Published in final edited form as:

J Phys Chem B. 2011 May 19; 115(19): 6119–6129. doi:10.1021/jp111258p.

Simple physics-based analytical formulas for the potentials of mean force of the interaction of amino-acid side chains in water. V. Like-charged side chains

Mariusz Makowski^{1,2,*}, Adam Liwo^{1,2}, Emil Sobolewski³, and Harold A. Scheraga²

¹Faculty of Chemistry, University of Gdańsk, Sobieskiego 18, 80-952 Gdańsk, Poland ²Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853-1301

³Laboratory of Biopolymer Structure, Intercollegiate Faculty of Biotechnology, University of Gdańsk, Kładki 24, 80-822 Gdańsk, Poland.

Abstract

A new model of side-chain – side-chain interactions for charged side-chains of amino acids, to be used in the UNRES force-field, has been developed, in which a side chain consists of a nonpolar and a charged site. The interaction energy between the nonpolar sites is composed of a Gay-Berne and a cavity term; the interaction energy between the charged sites consists of a Lennard-Jones term, a Coulombic term, a Generalized-Born term, and a cavity term, while the interaction energy between the nonpolar and charged sites is composed of a Gay-Berne and a polarization term. We parameterized the energy function for the models of all six pairs of natural like-charged amino-acid side chains, namely propionate-propionate (for the aspartic acid-aspartic acid pair), butyrate-butyrate (for the glutamic acid-glutamic acid pair), propionate-butyrate (for the aspartic acid-glutamic acid pair), pentylamine cation-pentylamine cation (for the lysine-lysine pair), 1-butylguanidine cation-1-butylguanidine cation (for the arginine-arginine pair), and pentylamine cation-1-butylguanidine cation (for the lysine-arginine pair). By using umbrella-sampling molecular dynamics simulations in explicit TIP3P water, we determined the potentials of mean force of the above-mentioned pairs as functions of distance and orientation and fitted analytical expressions to them. The positions and depths of the contact minima and the positions and heights of the desolvation maxima, including their dependence on the orientation of the molecules were well represented by analytical expressions for all systems. The values of the parameters of all the energy components are physically reasonable, which justifies use of such potentials in coarse-grain protein-folding simulations.

Keywords

charged side chains; new model of side-chain – side-chain interactions; potential of mean force; molecular dynamics; umbrella sampling

Introduction

For several years, we have been developing the coarse-grained UNRES model and force field for simulations of protein structure and dynamics.^{1–17} In the UNRES model, a polypeptide chain is represented by its α -carbon trace with united peptide groups (p) located halfway between two consecutive C $^{\alpha}$ atoms and united side chains (SC) attached to

*Corresponding author: momo@chem.univ.gda.pl.

them.^{1,2,4} The UNRES force field is a physics-based force field which is defined⁴ as a restricted energy function (RFE) or potential of mean force (PMF) of polypeptide chains immersed in water where all degrees of freedom except the positions of C ^{α} atoms and SC centers are averaged out. This formulation is a direct implementation of the Boltzmann law because the PMF determines the probability of a coarse-grained conformation to occur. The PMF is further factorized into contributions from single interactions, as well as pair, triads, etc. of interactions within and between coarse-grained sites. The factors can clearly be identified with the terms in the effective energy function that correspond to particular interactions,⁵ as opposed to the effective energy terms of the so-called statistical potentials derived by Boltzmann inversion of the distribution and correlation functions determined from structural databases.^{18–20}

In order to make the approach practical, the expansion is cut at the fourth order, which appears sufficient to represent protein structure and thermodynamics.⁷ Each factor can then be derived and parameterized by considering only the usually small subsystem to which it belongs; this has enabled us to derive essential terms corresponding to the coupling of the local and backbone-electrostatic interactions based on a generalized-cumulant expansion.⁵ and parameterize the local-interaction,^{5,8,13} backbone-electrostatic-interaction^{1,3} and correlation^{3,6} terms based on potential-energy surfaces calculated by ab initio molecular quantum mechanics. Further, the factors are scaled by temperature functions to account for the temperature dependence of the PMF.^{9,10} The terms which remain to be determined based on factor expansion are the mean-field potentials of solvent-mediated interactions between side chains.

In the present UNRES force field,^{1–17} the side chain – side chain interaction ($U_{SC_iSC_j}$) potentials were assigned Gay-Berne²¹ functional forms that take anisotropy of interactions into account, and their parameters were determined² by fitting to correlation functions and to side-chain-contact energies determined from the Protein Data Bank (PDB).²² These potentials are related to the physics of solvent-mediated interactions between side chains because the average free energies of side-chain side-chain contacts calculated from them are very well correlated with the free energies of transfer of amino-acid side chains from water to n-octanol.² The Gay-Berne-type potential assumes that the interacting sites are ellipsoids of revolution (also termed spheroids), with a single interaction site per side chain. This model is illustrated in Figure 1. However, the charged side chains contain both a charged “head” and a nonpolar “tail”; i.e., they are amphiphilic. The incorrect functional form, previously used² for $U_{SC_iSC_j}$, involving amphipolar side chains probably impairs the predictive power of the UNRES force field.^{1–17}

In our previous papers of this series,^{14–17} we presented the results of our work on the development of the new $U_{SC_iSC_j}$ potential which will replace the present knowledge-based potentials² used in UNRES.^{1–17} Previously we derived¹⁴ a simple analytical expression for the cavity terms of the potentials of mean force of nonpolar solutes in water, based on an analysis of the number and context of water molecules in different parts of the solvation sphere, and generalized it to spheroidal solute particles. Later,¹⁵ we showed that the derived expression reproduces the features of the cavity part of the potentials of mean force of simple spherical solutes in water very well. In our next two papers^{16,17} we considered models of homo- and heterodimers of nonpolar amino-acid side chains.

