Simple Physics-Based Analytical Formulas for the Potentials of Mean Force for the Interaction of Amino Acid Side Chains in Water. 2. Tests with Simple Spherical Systems

Mariusz Makowski,[†,‡] Adam Liwo,[†] Katarzyna Maksimiak,[§] Joanna Makowska,[†,‡] and Harold A. Scheraga*[,†]

Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853-1301, Faculty of Chemistry, University of Gdańsk, Sobieskiego 18, 80-952 Gdańsk, Poland, and Michigan State University, 218 Biochemistry, East Lansing, Michigan 48824

Received: September 11, 2006; In Final Form: December 13, 2006

Simple analytical functions consisting of electrostatic, polarization, Lennard-Jones or modified Lennard-Jones, and cavity terms are proposed to express the potentials of mean force analytically for spherical particles interacting in water. The cavity term was expressed either through the molecular-surface area of the solute or by using the Gaussian-overlap model of hydrophobic hydration developed in paper 1 of this series. The analytical expressions were fitted to the potentials of mean force of a methane homodimer, heterodimers composed of a methane molecule, and an ammonium cation or a chloride anion, respectively, and dimers consisting of two chloride anions, two ammonium cations, or a chloride ion and an ammonium cation. The potentials of mean force for these dimers were determined by umbrella-sampling molecular dynamics simulations with the AMBER 7.0 force field with TIP3P water either in our earlier work or in this work. For all systems, the analytical formulas fitted the potentials of mean force very well. However, using the molecular-surface area to express the cavity term provided a good fit only when the nonbonded interactions were expressed by an all-repulsive modified Lennard-Jones potential but also resulted in non-physical values of some of the parameters. Conversely, the use of our new Gaussian-overlap-based expression for the cavity term provided a good fit to the potentials of mean force (PMFs) with Lennard-Jones nonbonded potential, and the values of all parameters were physically reasonable.

Introduction

For several years, we have been developing a united-residue physics-based force field termed UNRES[1−7] for *ab initio* predictions of protein structure. Recently, the application of UNRES has been extended to simulate protein folding pathways.[8] UNRES has been derived[3] as a cluster cumulant expansion[9] of the restricted free energy (RFE) function of polypeptide chains in which the less important degrees of freedom (the degrees of freedom of the solvent molecules, the dihedral angles of the side chains, and the angles of rotation of the peptide groups about the virtual bond axes) are integrated out. However, some of the terms, including the potential of mean force (PMF) of interaction of united side chains ($U_{SC_iSC_j}$), have been derived on the basis of distribution functions determined from the Protein Data Bank (PDB)[10] and fitted[2] to arbitrary Gay−Berne[11] functional forms that take into account the anisotropy of the interactions. The computed $U_{SC_iSC_j}$ potentials are connected to the thermodynamics of protein interactions because they were rescaled[2] to reproduce the partition coefficients of amino acids between water and *n*-octanol. However, because the $U_{SC_iSC_j}$ components account for hydrophobic/hydrophilic interactions, which are one of the main factors governing protein folding, it is necessary to replace this part of UNRES in order to obtain a fully physics-based form. Moreover, such an approach will enable us to obtain functional forms that account better for the actual distance and orientation dependence of $U_{SC_iSC_j}$. The most undesirable feature of the present functional expressions is that they treat the interacting side chains as ellipsoids of revolution. Although this is reasonable for nonpolar side chains that interact uniformly through hydrophobic interactions, the so-called charged or polar side chains (e.g., lysine or glutamine) have polar heads and nonpolar hydrocarbon tails and are, therefore, amphiphilic. The symmetry of their effective interaction potentials should, therefore, be lower than that of the Gay−Berne potential.

In our earlier work,[12−16] we used umbrella-sampling molecular dynamics simulations to study the potential of mean force of hydrophobic interactions[12−14,16] as well as interactions between charged particles, and between nonpolar particles.[15] In those papers,[12−16] the resulting PMFs have been discussed and compared with earlier results obtained by other authors. In this paper, we describe simple physics-based functional forms that approximate the PMFs of charged and nonpolar systems and fit them to the PMFs of simple spherical systems determined by molecular dynamics simulations in this and in our earlier[14,16] work.

Theory. We consider three types of interactions: between nonpolar particles (also termed hydrophobic interactions), between charged particles, and between nonpolar and charged particles. In this work, we restrict the considerations to systems with spherical symmetry. For this reason, we leave the treatment of the interactions involving amphiphilic particles to our future work. We propose the following general analytical approxima-

* Corresponding author. E-mail: has5@cornell.edu. Phone: (607) 255-4034. Fax: (607) 254-4700.
[†] Baker Laboratory of Chemistry and Chemical Biology.
[‡] Faculty of Chemistry.
[§] Michigan State University.

10.1021/jp065917k CCC: $37.00
Published on Web 02/27/2007

tions to the corresponding PMFs:

$$W_{nn} = E_{vdW} + \Delta F_{cav} \quad (1)$$

$$W_{cc} = E_{vdW} + E_{el} + E_{pol}^{GB} + \Delta F_{cav} \quad (2)$$

$$W_{cn} = E_{vdW} + E_{pol} + E_{pol}^{GB} + \Delta F_{cav} \quad (3)$$

where W_{nn}, W_{cc}, and W_{cn} refer to the PMFs of neutral−neutral, charged−charged, and charged-neutral particle pairs, respectively, E_{vdW} is the van der Waals term, ΔF_{cav} is the difference between the cavity contribution to the free energy of hydration of the dimer and that of the isolated monomers, E_{el} is the electrostatic interaction (Coulombic) term, E_{pol} is the polarization energy coming from the interactions between the two solute particles but not from their interactions with the solvent, and E_{pol}^{GB} is the energy contribution due to solvent polarization; it is computed by using the generalized Born (GB) model.[17−21]

For E_{vdW}, we examined two types of expressions: the Lennard-Jones 6−12 potential and the Kihara[22] (modified Lennard-Jones) potential. They are given by eqs 4 and 5, respectively.

$$E_{LJ} = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \quad (4)$$

where r_{ij} is the distance between the centers of the interacting sites, ϵ_{ij} is the well depth, and σ_{ij} is the closest-approach distance (at which E_{LJ} becomes 0 and increases with decreasing distance).

$$E_{Kihara} = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij} - \sigma_{ij}}\right)^{m} + s\left(\frac{\sigma_{ij}}{r_{ij} - \sigma_{ij}}\right)^{n}\right] \quad (5)$$

where ϵ_{ij} and σ_{ij} are the parameters of the potential and $s = \pm 1$; for $s = -1$, ϵ_{ij} and σ_{ij} are the well-depth and closest-approach distance, respectively, and for $s = +1$, the potential has no minimum. The values of m and n are given in the Results and Discussion.

