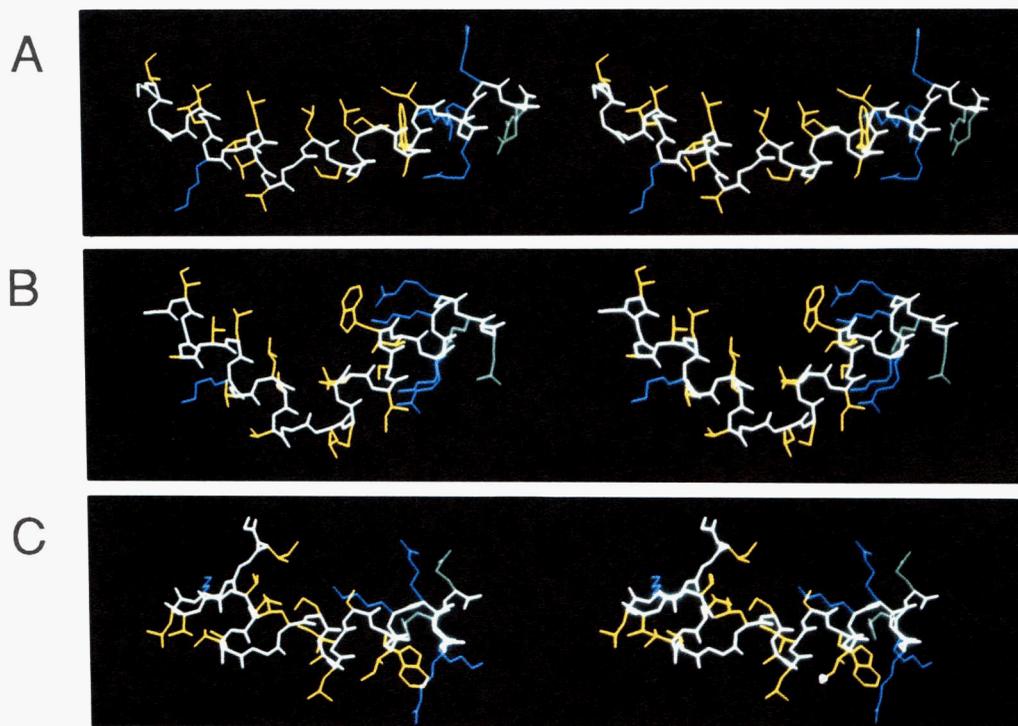


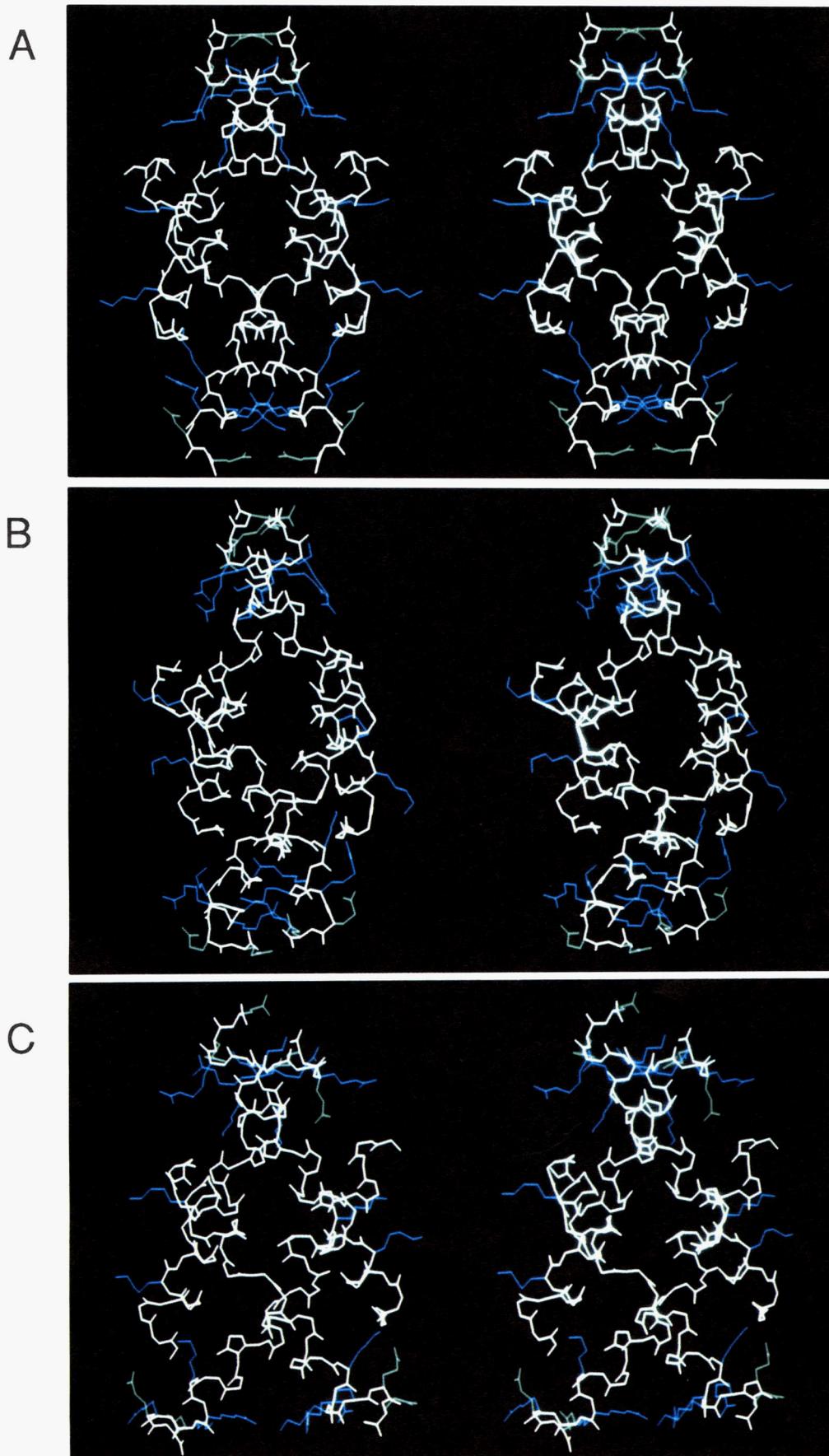
Fig. 3. Melittin monomer structures before and after CHARMM dynamics. The N-termini of the monomers are at the left. The main chain is shown in white; the charged Lys and Arg side chains are blue, Gln is shown in green, and hydrophobic and neutral side chains are yellow. Panel **A** shows the crystal structure of the monomer, which is helical with a kink. Panel **B** shows the melittin monomer after 110 ps of CHARMM dynamics. The kink has increased, and the helix has unwound slightly near the C-terminus. Panel **C** shows the melittin monomer after 110 ps of molecular dynamics with CHARMM with the solvation energy term added. The N-terminal half of the monomer has unwound and folded upon itself, forming a hydrophobic core that shields hydrophobic side chains from solvent. (See Kinemage 1.)

tures after the two dynamics runs are shown in Figure 4. The tetramer was stable during dynamics with both the original version of CHARMM and our version with the added ΔG_s energy term, which is in agreement with experiment (Terwilliger & Eisenberg, 1982). The helices did unwind slightly in both optimized models. The RMS displacement from the crystal structure was 2.98 Å for the tetramer that was optimized with CHARMM alone, and 3.41 Å for the tetramer that was optimized by CHARMM with the ΔG_s energy term. Again, CHARMM with the ΔG_s energy term moved the charged side chains into a more realistic extended conformation (shown in blue in Fig. 4C).