In this work and in the accompanying paper,²³ we present the potentials of mean force of pairs of models of all natural charged amino acids, namely propionate (Pr p[–]) butyrate (But[–]), pentylamine cation (Pna⁺), and N-butylguanidine cation (BtG⁺), which model the anionic bases of aspartic acid and glutamic acid or the cationic acids of lysine and arginine, respectively, as functions of distances between the molecules and their relative orientation

by means of umbrella-sampling molecular dynamics (MD) simulations; these potentials were subsequently fitted to the PMFs with analytical expressions. In each side-chain model, the C $^{\alpha}$ atom is considered to be part of a side chain, as in the UNRES model.¹⁻¹⁷

Several studies have been carried out to determine the PMFs of pairs of simple like-charged ions, such as lithium, sodium, potassium, and chloride as well as pairs of simple oppositely charged ions.²⁴⁻³⁷ Although, the results depend quantitatively on the approach and the water model used, they are qualitatively similar. The PMFs exhibit one (for like-charged ions) or two (for oppositely charged ions) minima as functions of the distance between the centers of the ions. One minimum in the PMF plots of like-charged and two minima for oppositely charged ions were explained in terms of locally-increased concentration of solvent molecules and counter-ions.^{30,31,34,35,38} The appearance of a minimum in the PMF curve for like-charged ions is manifested in PDB structures as a considerable number of arginine - arginine side-chain contacts.²

Several studies have also been carried out on the PMFs of systems modeling like-charged side chains of proteins,³⁸⁻⁴² such as, e.g., guanidinium cations^{38,39} (modeling the side chains of arginine), between side chains of two lysine residues⁴¹ and a pair of acetate - guanidine^{38,40} ions in water to model interactions in proteins of the aspartic or glutamic acids with the arginine side-chain. The PMFs of systems modeling the charged-nonpolar side chains were also determined.^{38,40} In these studies,³⁸⁻⁴² the PMFs depended only on the distance, i.e., they were averaged over all orientations. Masunov and Lazaridis³⁶ calculated the dependence of the PMFs of side-chain models on both distance and orientations. They performed a series of molecular dynamics simulations for the particular orientations as functions of the distance between the centers of interaction. However, these authors considered only selected orientations. Also, the depths of the PMF minima were probably overestimated in their study because they used an incorrect model of electrostatic interactions with solvent.

Here, we present a new side-chain model which assumes two centers per side chain: a nonpolar one placed in the middle of a side chain (represented by an ellipsoid of revolution) and a charged one placed in the “head” part of a side chain. In this work we present results for six pairs of like-charged solutes, i.e. Pr p⁻ ... Pr p⁻, But⁻ ... But⁻, Pr p⁻ ... But⁻, Pna⁺ ... Pna⁺, BtG⁺ ... BtG⁺, and Pna⁺ ... BtG⁺. These pairs model the side-chain – side-chain interactions of Asp...Asp, Glu...Glu, Asp...Glu, Lys-Lys, Arg...Arg, and Lys...Arg, pairs respectively. In our accompanying paper²³, we treat pairs of oppositely charged side chains, i.e. Asp...Lys, Glu...Lys, Asp...Arg and Glu...Arg.

Theory

The new model of interactions between charged side-chains is illustrated in Figure 2. Each such side chain is represented by two interaction sites: a nonpolar site (represented by an ellipsoid of revolution in Figure 2) and a charged/polar site located in the “head” part of the side chain (represented by shaded spheres in Figure 2). The geometric center of the side chain is located between the center of the nonpolar and that of the charged site (Figure 2).

The effective interaction (potential of mean force) of a pair of charged solutes in water, W_{cc} , is approximated by a sum of the energy of electrostatic interactions between the charged or polar parts, and the interaction energy between charged/polar and nonpolar sites, the interaction energy between the nonpolar sites, and a hydrophobic association term derived in our previous paper,¹ as expressed by eq 1.

$$W_{cc} = E_{GBerne} + E_{el} + E_{pol}^{GB} + E_{pol} + \Delta F_{cav}^{iso} + \Delta F_{cav} + E_{LJ} \quad (1)$$

where E_{GBerne} is the van der Waals term corresponding to the interactions between the nonpolar sites, E_{el} is the Coulombic term for the interaction between the charged sites, E_{pol} is the polarization energy coming from the interactions between the charged and nonpolar sites (but not from their interactions with the solvent), E_{pol}^{GB} is the energy contribution due to solvent polarization by the charged sites; it is computed by using the generalized Born (GB) model,^{34–38} ΔF_{cav}^{iso} is the cavity term corresponding to the charged sites (the superscript “iso” indicates that the sites involved are isotropic), ΔF_{cav} is the cavity term corresponding to the nonpolar sites, which is the difference between the cavity contribution to the free energy of hydration of the nonpolar sites of the dimer and those of the isolated monomers, and the isotropic Lennard-Jones potential (E_{LJ}) describes the van der Waals interactions between two charged headgroups.

For the E_{GBerne} energy term, we use the Gay-Berne-type potential¹³ expressed by eq 2. It should be noted that, previously, we used eq 2 to express the complete side chain-side chain interaction potential in UNRES,⁵ but now we consider more terms (eq 1).

$$E_{GBerne} = 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}^0}{r_{ij} - \sigma_{ij} + \sigma_{ij}^0} \right)^{12} - \left(\frac{\sigma_{ij}^0}{r_{ij} - \sigma_{ij} + \sigma_{ij}^0} \right)^6 \right] \quad (2)$$

where r_{ij} is the distance between the centers of the side chains, σ_{ij} is the distance corresponding to the zero value of E_{GBerne} for arbitrary orientation of the particles (σ_{ij}^0 is the distance corresponding to the zero value of E_{GBerne} for the side-to-side approach of the particles), ε_{ij} (depending on the relative orientation of the particles) is the van der Waals well depth. The dependence of ε_{ij} and σ_{ij} on the orientation of the particles is given by eqs 3–5 and eq 6, respectively.^{3–5}

$$\varepsilon_{ij} \equiv \varepsilon(\omega_{ij}^{(1)}, \omega_{ij}^{(2)}, \omega_{ij}^{(12)}) = \varepsilon_{ij}^0 \varepsilon_{ij}^{(1)} \varepsilon_{ij}^{(2)} \quad (3)$$