The Coulombic term is given by eq 6.

$$E_{el} = 332\left(\frac{q_i q_j}{r_{ij}}\right) \quad (6)$$

where q_i and q_j are the charges of sites i and j, respectively, and the coefficient 332 is introduced to express the energies in kcal/mol.

For the solvent-polarization part involving a pair of charged particles, we adopt the expression from the generalized Born model.[17−21]

$$E_{pol}^{GB} = 332 q_i q_j \left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_{out}}\right)\frac{1}{f_{GB}(r_{ij})} \quad (7)$$

where ϵ_{in} is the effective dielectric constant of the "inside" of the interacting particles, ϵ_{out} is the effective dielectric constant of the solvent, and f_{GB} is expressed by eq 8.

$$f_{GB}(r_{ij}) = \sqrt{r_{ij}^2 + R_i R_j \exp\left(-\frac{r_{ij}^2}{4R_i R_j}\right)} \quad (8)$$

where R_i and R_j are the Born radii of sites i and j, respectively. We set $\epsilon_{out} = 80$ (for water) and $\epsilon_{in} = 1$.

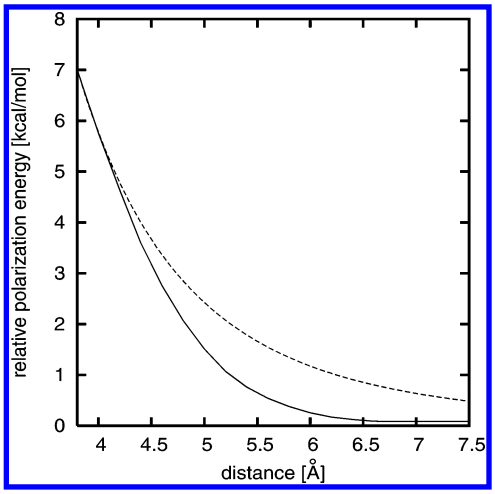

Figure 1. Comparison of the variation of the polarization component of the free energy of hydration of a pair composed of a particle with unit electron charge and a neutral spherical solute particle, each with radius 1.9 Å (this radius corresponds to a methane molecule or an ammonium cation), in a continuous dielectric computed by solving the Poisson equation numerically (solid line) or by using eq 9 (dashed line). The energy is expressed relative to the hydration energy of a spherical particle with a unit electron charge and radius of 1.9 Å, which is equal to $E = 332(1/\epsilon_{out} - 1/\epsilon_{in})(1/R)$, where R is the radius of the charged particle, ϵ_{out} is the dielectric constant of the solvent, and ϵ_{in} is the effective dielectric constant inside the particle. We assumed $\epsilon_{in} = 1$, $\epsilon_{out} = 80$, and that the surface area of the solute is the molecular surface illustrated in Figure 2 and expressed by eqs 11−17.

We propose eq 9 to express the polarization component involving the interaction between charged and nonpolar particles.

$$E_{pol} = \alpha_{ij}\left(\frac{1}{f_{GB}(r_{ij})}\right)^4 \quad (9)$$

where α_{ij} is related to the polarizability of the nonpolar particle.

At large distances, this expression varies as $1/r_{ij}^4$. The rationale for this expression is that a nonpolar particle replaces the solvent at distance r_{ij}, consequently removing a part of the polarization interaction with the solvent. The polarization interaction energy is proportional to the square of the electric field, which, by Coulomb's law, varies as $1/r_{ij}^4$. Figure 1 compares the distance dependence of E_{pol} computed for a pair composed of a charged and a neutral spherical solute in a continuous dielectric with the dielectric constant $\epsilon = 80$ (corresponding to water) calculated numerically in this work, by solving the Poisson equation by means of the boundary-element method,[23] with the distance dependence computed from eq 9. It can be seen that, in both cases, the polarization energy decreases with distance, although E_{pol} computed by solving the Poisson equation falls off faster with the distance than the expression of eq 9.

To express the cavity contribution to the free energy of hydrophobic association, ΔF_{cav}, we used the expression on the basis of molecular-surface area S and, alternatively (eq 18), the expression based on a Gaussian representation of the water density in the first hydration sphere derived in paper 1 of this series[24].

Using the molecular-surface area of the solute, ΔF_{cav} is expressed by eq 10.

$$\Delta F_{cav} = \gamma \Delta S \quad (10)$$

where γ is the local surface tension and ΔS is the difference between the molecular-surface area of the solutes at a given distance and that of the separate solutes. The value of γ can be

taken as measured for a given solvent or treated as an adjustable parameter; the second approach is used here because the dimensions of a solute are small and the macroscopic surface tension coefficient is not fully adequate. For a pair of spherical solute particles, the molecular-surface area can be expressed analytically by eq 11.[25]

$$S(x) = \begin{cases} 2\pi\left[R_{s1}\left(1+\frac{x_1}{R_1}\right) + r[\alpha_1(x)\sqrt{R_1^2 - x_1^2} - Z_{21}(x)] + \right. & \\ \left. R_{s2}\left(1+\frac{x_2}{R_2}\right) + r[\alpha_2(x)\sqrt{R_2^2 - x_2^2} - Z_{22}(x)] \right. & x \le R_1 + R_2 \\ 4\pi(R_{s1}^2 + R_{s2}^2) & x > R_1 + R_2 \end{cases} \quad (11)$$