A breakdown of the component energies of the different models is shown in Figure 5. In the monomer after dynamics with CHARMM with solvation energy added

(bar 3), hydrogen bonds and van der Waals contacts have been broken in order to solvate hydrophilic groups, as can be seen by comparison with the crystal structure after 200 cycles of CHARMM minimization (bar 1). In the melittin tetramer after dynamics with CHARMM and solvation energy (bar 6), van der Waals contacts have been broken and solvation energy has increased, but the hydrogen bonds have not been disrupted much by comparison with the minimized crystal structure (bar 4), indicating that the helices in the tetramer remain largely intact. In the melittin monomer and tetramer after dynamics with CHARMM alone (bars 2 and 5), van der Waals contacts have increased slightly over the minimized crystal structure, and the solvation energy has actually decreased. The electrostatic energy is small compared to hydrogen bonding energy in all of the models, indicating

Fig. 4. Melittin tetramer structures before and after molecular dynamics. The main chain is white, the charged Lys and Arg side chains are blue, and Gln side chains are green. The hydrophobic and neutral side chains have been omitted. Panel **A** shows the crystal structure of the melittin tetramer, looking down the z axis. Panel **B** shows the melittin tetramer after 110 ps of CHARMM molecular dynamics. Panel **C** shows the melittin tetramer after 110 ps of molecular dynamics with CHARMM with solvation energy (ΔG_s energy term) added. The secondary structure of the tetramer was preserved in both dynamics runs, in agreement with results of experiments. (See Kinemage 2.)



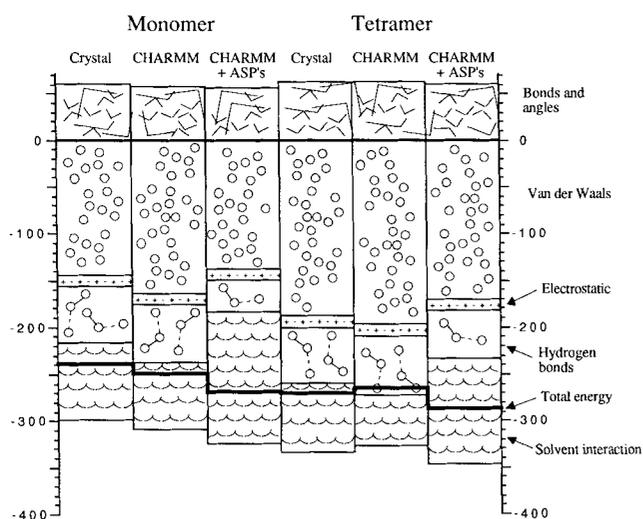


Fig. 5. Breakdown of the component energies of melittin models. The first three bars show energies for a melittin monomer after 200 cycles of CHARMM minimization, for the monomer after 110 ps of dynamics using CHARMM alone, and for the monomer after 110 ps of dynamics with CHARMM and solvation energy. The next three bars show energies for the melittin tetramer after 200 cycles of minimization, after 110 ps of dynamics with CHARMM, and after 110 ps of dynamics with CHARMM and solvation energy. The energies are in units of kcal per mole of monomer. The van der Waals, electrostatic, hydrogen bonding, and solvation (ΔG_s) energies are negative and are accumulated in this order below the zero line. The internal energy, which results from the deviation from ideality of bonds, angles, dihedral angles, and improper dihedral angles, is positive and is shown above the zero line. The heavy line near the bottom of the graph is the total energy.

that the dielectric constant of 40 that we used may be too large.

The hydrophobic and hydrophilic solvent-accessible surface areas for the melittin models are shown in Figure 6. We consider carbon atoms to be hydrophobic, and nitrogen, oxygen, and sulfur atoms to be hydrophilic. The melittin monomer after 110 ps of dynamics with CHARMM and solvation energy (bar 3) has increased hydrophilic surface area and decreased hydrophobic surface area, relative to the minimized crystal structure of the monomer (bar 1). Similarly, dynamics with CHARMM and solvation energy increased the hydrophilic exposed surface area of the tetramer (bar 6), relative to the minimized crystal structure (bar 4). In comparison, CHARMM alone decreased the hydrophilic exposed area in the monomer and tetramer (bars 2 and 5), while increasing the hydrophobic exposed area in the tetramer (bar 5).