$$\varepsilon_{ij}^{(1)} = \left[1 - \chi_{ij}^{(1)} \chi_{ij}^{(2)} \omega_{ij}^{(12)2} \right]^{-1/2} \quad (4)$$

$$\varepsilon_{ij}^{(2)} = \left[1 - \frac{\chi_{ij}^{(1)} \omega_{ij}^{(1)2} + \chi_{ij}^{(2)} \omega_{ij}^{(2)2} - 2\chi_{ij}^{(1)} \chi_{ij}^{(2)} \omega_{ij}^{(1)} \omega_{ij}^{(2)} \omega_{ij}^{(12)}}{1 - \chi_{ij}^{(1)} \chi_{ij}^{(2)} \omega_{ij}^{(12)2}} \right]^2 \quad (5)$$

$$\sigma_{ij} = \sigma_{ij}^0 \left[1 - \frac{\chi_{ij}^{(1)} \omega_{ij}^{(1)2} + \chi_{ij}^{(2)} \omega_{ij}^{(2)2} - 2\chi_{ij}^{(1)} \chi_{ij}^{(2)} \omega_{ij}^{(1)} \omega_{ij}^{(2)} \omega_{ij}^{(12)}}{1 - \chi_{ij}^{(1)} \chi_{ij}^{(2)} \omega_{ij}^{(12)2}} \right] \quad (6)$$

with

$$\omega_{ij}^{(1)} = \hat{\mathbf{u}}_{ij}^{(1)} \cdot \hat{\mathbf{r}}_{ij} = \cos \theta_{ij}^{(1)} \quad (7)$$

$$\omega_{ij}^{(2)} = \hat{\mathbf{u}}_{ij}^{(2)} \cdot \hat{\mathbf{r}}_{ij} = \cos \theta_{ij}^{(2)} \quad (8)$$

$$\omega_{ij}^{(12)} = \hat{\mathbf{u}}_{ij}^{(1)} \cdot \hat{\mathbf{u}}_{ij}^{(2)} = \cos \theta_{ij}^{(1)} \cos \theta_{ij}^{(2)} + \sin \theta_{ij}^{(1)} \sin \theta_{ij}^{(2)} \cos \phi_{ij} \quad (9)$$

where $\hat{\mathbf{u}}_{ij}^{(1)}$ and $\hat{\mathbf{u}}_{ij}^{(2)}$ are unit vectors along the principal axes of the interacting sites (identified in this work with the C^α-SC axes), $\hat{\mathbf{r}}_{ij}$ is the unit vector pointing from the center of site *i* to that of site *j*, r_{ij} is the distance between the side-chain centers (Figs. 1–2), the parameters $\chi_{ij}^{(1)}$ and $\chi_{ij}^{(2)}$ are the anisotropies of the van der Waals distance, the parameters $\chi_{ij}^{'(1)}$ and $\chi_{ij}^{'(2)}$ are the anisotropies of the van der Waals well depth and the parameter ε_{ij}^0 is the well-depth corresponding to the side-to-side orientation of the interacting particles. In this work, as previously,^{2–4} the parameters mentioned above were determined by least-squares fitting of the analytical expression for the free energy of two side chains interacting in water (eq 1) to the potentials of mean force determined from MD simulations.

The Coulombic term is given by eq 10.

$$E_{el} = 332 \left(\frac{q_i q_j}{r'_{ij}} \right) \quad (10)$$

where q_i and q_j are the charges of sites *i* and *j*, respectively, r'_{ij} is the distance between the centers of the charged sites of side chains SC_{*i*} and SC_{*j*} (see Figure 2), and the coefficient 332 is introduced to express the energies in kcal/mol.

The components E_{pol}^{GB} and ΔF_{cav}^{iso} correspond to the “bulk dielectric” solvent-polarization and solvent-restructuring component, respectively. For the “bulk dielectric” solvent-polarization part involving a pair of charged particles, we adopt the expression from the Generalized Born model.^{34–38}

$$E_{pol}^{GB} = 332 q_i q_j \left(\frac{1}{\varepsilon_{in}} - \frac{1}{\varepsilon_{out}} \right) \frac{1}{f_{GB}(r'_{ij})} \quad (11)$$

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 12.

$$f_{GB}(r'_{ij}) = \sqrt{r_{ij}^{'2} + a_i a_j} \exp \left(-\frac{r_{ij}^{'2}}{4 a_i a_j} \right) \quad (12)$$

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

The Generalized Born model is an analytical approximation to the solvent-polarization energy calculated based on the solution of the Poisson equation (or linearized Poisson-Boltzmann equation) for the solute-solvent system;^{46–49} this model has recently been extended to systems containing multipoles.⁴⁹ It can also be considered as an interpolation of the solvent-polarization energy between two charges at zero and infinite separation immersed in a solvent; for these boundary configurations, the solution of the Poisson equation is given by the Born model.^{43,46,49} Solving the Poisson-Boltzmann equation is too expensive in molecular simulations and is applied only when performing single-point energy evaluations. The solvation energies computed with the Generalized Born model are in good agreement both with those computed by solving the Poisson equation⁴⁸ and with experimental values.^{44–48} Moreover, the pK_a values computed with the Generalized Born model are consistent with experiment.⁴⁷ On the other hand, it must be noted that the Poisson equation and, consequently, the Generalized Born model accounts for the only linear response of solvent polarization to the electric field of the solute.^{50,51} However, because the basic/acidic side chains bear the charge of ± 1 and the charge is distributed on the entire side-chain headgroup, the ionic potential of a charged side-chain group is not high; therefore, the non-linearity effects that are ignored by the Born model and the parent Poisson equation are not significant, as proved by the good agreement between the calculated and experimental solvation energies^{44–48} and pK_a values.⁴⁷ The nonlinearity, however, becomes remarkable in systems that include nucleic acids and multivalent metal ions.⁵¹

More rigorous models were recently proposed in the literature to account for the “bulk solvent” polarization in biomolecular systems,^{50,51} which assume that the solvent is represented by Langevin dipoles. This approach results in the dependence of the dielectric permittivity on the distance between the charged sites;⁵⁰ however, finding this distance dependence involves the solution of a transcendental equation or numerical integration and, therefore, seems too expensive to implement in molecular simulations.