$$\alpha_i(x) = \begin{cases} \cos^{-1}\left(\sqrt{1 - \frac{x_i^2}{R_i^2}}\right) & x \le \sqrt{R_1^2 - r^2} + \sqrt{R_2^2 - r^2} \\ \cos^{-1}\left(\frac{R_i^2 + r^2 - (x_i - Z_{1i}(x))^2}{2R_i r}\right) & x > \sqrt{R_1^2 - r^2} + \sqrt{R_2^2 - r^2} \end{cases} \quad (12)$$

$$Z_{1i}(x) = \begin{cases} 0 & x \le \sqrt{R_1^2 - r^2} + \sqrt{R_2^2 - r^2} \\ \sqrt{r^2 - R_i^2 + x_i^2} & x > \sqrt{R_1^2 - r^2} + \sqrt{R_2^2 - r^2} \end{cases} \quad (13)$$

$$Z_{2i}(x) = \frac{r}{R_i} x_i - Z_{1i}(x) \quad (14)$$

$$R_i = R_{si} + r \quad (15)$$

$$x_1 = \frac{1}{2}\left(x + \frac{R_1^2 - R_2^2}{x}\right) \quad (16)$$

$$x_2 = \frac{1}{2}\left(x + \frac{R_2^2 - R_1^2}{x}\right) \quad (17)$$

where R_{s1} and R_{s2} are the radii of the first and the second solute molecule, respectively, r is the radius of the solvent molecule (the probe in Figure 2), x_1 and x_2 are the distances of the projection of the center of the solvent molecule on the line linking the centers of the solute molecules from the center of the first and the second solute molecules, respectively, and $x = x_1 + x_2$ is the distance between the centers of the solute molecules (Figure 2).

The notation in eq 11 and in Figure 2 follows that of ref 25, although the two solute molecules are assumed to have equal radii in ref 25; it can easily be shown that eq 11 becomes eq 17 in ref 25 if we set $R_{s1} = R_{s2} = R_s$. The index i in eqs 12−17 runs from 1 to 2; the index 1 denotes the first, and the index 2 the second, solute particle in Figure 2. As can be seen from eq 11, the molecular-surface area $S(x)$ and, consequently, ΔF_{cav} expressed by eq 10 has a discontinuous gradient at $x = R_1 + R_2 = R_{s1} + R_{s2} + 2r$. In our earlier work,[26] we discussed the discontinuity of the gradient of the molecular-surface area for clusters of more than two solute molecules. Moreover, the molecular-surface area cannot be expressed analytically for solute particles with the symmetry of an ellipsoid of revolution, such as that of the side-chain model of UNRES. Therefore, as an alternative to using eq 11, in paper 1[24] of this series we derived an alternative expression (eq 16 in ref 24) for the cavity term, which is cited here as eq 18. This equation is based on hydration-shell volume[27−33] and not molecular-surface area. In ref 24, we implemented the Gaussian-overlap model introduced by Berne and Pechukas[34] to express the change of the free energy of hydration (ΔF_{cav}) upon the formation of dimers and larger clusters of spherical nonpolar molecules in water and, consequently, the free energy of hydration of the nonpolar-solute dimer relative to a pair of isolated spherical solute particles. Details are provided in paper 1[24] of this series

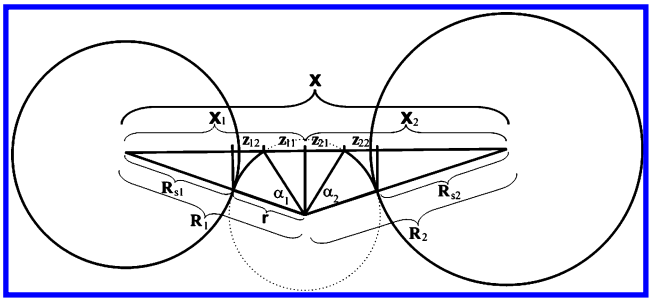

Figure 2. Illustration of analytical molecular-surface area calculation. The probe (marked with a dotted circle except for the part that becomes a part of the molecular surface of the dimer, which is shown as a solid arc), which has a radius r, is in contact with two solute particles, with radii R_{s1} and R_{s2}, respectively, and penetrates between two particles to the line connecting the solute centers (which are at distance x). Z_1 is the distance between the projection of the center of the probe on the line linking the centers of the solute molecules and the point of the intersection of the probe sphere with the line linking the solute centers, expressed by eq 13, Z_2 is the distance between the point of the intersection of the probe sphere with the line linking the solute centers and the projection of the tangent circle of the probe sphere and the solute molecule on the line linking the solute centers (eq 14), and α is the angle between the line pointing from the center of the probe sphere to the center of a solute molecule, expressed by eq 12. Z_1, Z_2, and α are functions of the distance between solute centers, x.

$$\Delta F_{cav} = \frac{\alpha_{ij}^{(1)} x^{1/2} + \alpha_{ij}^{(2)} x - \alpha_{ij}^{(3)}}{1 + \alpha_{ij}^{(4)} x^{12}} \quad (18)$$

where $\alpha_{ij}^{(1)}$, $\alpha_{ij}^{(2)}$, $\alpha_{ij}^{(3)}$, and $\alpha_{ij}^{(4)}$ are constants of the expression and $x = r_{ij}/\sqrt{\sigma_i^2 + \sigma_j^2}$ where σ_i can be identified with the solvation radius of particle i.

Methods

We considered the following pairs of particles: methane−methane (m−m), methane−chloride ion (m−Cl^-), methane−ammonium ion (m−NH_4^+), a pair of chloride anions (Cl^-...Cl^-), a pair of ammonium cations (NH_4^+...NH_4^+), and an ammonium−chloride ion pair (NH_4^+...Cl^-) to model the effective interactions of nonpolar, nonpolar and charged, like-charged, and oppositely charged solutes in water, respectively. The PMFs of the m−m, m−NH_4^+, Cl^-...Cl^-, and NH_4^+...NH_4^+ pairs were taken from our previous work[15] where they were determined by umbrella-sampling molecular dynamics simulations with the TIP3P[35] water model; the PMFs of the m−Cl^- and NH_4^+...Cl^- pairs were determined in this work by the procedure used in our earlier work,[15,36,37] as summarized briefly below.