Conclusions

We have presented an analytical procedure for combining the powerful programs for molecular modeling of protein energetics with a term that accounts for solvent

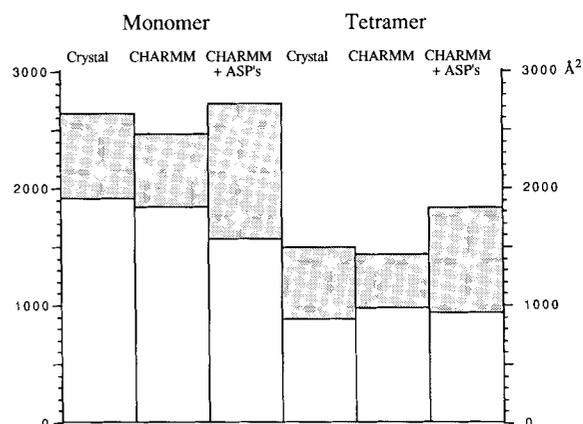


Fig. 6. Hydrophilic and hydrophobic solvent-accessible surface areas in six melittin models, in units of \AA^2 per monomer, as computed by ANAREA (Richmond, 1984). The solvent accessible areas of carbon atoms are unshaded; the areas of nitrogen, oxygen, and sulfur atoms are shaded. The first three bars show areas for a melittin monomer after 200 cycles of CHARMM minimization, for the monomer after 110 ps of dynamics using CHARMM alone, and for the monomer after 110 ps of dynamics with CHARMM and solvation energy. The next three bars show areas for the melittin tetramer after 200 cycles of minimization, after 110 ps of dynamics with CHARMM, and after 110 ps of dynamics with CHARMM and solvation energy.

interactions. To describe the solvent interactions, we have derived atomic solvation parameters for the free energy of transfer of protein groups from vacuum to water. When applied to a simple test calculation on melittin, the procedure gives reasonable results. Clearly, further tests are required, but the ones presented here suggest that the addition of atomic solvation terms to conventional molecular dynamics may provide a more realistic description of proteins in water.

Acknowledgments

We thank NIH and NSF for support.

References

- Bernstein, F.C., Koetzle, T.F., Williams, G.J., Meyer, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977). The Protein Data Bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112(3), 535-542.
- Blevins, R.A. & Tulinsky, A. (1985). The refinement and the structure of the dimer of alpha-chymotrypsin at 1.67 Å resolution. *J. Biol. Chem.* 260(7), 4264-4275.
- Brooks, B.R., Brucoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., & Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4(2), 187-217.
- Brown, L.R., Lauterwein, J., & Wuethrich, K. (1980). High-resolution NMR studies of self-aggregation of melittin in aqueous solution. *Biochim. Biophys. Acta* 622(2), 231-244.
- Connolly, M.J. (1983). Analytical molecular surface calculation. *J. Appl. Crystallogr.* 16, 548-558.