We propose eq 13 to express the polarization component involving the interaction between charged and nonpolar particles (see Figure 2).

$$E_{pol} = \alpha_{ij}^{pol} \left(\frac{I}{f_{GB}(r_{ij}'')} \right)^4 + \alpha_{ji}^{pol} \left(\frac{I}{f_{GB}(r_{ji}'')} \right)^4 \quad (13)$$

where α_{ij}^{pol} and α_{ji}^{pol} are related to the polarizability of the nonpolar parts of side chain i and side chain j , respectively. Previously², we demonstrated that, at large distances, this contribution of polarization energy varies as $1/r^4$. The rationale for expression 13 is that a nonpolar particle replaces the solvent at distance r , 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 $1/r^4$.

The expression for isotropic cavity creation (or solvent-restructuring) term ΔF_{cav}^{iso} was derived previously¹ based on the Gaussian overlap model and is expressed by eq. 14. This term enables us to differentiate between head-to-head and tail-to-tail orientations of side chains in our analytical PMF curves. In the original formulation of the Generalized-Born model,^{44,46} this term is proportional to molecular surface area, hence the complete name of Generalized Born Surface Area (GBSA) which is, however more difficult to compute less numerically stable compared to eq 14.

$$\Delta F_{cav}^{iso} = \frac{\alpha_{ij}^{iso(1)} [(x)^{\frac{1}{2}} + \alpha_{ij}^{iso(2)} x - \alpha_{ij}^{iso(3)}]}{1 + \alpha_{ij}^{iso(4)} \cdot x^{1/2}} \quad (14)$$

with

$$x = \frac{r'_{ij}}{\sqrt{(\sigma_i^{iso})^2 + (\sigma_j^{iso})^2}} \quad (15)$$

where r'_{ij} is the distance between two charged parts (see Figure 2) of particles i and j , σ_i^{iso} and σ_j^{iso} can be identified with the minimum distance between the center of charge of particle i or j , respectively. The parameters $\alpha_{ij}^{iso(1)}$, $\alpha_{ij}^{iso(2)}$, $\alpha_{ij}^{iso(3)}$, $\alpha_{ij}^{iso(4)}$, σ_i^{iso} , and σ_j^{iso} are determined by least-squares fitting of the analytical expression for the free energy of two side chains interacting in water (eq 1) to the potentials of mean force determined from MD simulations.

The Gaussian-overlap model assumes that the solvent and, for charged systems, also the counter-ion density around a solute has the form of a Gaussian differential; the free energy of solvation due to solvent restructuring is computed by integrating the overlap of this density with the density of the solute (assumed to be Gaussian) and that of the solvation sphere of another solvent molecule; these contributions are assigned different weights. The model reproduces the desolvation maxima.^{14,15} The Gaussian overlap model of the solvent and counter-ion density in the first solvation sphere is supported by the spatial distribution of water molecules and counter-ions obtained from simulation of model systems, where counter-ions and water density around a solute have clear maxima in the center of the first solvation sphere (Figures 9 and 10 in ref 38).

More sophisticated models were developed recently to account for the solvent-restructuring term in systems containing charges immersed in solvent;^{34,35,50} however, they involve numerical integration and seem, therefore, to be too expensive for molecular simulations (see, e.g., eq 18 in ref 50) or numerical solution of integral equations.^{34,35}

The expression for ΔF_{cav} of spheroidal particles was derived in paper I¹ of our work on the side-chain - side-chain interaction potential in the UNRES force field and is given by eq 16. This term accounts for the free-energy contribution due to restructuring water molecules around a hydrophobic dimer and has been derived and discussed in detail in ref 14.

$$\Delta F_{cav} = \frac{\alpha_{ij}^{(1)} [(x \cdot \lambda)^{\frac{1}{2}} + \alpha_{ij}^{(2)} x \cdot \lambda - \alpha_{ij}^{(3)}]}{1 + \alpha_{ij}^{(4)} (x \cdot \lambda)^{1/2}} \quad (16)$$

with

$$x = \frac{r_{ij}}{\sqrt{\sigma_i^2 + \sigma_j^2}} \quad (17)$$

$$\lambda = \left[1 - \frac{\chi_{ij}^{(1)} \omega_{ij}^{(1)2} + \chi_{ij}^{(2)} \omega_{ij}^{(2)2} - 2\chi_{ij}^{(1)} \chi_{ij}^{(2)} \omega_{ij}^{(1)} \omega_{ij}^{(2)} \omega_{ij}^{(12)}}{1 - \chi_{ij}^{(1)} \chi_{ij}^{(2)} \omega_{ij}^{(12)2}} \right]^2 \quad (18)$$

where the symbols $\omega_{ij}^{(1)}$, $\omega_{ij}^{(2)}$, and $\omega_{ij}^{(12)}$ are defined by eqs 7–9, r_{ij} is the distance between the centers of the particles, $\chi_{ij}^{(1)}$ and $\chi_{ij}^{(2)}$ are anisotropies pertaining to ΔF_{cav} , and σ_i and σ_j can be identified with the minimum distance between the center of particle i or j , respectively. The parameters $\alpha_{ij}^{(1)}$, $\alpha_{ij}^{(2)}$, $\alpha_{ij}^{(3)}$, and $\alpha_{ij}^{(4)}$, σ_i and σ_j and the anisotropies are determined by least-squares fitting of the analytical expression for the free energy of two side chains interacting in water (eq 1) to the potentials of mean force determined from MD simulations.

The isotropic Lennard-Jones potential (E_{LJ}) describing the van der Waals interactions between two charged headgroups is expressed by eq 19:

$$E_{LJ} = 4 \cdot \epsilon'_{ij} \cdot \left[\left(\frac{\sigma'_{ij}}{r'_{ij}} \right)^{12} - \left(\frac{\sigma'_{ij}}{r'_{ij}} \right)^6 \right] \quad (19)$$

where r'_{ij} is the distance between the centers of the charged headgroups, σ'_{ij} is the distance corresponding to the zero value of E_{LJ} , and ϵ'_{ij} is the van der Waals well depth.