For each of the m−Cl^- and NH_4^+...Cl^- systems, a series of six umbrella-sampling molecular dynamics (MD) simulations were carried out with harmonic restraints imposed on the distance between the centers of the solute particles. The force constant was $k = 2.0$ kcal/(mol Å^2), and the restraints were centered at 4.0, 5.0, 6.0, 7.0, 8.0, and 9.0 Å in simulations 1−6, respectively. Each simulation carried out with a given restraint

potential is termed a *window*. The data from all windows were collected and processed together by using the Weighted Histogram Analysis Method (WHAM).[38,39] In our earlier work,[12,40] we demonstrated that umbrella sampling followed by processing of the simulation results with WHAM gives much more reliable PMFs compared with thermodynamic integration, the free-energy perturbation method and the particle-insertion method when the solute molecules are of the size of a methane molecule.

Molecular dynamics simulations were carried out with the AMBER[41] suite of programs, using the AMBER 7.0 force field[42] in a cubic box with a 32 Å side. With the dimensions of the box largely exceeding that of the solute molecules, the possible dependence of the results on box shape is not a concern; however, it should be noted that this dependence is remarkable when the solute molecule is comparable with the box size, e.g., in the case of all-atom MD simulations of proteins.[43] The number of water molecules corresponded to the experimental solvent density at 298 K. The number of water molecules in a box was 792 for the m−Cl^- system and 742 for the NH_4^+−Cl^- system. The TIP3P[35] model of water was used. The charges on the atoms of the ammonium cation were taken from ref 15. The methane molecule was treated as a united atom with the parameters of the nonbonded Lennard-Jones potential taken from ref 44, and AMBER-standard values of the other parameters were used. MD simulations were carried out in two steps. In the first step, each system was equilibrated in the *NPT* scheme (constant number of particles, pressure, and temperature) at 298 K for 100 ps (picoseconds), and the integration step was 2 fs (femtoseconds). After equilibration, production MD simulations were carried out in the *NVT* scheme (constant number of particles, volume, and temperature) for 2 ns, and the integrations step was 1 fs. A 9 Å cutoff for nonbonded interactions was imposed. The electrostatic energy was evaluated using the particle-mesh Ewald summation,[45] and counterions (sodium, if the solute had a net negatively charge or chloride, if the solute had a net positive charge) were added to the systems to zero the net charge. A total of 2 000 000 configurations were generated and subsequently used to calculate the PMF for each of the m−Cl^- and NH_4^+...Cl^- systems.

Fitting of analytical formulas to the PMFs was accomplished by minimizing the sum of the squares (Φ), defined by eq 19, of the differences between the PMF values computed from the analytical formulas and those determined from MD simulations. Marquardt's method[46] was used to minimize Φ:

$$\min_x \Phi(\mathbf{y}) = \sum_i [W^{\mathrm{MD}}(r_i) - W^{\mathrm{anal}}(r_i;\mathbf{y})]^2 \qquad (19)$$

where $W^{\mathrm{MD}}(r_i)$ is the PMF value determined by simulations for distance r_i and $W^{\mathrm{anal}}(r_i;\mathbf{y})$ is the analytical approximation to the PMF at distance r_i calculated using adjustable parameters of eqs 4−5, 7−11, and 18, given by the components of the vector $\mathbf{y}$.

To avoid unreasonable values of the parameters, a simple model consisting either of the nonbonded (Lennard-Jones or Kihara) or nonbonded + electrostatic + GB potential was fitted first, and the surface-solvation contribution (expressed either by eq 11 or eq 18, respectively), was subsequently added and all parameters were re-fitted.

Results and Discussion

The PMFs of the pairs considered, together with curves computed with the fitted analytical expressions, are shown in Figure 3. As discussed in our previous work,[15] all PMFs including those of pairs of like-charged ions, exhibit fine structure with minima and maxima present. For the m−m and NH_4^+...Cl^- pairs, the first minimum is termed the contact minimum and the second one the solvent-separated minimum, and the contact minimum occurs at distances shorter than the sum of the vdW radii of the interacting particles. We note, however, that the contact minimum of the NH_4^+...Cl^- pair is remarkably shallower than that of the K^+...Cl^- pair studied in our earlier work,[15] the difference between the PMF at the contact minimum and the first maximum being about 0.25 kcal/mol as opposed to 2 kcal/mol for the K^+...Cl^- pair. Also, the PMF of the NH_4^+...Cl^- pair falls off faster with distance than would be expected for a pair of opposite unit charges in a continuous dielectric, which means that the charges are very efficiently screened by the solvent probably because of effective hydrogen-bond formation between water molecules and the ammonium cation. For like-charged particles, the first maximum originates from a local increase of the concentration of water molecules and counter-ions near the intersection of the molecular surfaces of the particles involved and occurs at distances longer than the sum of the vdW radii.[15]

We first tried to fit the PMFs to analytical expressions with ΔF_{cav} expressed in terms of the molecular-surface area (eq 11). Use of the Lennard-Jones potential or the Kihara potential with a minimum ($s = -1$ in eq 5) gave poor results for all systems. A good fit was obtained with the all-repulsive Kihara potential ($s = +1$ in eq 5) with exponents $m = 6$ and $n = 3$ (Figure 3). The resulting fitting curves, together with the PMF curves determined by simulations, are shown in Figure 3a−f, and the parameters of the analytical expressions are shown in Table 1.