- Eisenberg, D. & McLachlan, A.D. (1986). Solvation energy in protein folding and binding. *Nature* 319(16), 199–203.
- Eisenberg, D., Wesson, M., & Yamashita, M. (1989). Interpretation of protein folding and binding with atomic solvation parameters. *Chem. Scr.* 29A, 217–221.
- Habermann, E. (1972). Bee and wasp venoms. *Science* 177, 314–322.
- Habermann, E. & Jentsch, J. (1967). Sequential analysis of melittin from tryptic and peptic fragments. *Hoppe-Seyler's Z. Physiol. Chem.* 348(1), 37–50.
- 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.
- Hendrickson, W.A. & Teeter, M.M. (1981). Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulphur. *Nature* 290, 107–113.
- Jorgensen, W.L. & Tirado-Rives, J. (1988). The OPLS potential function for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* 110(6), 1657–1666.
- Kauzmann, W. (1959). Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.* 14, 1–63.
- Kyte, J. & Doolittle, R.F. (1982). A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157(1), 105–132.
- Lauterwein, J., Brown, L.R., & Wuethrich, K. (1980). High-resolution proton NMR studies of monomeric melittin in aqueous solution. *Biochim. Biophys. Acta* 622(2), 219–230.
- Lee, B. & Richards, F.M. (1971). Interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.* 55(3), 379–400.
- Momany, F.A., McGuire, R.F., Burgess, A.W., & Scheraga, H.A. (1975). Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. *J. Phys. Chem.* 79, 2361–2381.
- Nemethy, G., Pottle, M.S., & Scheraga, H.A. (1983). Energy parameters in polypeptides. 9. Updating of geometrical parameters, nonbonded interactions, and hydrogen bond interactions for the naturally occurring amino acids. *J. Phys. Chem.* 87, 1883–1887.
- Ooi, T., Oobatake, M., Nemethy, G., & Scheraga, H.A. (1987). Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides. *Proc. Natl. Acad. Sci. USA* 84, 3086–3090.
- Poulos, T.L., Freer, S.T., Alden, R.A., Edwards, S.L., Skogland, U., Takio, K., Eriksson, B., Xuong, N.H., Yonetani, T., & Kraut, J. (1980). The crystal structure of cytochrome c peroxidase. *J. Biol. Chem.* 255(2), 575–580.
- Rees, D. (1980). Experimental evaluation of the effective dielectric constant of proteins. *J. Mol. Biol.* 141(1), 323–326.
- Richards, F.M. (1977). Areas, volumes, packing, and protein structure. *Annu. Rev. Biophys. Bioeng.* 6, 151–176.
- 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(1), 63–89.
- Schiffer, C.A., Caldwell, J.W., Stroud, R.M., & Kollman, P.A. (1992). Inclusion of solvation free energy with molecular mechanics energy: Alanine dipeptide as a test case. *Protein Sci.* 1, in press.
- Sharp, K., Nicholls, A., Friedman, R., & Honig, B. (1991). Extracting hydrophobic free energies from experimental data: Relationship to protein folding and theoretical models. *Biochemistry*, in press.
- Sippl, M.J., Nemethy, G., & Scheraga, H.A. (1984). Intermolecular potentials from crystal data. 6. Determination of empirical potentials for O–H...O=C hydrogen bonds from packing configurations. *J. Phys. Chem.* 88, 6231–6233.
- Terwilliger, T.C. & Eisenberg, D. (1982). The structure of melittin. II. Interpretation of the structure. *J. Biol. Chem.* 257(11), 6016–6022.
- Van Gunsteren, W.F. & Berendsen, H.J.C. (1987). *Groningen Molecular Simulation (GROMOS) Library Manual*. Biomos, Groningen, The Netherlands.
- Vila, J., Williams, R.L., Vasquez, M., & Scheraga, H.A. (1991). Empirical solvation models can be used to differentiate native from near-native conformations of bovine pancreatic trypsin inhibitor. *Proteins Struct. Funct. Genet.* 10, 199–218.
- Weber, P.C., Howard, A., Xuong, N.H., & Salemme, F.R. (1981). Crystallographic structure of *Rhodospirillum molischianum* ferricytochrome c' at 2.5 Å resolution. *J. Mol. Biol.* 153(2), 399–424.
- Weiner, P.K. & Kollman, P.A. (1981). AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions. *J. Comput. Chem.* 2(3), 287–303.
- Wodak, S.J. & Janin, J. (1980). Analytical approximation to the surface area of proteins. *Proc. Natl. Acad. Sci. USA* 77(4), 1736–1740.
- Wolfenden, R., Andersson, L., Cullis, P.M., & Southgate, C.C.B. (1981). Affinities of amino acid side chains for solvent water. *Biochemistry* 20, 849–855.