It should be noted that the analytical expressions used here for E_{pol}^{GB} , ΔF_{cav} , and ΔF_{cav}^{iso} are not as rigorous as in more refined models proposed recently.^{50,51} However, they have counterparts in the free-energy terms obtained from more refined models, e.g., that proposed recently by Hassan⁵⁰ and, while simple, capture the physics of solvent-solute interactions, as explained earlier in this section. The simplifications are necessary for developing a force field usable in large size- and time-scale molecular simulations. More refined approaches involve numerical solution of the Poisson equation,^{52,53} numerical solution of integral equations,^{34,35} or numerical integration and numerical solution of transcendental equations;^{50,51} only single-point energy evaluation is practical when using them.

Methods

Molecular dynamics simulations were carried out with the AMBER⁵⁴ suite of programs, using the AMBER 9.0 force field.⁵⁵ Each system was placed in a periodic box containing explicit water molecules in an amount corresponding to the experimental water density at 298 K. The box dimensions as well as number of water molecules for the pair studied are collected in Table 1. The systems shown in Table 1 were used to model the effective interactions between the like-charged side chains in peptides and proteins. The TIP3P water model⁵⁶ was used in the molecular dynamics (MD) simulations.

The charges on the atoms of the solute molecules, needed for the AMBER 9.0 force field, were determined for each side-chain model by using a standard procedure,⁵⁷ i.e., by fitting the point-charge electrostatic potential to the molecular electrostatic potential computed using the electronic wave function calculated at the restricted Hartree-Fock (RHF) level with the 6–31 G* basis set. The program GAMESS⁵⁸ was used to carry out quantum-mechanical

calculations, while the program RESP⁴² of the AMBER 9.0 package was used to compute the fitted charges. The charges and the AMBER atom types are shown in Figure 3. Because the systems studied bear a net charge, chloride or sodium counter ions, respectively, were added to neutralize a system.

MD simulations were carried out in two steps. In the first step; each system was equilibrated in the NPT ensemble (constant number of particles, pressure, and temperature) at 298 K for 100 ps (picoseconds), and the integration step was 2 fs (femtoseconds). After equilibration, MD simulations were run in the NVT ensemble (constant number of particles, volume, and temperature) for 10 ns, and the integration step was 2 fs. A 9 Å cut-off for all nonbonded interactions, including electrostatic interactions, was imposed. For each system, a series of 10 windows of 10 ns simulations was run with different harmonic-restraint potentials imposed on the distance between the atoms closest to the center of the mass of each of the molecules, as given by eq 20.

$$V = \frac{1}{2}k(r - r_i^o)^2 \quad (20)$$

where k is the force constant [assumed to be 2 kcal/(mol $\times$ Å²)], r is the inter-particle distance, and r_i^o is the center of the restraint in the i th simulation (window); the values of r^o were 4, 5, 6, 7, 8, 9, 10, 11, 12, and 13 Å. A total number of 50,000 configurations were collected for each window.

To determine the potentials of mean force (PMFs) of the systems studied, we processed the results of all restrained MD simulations for each system by using the Weighted Histogram Analysis Method (WHAM).^{59,60} For a given system, four-dimensional histograms in r_{ij} ,

$\theta_{ij}^{(1)}$, $\theta_{ij}^{(2)}$, and ϕ_{ij} (Figure 1) were constructed. The ranges and bin sizes were 4.0 Å $\leq r_{ij} \leq$ 13 Å with one bin side distance 0.2 Å, with angles $0^\circ \leq \theta_{ij}^{(1)} \leq 180^\circ$ with bin side 60° , with angles $0^\circ \leq \theta_{ij}^{(2)} \leq 180^\circ$ with bin side 60° , and with angles $-180^\circ \leq \phi_{ij} < 180^\circ$ with bin side 60° . Consequently, each distance corresponded to 54 bins, each containing counts from different orientations of the molecules.

Fitting of analytical formulas to the PMFs was accomplished by minimizing the sum of the squares of the differences between the PMF values computed from the analytical formulas and determined from MD simulations, (Φ) defined by eq 21 by using the Marquardt method.⁶¹

$$\min_{\mathbf{x}} \Phi(\mathbf{y}) = \sum_i w_i \left[W^{MD}(r_i, \theta_{ij}^{(1)}, \theta_{ij}^{(2)}, \phi_{ij}) - W^{anal}(r_i, \theta_{ij}^{(1)}, \theta_{ij}^{(2)}, \phi_{ij}; \mathbf{y}) \right]^2 \quad (21)$$

$$w_i = \exp \left[- \frac{W^{MD}(r_i, \theta_{ij}^{(1)}, \theta_{ij}^{(2)}, \phi_{ij}) - W_{min}}{RT} \right] \quad (22)$$

where $W^{MD}(r_i, \theta_{ij}^{(1)}, \theta_{ij}^{(2)}, \phi_{ij})$ the PMF value determined by simulations for distance r_{ij} and orientation $(\theta_{ij}^{(1)}, \theta_{ij}^{(2)}, \phi_{ij})$; $W^{anal}(r_i, \theta_{ij}^{(1)}, \theta_{ij}^{(2)}, \phi_{ij}; \mathbf{y})$ the analytical approximation to the PMF at distance r_{ij} calculated with parameters given by the vector $\mathbf{y}$, whose components are the

adjustable parameters of eqs (2–6) and (10–19), and w_i defined by eq 22 (in which W_{min} is the minimum PMF obtained in the simulations, R is the gas constant and $T = 298$ K is the absolute temperature) is the weight of the i th point. Weighting the data points by the Boltzmann factor (eq 22) gives greater importance to low-energy regions of the free-energy surface.

In Figure 4 we show a scheme for four selected orientations: “side-to-side” (parallel), “head-to-head” (linear), “head-to-side” (perpendicular), and “side-to-head” (perpendicular) for which the distance dependence of the PMF is discussed in the next section.