It can be seen from Figure 3 that the simulated PMFs are reproduced fairly well including the desolvation maximum. The solvent-separated minimum is not reproduced by eq 11. This is not a problem, however, because the oscillations of the PMFs beyond the desolvation maximum are smaller than the expected accuracy of the UNRES force field. However, the best-fitting nonbonded component is expressed as an all-repulsive Kihara potential and, consequently, is unphysical. Moreover, it would be difficult to extend the surface-solvation component expressed by eq 11 to a non-spherical side-chain representation of UNRES,[2] and furthermore, eq 11 can lead to a non-continuous gradient of the energy,[26] this being unacceptable for local minimization or MD simulations with the UNRES force field. Another problem is that the values of the surface tension γ are negative for the Cl^-...Cl^- and the NH_4^+...NH_4^+ pair, which means that the ΔF_{cav} term loses its physical meaning. Attempts at forcing positive values of γ for the two systems mentioned above resulted in poor fits. For the remaining systems, the values of γ are reasonable [the value of γ of water at 298 K is 72 dyn/cm = 0.10 kcal/(mol Å^2)].[47] It should be noted that the effective value of the surface-tension parameter is also lower than the macroscopic value for water in other surface-solvation models.[18]

The results of fitting the PMFs, using the Gaussian-overlap-based expression for the cavity component of the potential (eq 18), are also shown in Figure 3, and the parameters of the corresponding analytical expressions for the PMF components are collected in Table 2.

As opposed to using the molecular-surface area approximation to the cavity component (eq 11), a good fit was obtained with the Lennard-Jones potential. For the methane−methane pair (Figure 3a), the analytical approximation (solid line) reproduced the simulated PMF (dashed line) very well, including the

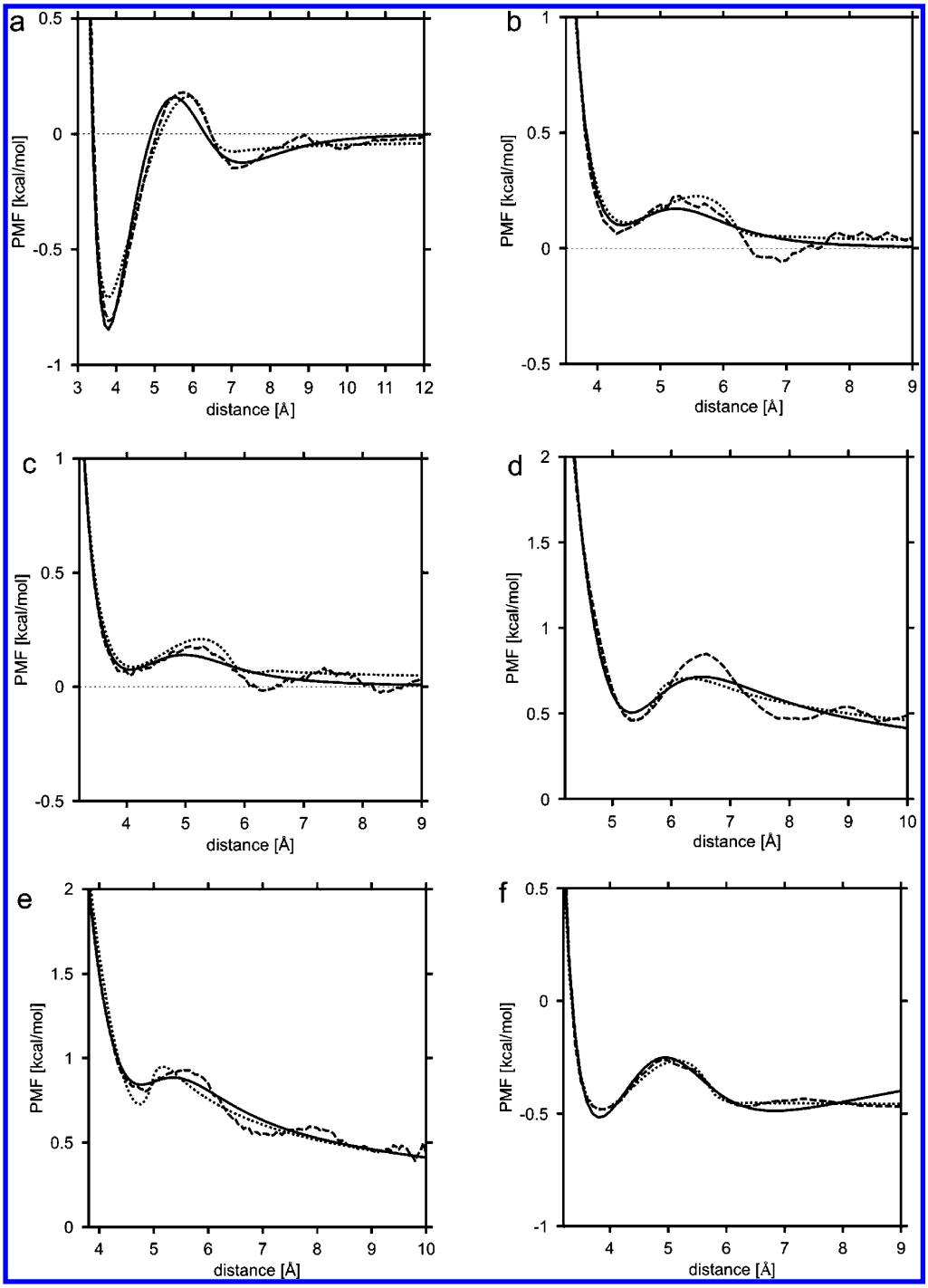

Figure 3. The PMF curves for systems studied in this work: methane homodimer (a), methane−chloride anion heterodimer (b), methane−ammonium cation heterodimer (c), chloride−anion homodimer (d), ammonium−cation homodimer (e), and ammonium cation−chloride anion heterodimer (f). The dashed lines correspond to PMFs obtained by molecular simulations. The solid lines correspond to the analytical approximation to the PMFs (with parameters determined by least-square fitting to the PMF determined by simulations) in which the Lennard-Jones potential (eq 4) was used to express the van der Waals Lennard-Jones 6−12 potential, and the cavity component (ΔF_{cav}) was expressed by eq 18. The dotted lines correspond to the analytical approximations to the PMFs in which the Kihara potential (eq 5) was used to express the van der Waals energy, and ΔF_{cav} was expressed in terms of molecular-surface area with eqs (11−17). The thin dotted lines show the asymptotic behavior of the PMFs.

solvent-separated minimum, which makes the model with the cavity potential expressed by eq 18 superior to that with the molecular-surface area approximation. However, for systems involving charged particles (Figure 3b−f), the use of eq 11 results in a comparable or better fit in the region of the desolvation maximum. Nevertheless, the differences between the two models are of the order of 0.1 kcal/mol, which does not seem to be an issue given the inaccuracies inherent in coarse-grained force fields. The analytical expressions for the PMF containing the cavity potential expressed by eq 18 has a continuous gradient over the entire range of distances. We, therefore, consider the models with the cavity component of the PMF expressed by eq 18 as the best candidates for the effective physics-based side-chain−side-chain interaction potentials for the UNRES force field.

We also tried to treat the local dielectric constant (ϵ_{in} in eq 7) as an adjustable parameter. However, the best-fitting values of ϵ_{in} were always close to 1.

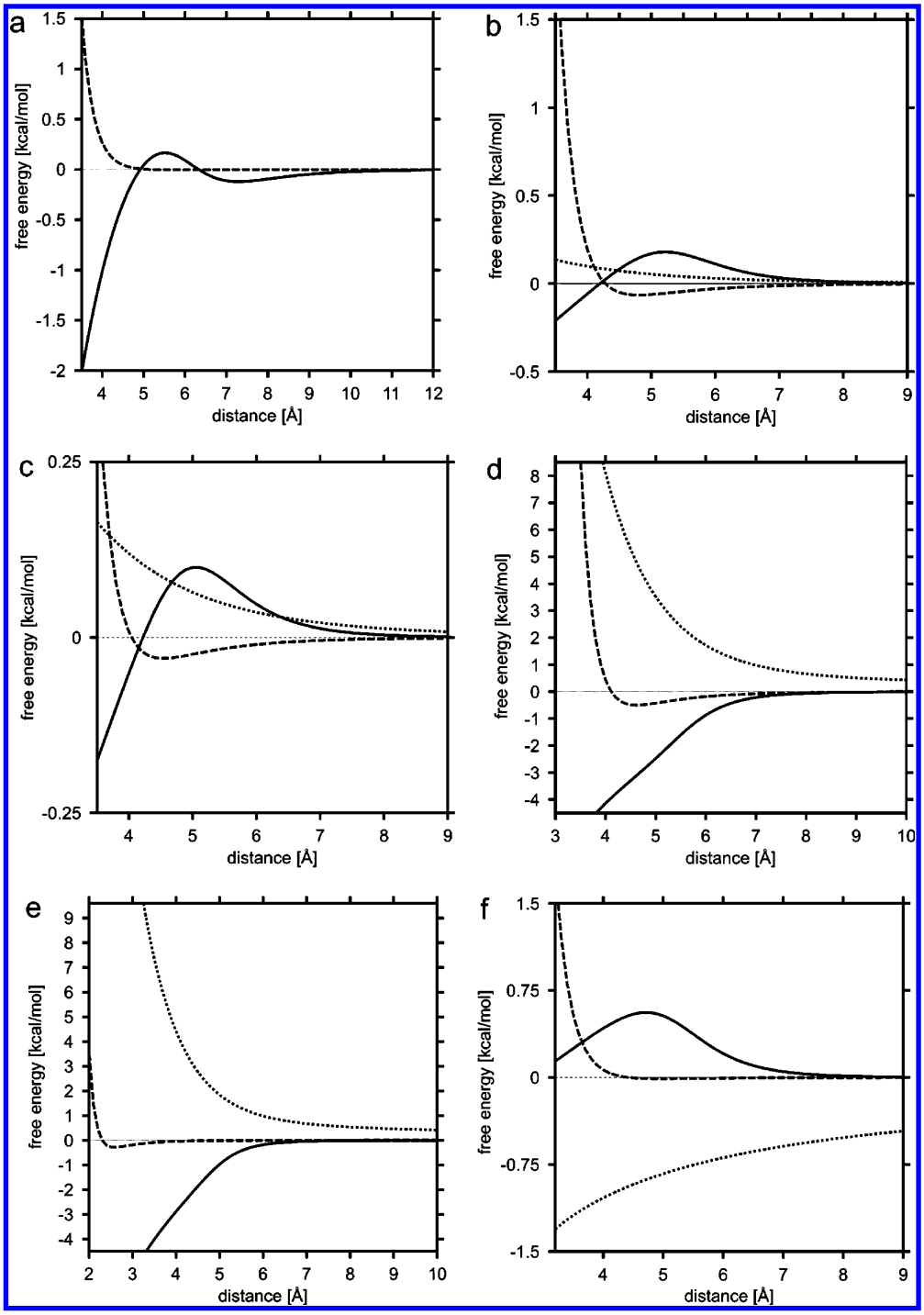

Figure 4. Plots of the energy components of the analytical approximations to the PMFs for systems studied in this work: methane homodimer (a), methane−chloride anion heterodimer (b), methane−ammonium cation heterodimer (c), chloride−anion homodimer (d), ammonium−cation homodimer (e), and ammonium cation−chloride anion heterodimer (f). The energy components of the PMF are as follows: cavity contribution to the free energy of hydrophobic association (ΔF_{cav}; solid line), Lennard-Jones potential (E_{LJ}; dashed line), and the sum of the electrostatic and generalized Born potential energy ($E_{el} + E_{pol}^{GB}$; dotted line), respectively. The polarization component is not present in Figure 4a, because the two solute particles are neutral.