Results and Discussion

The PMFs of the pairs considered, together with curves computed with the fitted analytical expressions using eq 21, are plotted as functions of the distance between the centers of the molecules in Figure 5a–f for four selected orientations: “side-to-side” (parallel), “head-to-head” (linear), “head-to-side” (perpendicular), and “side-to-head” (perpendicular); these orientations are shown in Figure 4. It should be noted, however, that all orientations, not only those of Figure 4 selected for illustration, were considered in fitting with eqs. 21 and 22. The dashed: black, red, blue, and green lines correspond to PMFs determined for the side-to-side (Figure 4a), head-to-head (Figure 4b), head-to-side (Figure 4c), and side-to-head (Figure 4d), respectively, obtained by MD simulations. The solid lines correspond to the analytical approximation to the PMFs, with coefficients determined by least-squares fitting (eq 21) of the analytical expression to the PMF determined by MD simulations.

It can be seen for side-to-side orientations (black lines) that each plot possesses a deep contact minimum (occurring at the shortest distances), and a desolvation maximum separating the two minima. Except for the $\text{BtG}^+ \dots \text{BtG}^+$ pair, the contact minima either shift up in free energy or disappear for the head-to-head and the head-to-side orientations which results from a close approach of the charged head groups. The curves corresponding to the head-to-head and head-to-side orientations are not smoother than those corresponding to the side-to-side orientations because of poorer statistics in fitting with eqs. 21 and 22. The highest solvent-separated minima occurred for negatively charged pairs (Figures 5a–c) than for the remaining systems. This observation is in agreement with our earlier results on distance-dependent PMFs.^{15,38,62,63} The contact minima occur at the shortest distances for the side-to-side and at the longest distances for the head-to-head orientation. The positions of the desolvation maxima occur at different distances between the centers of the interacting particles, as do those of the contact minima; however, their heights remain constant within the simulation error. For this reason, the multiplicative factor $\alpha_{ij}^{(l)}$ in eq 16 does not depend on orientation, whereas the van der Waals well depth, and ϵ_{ij} , in the Gay-Berne potential (eqs 3–5) do.

It can be seen from Figures 5 that the contact minima in the side-to-side orientation are wider and deeper for the pairs of larger like-charged molecules: (Figures 5d–f) than for those of negatively charged molecules (Figures 5a–c). The existence of these minima for this orientation is caused by the fact that more water molecules are eliminated from the side-to-side orientation and the particles are in contact through their hydrophobic parts.

The PMFs of the $\text{BtG}^+ \dots \text{BtG}^+$ pair (Figure 5e) are very similar to those corresponding to pairs of nonpolar side chains with the depth of the minimum almost unchanged and only its position shifted with orientation, as opposed to those of the other pairs of like-charged side chains, particularly the $\text{Pna}^+ \dots \text{Pna}^+$ pair (where the guanidine group is replaced with a smaller ammonium group). The reason for this is the presence of the water or counter-ion bridges between guanidine groups as concluded by Magalhaes et al.⁶⁴ based on the analysis

of arginine-arginine contact pairs in protein crystal structures and in our earlier simulation studies of guanidine-cation pairs.³⁸

The results of fitting the PMFs, using the analytical expressions given by eq 1 with components defined by eqs 2, 11, 13, 14 and 16, respectively, are also shown in Figure 5 (solid lines). We present the fitted parameters of the expressions for E_{GBerne} (eq 2), E_{pol}^{GB} (eq 11), E_{pol} (eq 13), ΔF_{cav}^{iso} (eq 14), ΔF_{cav} (eq 16), and E_{LJ} (eq 19) PMF components in Table 2 together with the distances of $d_i^{(1)}$, $d_i^{(2)}$, $d_j^{(1)}$, and $d_j^{(2)}$ defined in Figure 2.

It can be seen from Figure 5a–f that the analytical expression fits side-to-side orientation very well for all pairs studied in this work. The best fit is observed for the $\text{Pr p}^- \dots \text{Pr p}^-$ pair (Figure 5a). For all systems the fits are satisfactory and reproduce the order of the PMF curves corresponding to different orientations. Somewhat poorer fits are obtained for head-to-head and head-to-tail orientations, where the number of simulation data to determine the PMF was the smallest (this number is proportional to $\sin \theta_{ij}^{(1)} \cdot \sin \theta_{ij}^{(2)}$) and, therefore, the quality of the data is worse in this region (as mentioned earlier in this section, the PMFs determined for these orientations are not smooth). Consequently, the poorer quality of the fits for the above-mentioned orientations seems to be caused by the poorer quality of data and not by the inadequacy of the approximate expression for the free energy.

To illustrate the dependence of different energy components on the distance and orientation, the components resulting from fitting the eq 1 to the PMF of $\text{Pna}^+ \dots \text{BtG}^+$ by using eq 21 are shown in Figure 6. It can be seen that the energy components vary with the distance as expected from the physics of the respective interactions, as was also observed earlier,^{15–17} i.e., the E_{GBerne} components possess minima (Figure 6a), $E_{el} + E_{pol}^{GB}$ and E_{pol} decrease with the distance (Figure 6b and 6c, respectively), and ΔF_{cav} reproduces the desolvation maximum (Figure 6e). All energy components tend to zero at large distances. All energy components shift to larger distances when starting from side-to-side orientation to head-to-head orientation. Using the parameters collected in Table 2, the energy components can be plotted for all remaining systems, which have not been shown in Figure 6.

Conclusions

In this work, we proposed a new coarse-grained model for charged side chains, in which a side chain is represented by a nonpolar site represented by an ellipsoid of revolution and a point charge located on the long axis of the ellipsoid, generally away from the center of the ellipsoid. These potentials will be used in the UNRES force-field. We parameterized the model for the six pairs of like-charged natural amino-acid side chains by fitting the analytical expression to the potential of mean force for these pairs determined from umbrella-sampling molecular dynamics simulations. The approximate analytical expression consists of the anisotropic Gay-Berne potential to represent van der Waals interactions, electrostatic potential, generalized Born, polarization energy term, isotropic cavity, and anisotropic cavity terms. We demonstrated that the analytical expression fits the PMFs of like-charged side chains reasonably well, suggesting that it is a good candidate for the physics-based mean-field potentials of side chain-side chain interactions in our mesoscopic UNRES force field^{1–17} for simulation of the structure and dynamics of proteins. This expression reproduces the amphiphilic character of charged side chains, as opposed to the simple single-center axially-symmetric Gay-Berne potentials used in the present UNRES for all side-chain pairs.² Consequently, incorporating the new potentials should improve the performance of the UNRES force-field.