The PMFs of all systems studied in this work, fitted by using the expression for the cavity potential of eq 18, are broken down into components in Figure 4. It can be seen that the components vary with the distance as expected from the physics of the respective interactions; i.e., the E_{LJ} components possess minima, the ΔF_{cav} components increase with distance at small distances, $E_{el} + E_{pol}^{GB}$ decreases with distance for like-charged pairs and increases for the oppositely charged NH_4^+...Cl^- pair, E_{pol} decreases with distance (Figure 4b−f), and all components tend to zero at large distances. A maximum is present in ΔF_{cav} for all pairs except that of the NH_4^+...NH_4^+ pair (Figure 4e); however, in this case, the first maximum in the simulated PMF is weak (Figure 3e, dashed line), which suggests that the contribution from different ordering of water molecules at the interface of the hydration shells of the monomers (and creating a small cavity in the region of the desolvation maximum) plays a minor role. This might, in turn, be caused by the fact that water makes well-defined hydrogen bonds with the ammonium

TABLE 1: Potential Parameters Calculated by Minimization of the Target Function Defined by eq 19[a]

system[b]	α_{ij} [kcal/mol Å^4]	ϵ_0 [kcal/mol]	σ_0 [Å]	γ [kcal mol^{-1} Å^{-2}]	R_{s1} [Å]	R_{s2} [Å]	r [Å]
m−m		0.273	1.802	0.066	2.018	2.018	1.557
m−Cl^-	0.0125[c]	0.042	2.090	0.049	2.052	2.014	1.400[d]
m−NH_4^+	0.0125[c]	0.058	1.815	0.036	1.894	1.945	1.400[d]
Cl^-...Cl^-		0.073	2.339	−0.162	1.667	1.667	1.768
NH_4^+...NH_4^+		11.436	0.875	−0.232	1.668	1.668	1.028
NH_4^+...Cl^-		0.006	2.123	0.053	1.413	2.254	1.400[c]

[a] The van der Waals energy was calculated on the basis of the Kihara expression (eq 5), and the molecular-surface area was expressed by eq 11. [b] m denotes a methane molecule. [c] This value was fixed during minimization. [d] Fixed at the value of the water radius during minimization.

TABLE 2: Potential Parameters Calculated by Minimization of the Function Defined by eq 19[a]

system[b]	α_{ij} [kcal/mol·Å^4]	ϵ_0 [kcal/mol]	σ_0 [Å]	σ_i [Å]	σ_j [Å]	R_i [Å]	R_j [Å]	$\alpha_{ij}^{(1)}$ [kcal/mol]	$\alpha_{ij}^{(2)}$ [kcal/mol]	$\alpha_{ij}^{(3)}$ [kcal/mol]	$\alpha_{ij}^{(4)}$
m−m		0.004	5.131	4.405	4.405			90.00	−39.90	50.55	0.013
m−Cl^-	45.979	0.065	4.278	3.231	4.193	3.467	2.545	1.48	0.08	1.23	0.898
m−NH_4^+	54.415	0.030	4.041	3.086	4.004	3.424	2.486	1.41	−0.10	1.10	1.008
Cl^-...Cl^-		0.502	4.113	4.781	4.781	2.468	2.468	89.04	−40.92	51.43	0.162
NH_4^+...NH_4^+		0.284	2.303	4.589	4.589	1.977	1.977	89.34	−40.62	51.13	0.422
NH_4^+...Cl^-		0.011	4.422	3.235	4.061	0.168	2.851	1.70	0.19	0.96	0.872

[a] The van der Waals energy was calculated on the basis of the Lennard-Jones 6−12 potential (eq 4), and the cavity component of the potential was expressed by eq 18. The parameters pertain to the best fitted expression. [b] m denotes a methane molecule.

cations, and consequently, the water molecules of the first hydration sphere remain well ordered regardless of whether they are in the hydration shell of the monomer or of the dimer.

Conclusions

We have proposed simple physics-based analytical forms to express the PMF for pairs of interacting spherical charged particles, for pairs composed of charged and nonpolar particles, and for pairs of nonpolar particles and applied them to express the PMF of pairs composed of the methane molecule, chloride ion, and ammonium ion. The formulas consist of Lennard-Jones, electrostatic, solvent-polarization, and cavity terms. To calculate the free energy of cavity creation, we implemented a simplified analytical expression derived in paper 1[24] of this series, which is based on a Gaussian-overlap model. The advantage of this expression over the use of molecular-surface or solvent-accessible surface area is that it can be extended to non-spherical particles (as outlined in our paper 1[24] of this series); also, $\Delta F_{cav,}$ defined by eq 18, has continuous derivatives. The proposed analytical expressions fit the PMFs of model systems in water calculated by umbrella-sampling molecular dynamics very well, yielding physically reasonable parameters, suggesting that they can be extended to more complex anisotropic interaction sites, as will be demonstrated in paper 3 of this series.[48]

Acknowledgment. This work was supported by grants from the U.S. National Institutes of Health (GM-14312), the U.S. National Science Foundation (MCB05-41633), the National Institutes of Health Fogarty International Center Grant TW007193, and the Polish Ministry of Science and Informatization (1 T09A 099 30). Mariusz Makowski was supported by a grant from the "Homing" program of the Foundation for Polish Science (FNP). This research was conducted by using the resources of (a) our 818-processor Beowulf cluster at the Baker Laboratory of Chemistry and Chemical Biology, Cornell University, (b) the National Science Foundation Terascale Computing System at the Pittsburgh Supercomputer Center, (c) our 45-processor Beowulf cluster at the Faculty of Chemistry, University of Gdańsk, (d) the Informatics Center of the Metropolitan Academic Network (IC MAN) in Gdańsk, and (e) the Interdisciplinary Center of Mathematical and Computer Modeling (ICM) at the University of Warsaw.

References and Notes

(1) Liwo, A.; Pincus, M. R.; Wawak, R. J.; Rackovsky, S.; Scheraga, H. A. *Protein Sci.* **1993**, *2*, 1697.
(2) Liwo, A.; Ołdziej, S.; Pincus, M. R.; Wawak, R. J.; Rackovsky, S.; Scheraga, H. A. *J. Comput. Chem.* **1997**, *18*, 849.
(3) Liwo, A.; Czaplewski, C.; Pillardy, J.; Scheraga H. A. *J. Chem. Phys.* **2001**, *115*, 2324.
(4) Liwo, A.; Ołdziej, S.; Czaplewski, C.; Kozłowska, U.; Scheraga, H. A. *J. Phys. Chem. B* **2004**, *108*, 9421.
(5) Czaplewski, C.; Liwo, A.; Ołdziej, S.; Scheraga, H. A. *Polymer* **2004**, *45*, 677.
(6) Ołdziej, S.; Łagiewka, J.; Liwo, A.; Czaplewski, C.; Chinchio, M.; Nanias, M.; Scheraga, H. A. *J. Phys. Chem. B* **2004**, *108*, 16950.
(7) Scheraga, H. A.; Liwo, A.; Ołdziej, S.; Czaplewski, C.; Pillardy, J.; Ripoll, D. R.; Vila, J. A.; Kaźmierkiewicz, R.; Saunders, J. A.; Arnautova, Y. A.; Jagielska, A.; Chinchio, M.; Nanias, M. *Front. Biosci.* **2004**, *9*, 3296.
(8) Liwo, A.; Khalili, M.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2362.
(9) Kubo, R. *J. Phys. Soc. Jpn.* **1962**, *17*, 1100.
(10) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
(11) Gay, J. G.; Berne, B. J. *J. Chem. Phys.* **1981**, *74*, 3316.
(12) Czaplewski, C.; Rodziewicz-Motowidło, S.; Liwo, A.; Ripoll, D. R.; Wawak, R. J.; Scheraga, H. A. *Protein Sci.* **2000**, *9*, 1235.
(13) Czaplewski, C.; Ripoll, D. R.; Liwo, A.; Rodziewicz-Motowidło, S.; Wawak, R. J.; Scheraga, H. A. *Int. J. Quantum Chem.* **2002**, *88*, 41.
(14) Czaplewski, C.; Rodziewicz-Motowidło, S.; Dąbal, M.; Liwo, A.; Ripoll, D. R.; Scheraga, H. A. *Biophys. Chem.* **2003**, *105*, 339.
(15) Maksimiak, K.; Rodziewicz-Motowidło, S.; Czaplewski, C.; Liwo, A.; Scheraga, H. A. *J. Phys. Chem. B* **2003**, *107*, 13496.
(16) Czaplewski, C.; Liwo, A.; Ripoll, D. R.; Scheraga, H. A. *J. Phys. Chem. B* **2005**, *109*, 8108.
(17) Born, M. *Z. Phys.* **1920**, *1*, 45.
(18) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T, *J. Am. Chem. Soc.* **1990**, *112*, 6127.
(19) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 19824.
(20) Bashford, D.; Case, D. A. *Annu. Rev. Phys. Chem.* **2000**, *51*, 129.
(21) Wojciechowski, M.; Lesyng, B. *J. Phys. Chem. B* **2004**, *108*, 18368.
(22) Cited in: Margenau, H.; Kestner, N. R. *Theory of Intermolecular Forces*, 1st ed.; Pergamon Press: Oxford, U.K., 1969; p 107.
(23) Vorobjev, Y. N.; Grant, J. A. and Scheraga, H. A. *J. Am. Chem. Soc.* **1992**, *114*, 3189.
(24) Makowski, M.; Liwo, A.; Scheraga, H. A. *J. Phys. Chem. B* **2007**, *111*, 2910.
(25) Rank, J. A.; Baker, D. *Protein Sci.* **1997**, *6*, 347.

(26) Wawak, R. J.; Gibson, K. D.; Scheraga, H. A. *J. Math. Chem*. **1994**, *15*, 207.

(27) Gibson, K. D.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 420.

(28) Kang, Y. K.; Némethy, G.; Scheraga, H. A. *J. Chem. Phys* **1987**, *91*, 4105.

(29) Kang, Y. K.; Némethy, G.; Scheraga, H. A. *J. Chem. Phys* **1987**, *91*, 4109.

(30) Kang, Y. K.; Némethy, G.; Scheraga, H. A. *J. Chem. Phys*. **1987**, *91*, 4118.

(31) Kang, Y. K.; Némethy, G.; Scheraga, H. A. *J. Phys. Chem*. **1988**, *92*, 4739.

(32) Stouten, P. F. W.; Frömmel, C.; Nakamura, H.; Sander, C. *Mol. Simul*. **1993**, *102*, 97.

(33) Augspurger, J. D.; Scheraga, H. A. *J. Comput. Chem*. **1996**, *13*, 1549.

(34) Berne, B. J.; Pechukas, P. *J. Chem. Phys*. **1972**, *56*, 4213.

(35) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys*. **1983**, *79*, 926.

(36) Makowska, J.; Makowski, M.; Giełdoń, A.; Liwo, A.; Chmurzyński, L. *J. Phys. Chem. B* **2004**, *108*, 12222.

(37) Makowska, J.; Makowski, M.; Liwo, A.; Chmurzyński, L. *J. Comput. Chem*. **2004**, *26*, 235.

(38) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem*. **1992**, *13*, 1011.

(39) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem*. **1995**, *16*, 1339.

(40) Czaplewski, C.; Kalinowski, S.; Liwo, A.; Scheraga, H. A. *Mol. Phys.* **2005**, 3157.

(41) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. A. *Comput. Phys. Commun*. **1995**, *91*, 1.

(42) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Chaetham, T. E., III; Wang, J.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M; Tsui, V.; Gohlke, H; Radmer, R. J.; Duan, Y; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 7*; University of California, San Francisco, 2002.

(43) Wassenaar, T. A.; Mark, A. E. *J. Comput. Chem.* **2006**, *27*, 316.

(44) Jorgensen, W. L.; Madura, J. D.; Swenson C. J. *J Am Chem Soc*. **1984**, *106*, 6638.

(45) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.

(46) Marquardt, D. W. *J. Soc. Ind. Appl. Math*. **1963**, *11*, 431.

(47) Compostizo, A.; Cancho, S. M.; Rubio, R. G.; Colin, A. C. *Phys. Chem. Chem. Phys.* **2001**, *3*, 1861.

(48) Makowski, M.; Sobolewski, E.; Liwo, A.; Czaplewski, C.; Ołdziej, S.; Scheraga, H. A. *J. Phys. Chem. B* **2007**, *111*, 2925.