Acknowledgments

This work was supported by grants from the Polish Ministry of Science and Higher Education (N N204 152836), from the U.S. National Institutes of Health (GM-14312), and the U.S. National Science Foundation (MCB05-41633). This research was conducted by using the resources of (a) our 880-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 Gdansk, (d) the Informatics Center of the Metropolitan Academic Network (IC MAN) in Gdansk, and (e) the Interdisciplinary Center of Mathematical and Computer Modeling (ICM) at the University of Warsaw.

References and Notes

1. Liwo A, Pincus MR, Wawak RJ, Rackovsky S, Scheraga HA. *Prot. Sci.* 1993; 2:1697.
2. Liwo A, Oldziej S, Pincus MR, Wawak RJ, Rackovsky S, Scheraga HA. *J. Comput. Chem.* 1997; 18:849.
3. Liwo A, Kaźmierkiewicz R, Czaplewski C, Groth M, Oldziej S, Wawak RJ, Rackovsky S, Pincus MR, Scheraga HA. *J. Comput. Chem.* 1998; 19:259.
4. Liwo A, Czaplewski C, Pillardy J, Scheraga HA. *J. Chem. Phys.* 2001; 115:2323.
5. Oldziej S, Kozłowska U, Liwo A, Scheraga HA. *J. Phys. Chem. A.* 2003; 107:8035.
6. Liwo A, Oldziej S, Czaplewski C, Kozłowska U, Scheraga HA. *J. Phys. Chem. B.* 2004; 108:9421.
7. Oldziej S, Łągiewka J, Liwo A, Czaplewski C, Chinchio M, Nanas M, Scheraga HA. *J. Phys. Chem. B.* 2004; 108:16950.
8. Kozłowska U, Liwo A, Scheraga HA. *J. Phys.: Condens. Matter.* 2007; 19:285203.
9. Liwo A, Khalili M, Czaplewski C, Kalinowski S, Oldziej S, Wachucik K, Scheraga HA. *J. Phys. Chem. B.* 2007; 111:260. [PubMed: 17201450]
10. Shen H, Liwo A, Scheraga HA. *J. Phys. Chem. B.* 2009; 113:8738. [PubMed: 19480420]
11. Liwo, A.; Czaplewski, C.; Oldziej, S.; Rojas, AV.; Kaźmierkiewicz, R.; Makowski, M.; Murarka, RK.; Scheraga, HA. *Coarse-Graining of Condensed Phase and Biomolecular Systems*. Voth, G., editor. Vol. Chapter 8. Taylor & Francis: 2008. p. 107-122.
12. He Y, Xiao Y, Liwo A, Scheraga HA. *J. Comput. Chem.* 2009; 30:2127. [PubMed: 19242966]
13. Kozłowska U, Maisuradze GG, Liwo A, Scheraga HA. *J. Comput. Chem.* 2010; 31:1143. [PubMed: 20073062]
14. Makowski M, Liwo A, Scheraga HA. *J. Phys. Chem. B.* 2007; 111:2910–2916. [PubMed: 17388416]
15. Makowski M, Liwo A, Maksimiak K, Makowska J, Scheraga HA. *J. Phys. Chem. B.* 2007; 111:2917–2924. [PubMed: 17388417]
16. Makowski M, Sobolewski E, Czaplewski C, Liwo A, Oldziej S, Scheraga HA. *J. Phys. Chem. B.* 2007; 111:2925–2931. [PubMed: 17388418]
17. Makowski M, Sobolewski E, Czaplewski C, Oldziej S, Liwo A, Scheraga HA. *J. Phys. Chem. B.* 2008; 112:11385–11395. [PubMed: 18700740]
18. Thomas PD, Dill KA. *J. Mol. Biol.* 1996; 257:457. [PubMed: 8609636]
19. Ben Naim A. *J. Chem. Phys.* 1997; 107:3698.
20. Czaplewski, C.; Liwo, A.; Makowski, M.; Oldziej, S.; Scheraga, HA. *Multiscale Approaches to Protein Modeling*. Koliński, A., editor. Vol. Chapter 3. Springer; 2010. p. 35-84.
21. Gay JG, Berne BJ. *J. Chem. Phys.* 1981; 74:3316.
22. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. *Nucleic Acid Research.* 2000; 28:235.
23. Makowski M, Liwo A, Scheraga HA. *J. Phys. Chem. B.* 2010 accompanying paper.
24. Pettitt BM, Rossky PJ. *J. Chem. Phys.* 1986; 84:5836.
25. Buckner JK, Jorgensen WL. *J. Am. Chem. Soc.* 1989; 111:2507.
26. Dang LX, Pettitt BM. *J. Phys. Chem.* 1990; 94:4303.
27. Dang LX, Rice JE, Kollman PA. *J. Chem. Phys.* 1990; 93:7528.

28. Yu H-A, Roux B, Karplus M. J. Chem. Phys. 1990; 92:5020.
29. Guardiola E, Rey R, Padró JA. J. Chem. Phys. 1991; 95:2823.
30. Guardiola E, Rey R, Padró JA. Chem. Phys. 1991; 155:187.
31. Dang LX, Pettitt BM, Rossky PJ. J. Chem. Phys. 1992; 96:4046.
32. Karim OA. J. Chem. Phys. 1992; 96:9237.
33. Friedman RA, Mezei M. J. Chem. Phys. 1995; 102:419.
34. Kovalenko A, Hirata F. J. Chem. Phys. 2000; 112:10391.
35. Kovalenko A, Hirata F. J. Chem. Phys. 2000; 112:10403.
36. Masunov A, Lazaridis T. J. Am. Chem. Soc. 2003; 125:1722. [PubMed: 12580597]
37. Shinto H, Morisada S, Miyahara M, Higashitani K. J. Chem. Eng. Japan. 2003; 36:57.
38. Maksimiak K, Rodziewicz-Motowidło S, Czaplewski C, Liwo A, Scheraga HA. J. Phys. Chem. B. 2003; 107:13496.
39. Soetens J-C, Millot C, Chipot C, Jansen G, Angyan JG, Maigret B. J. Phys. Chem. B. 1997; 101:10910.
40. Rozanska X, Chipot C. J. Chem. Phys. 2000; 112:9691.
41. Villarreal M, Montich G. Protein Sci. 2002; 11:2001. [PubMed: 12142454]
42. Hassan SA. J. Phys. Chem. B. 2004; 108:19501.
43. Born M. Z. Phys. 1920; 1:45.
44. Still WC, Tempczyk A, Hawley RC, Hendrickson T. J. Am. Chem. Soc. 1990; 112:6127.
45. Hawkins GD, Cramer CJ, Truhlar DG. J. Phys. Chem. 1996; 100:19824.
46. Bashford D, Case DA. Annu. Rev. Phys. Chem. 2000; 51:129. [PubMed: 11031278]
47. Onufriev A, Bashford D, Case DA. J. Phys. Chem. B. 2000; 104:3172.
48. Wojciechowski M, Lesyng B. J. Phys. Chem. B. 2004; 108:18368.
49. Schnieders MJ, Ponder J. J. Chem. Theory Comput. 2007; 3:2083.
50. Hassan SA. J. Phys. Chem. B. 2007; 111:227. [PubMed: 17201447]
51. Gong H, Hocky G, Freed KF. Proc. Natl. Acad. Sci. U.S.A. 2008; 105:11146. [PubMed: 18678891]
52. Nicholls A, Honig B. J. Comp. Chem. 1991; 12:435.
53. Vorobjev YN, Grant A, Scheraga HA. J. Am. Chem. Soc. 1992; 114:3189.
54. Case DA, Pearlman DA, Caldwell JW, Ross WS, Cheatham TE III, DeBolt S, Ferguson D, Seibel G, Kollman PA. Comput. Phys. Commun. 1995; 91:1.
55. Case, DA.; Darden, TA.; Cheatham, TE., III; Simmerling, CL.; Wang, J.; Duke, RE.; Luo, R.; Merz, KM.; Pearlman, DA.; Crowley, M.; Walker, RC.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, KF.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, DH.; Schafmeister, C.; Ross, WS.; Kollman, PA. AMBER 9. San Francisco: University of California; 2006.
56. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. J. Chem. Phys. 1983; 79:926.
57. Jorgensen WL, Briggs JM. Mol. Phys. 1988; 63:547.
58. Schmidt MW, Baldrige KK, Boatz JA, Elbert ST, Gordon MS, Jensen JA, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA. J. Comput. Chem. 1993; 14:1347.
59. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM. J. Comput. Chem. 1992; 13:1011.
60. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. J. Comput. Chem. 1995; 16:1339.
61. Marquardt DW. J. Soc. Indust. Appl. Math. 1963; 11:431.
62. Makowska J, Makowski M, Gieldoń A, Liwo A, Chmurzyński L. J. Phys. Chem. B. 2004; 108:12222.
63. Makowska J, Makowski M, Chmurzyński L, Liwo A. J. Comput. Chem. 2005; 26:235. [PubMed: 15599950]
64. Magalhaes A, Maigret B, Hoflack J, Gomes JN, Scheraga HA. J. Prot. Chem. 1994; 13:195.

Figure 1.

Definition of variables describing the location of two spheroidal particles (i and j) with respect to each other. The vector $\hat{\mathbf{u}}_{ij}^{(1)}$ is the unit vector of the long axis of particle i , $\hat{\mathbf{u}}_{ij}^{(2)}$ is the unit vector of the long axis of particle j , $\hat{\mathbf{r}}_{ij}$ is the unit vector pointing from particle i to particle j , $\theta_{ij}^{(1)}$ and $\theta_{ij}^{(2)}$ are the angles between the vector $\hat{\mathbf{r}}_{ij}$ and the vectors $\hat{\mathbf{u}}_{ij}^{(1)}$ and $\hat{\mathbf{u}}_{ij}^{(2)}$, respectively, and ϕ_{ij} is the angle of counterclockwise rotation of the vector $\hat{\mathbf{u}}_{ij}^{(2)}$ about the vector $\hat{\mathbf{r}}_{ij}$ from the plane defined by the vector $\hat{\mathbf{u}}_{ij}^{(1)}$ and the vector $\hat{\mathbf{r}}_{ij}$ when looking from the center of particle j toward the center of particle i . This model is being used in the present UNRES force-field.

Figure 2.

Illustration of the new model for the interactions of charged and polar side chains. A side chain of this type consists of a nonpolar site (represented by an ellipsoid of revolution) and a polar/charged site (represented by a shaded sphere). The center of the polar/charged site of side chain i is at the distance $d_i^{(1)}$ from the geometric center of that side chain (SC_i) (which is located between the polar/charged and nonpolar center and represented by a small sphere in the figure), and that of side chain j is at the distance $d_j^{(1)}$ from the side-chain center (SC_j), while the centers of the nonpolar sites of side chains i and j are at distances $d_i^{(2)}$ and $d_j^{(2)}$, respectively, from their geometric centers (SC_i and SC_j , respectively). The vector $\hat{u}_{ij}^{(1)}$ is the unit vector of the long axis of the nonpolar site of side chain i , $\hat{u}_{ij}^{(2)}$ is the unit vector of the long axis of the nonpolar site of side chain j , $\hat{r}_{ij}$ is the unit vector pointing from the geometric center of the nonpolar site of side chain i to that of side chain j , r_{ij} is the distance between these two centers, r_{ij}' is the distance between the center of the charged/polar site of side chains i and j , r_{ij}'' is the distance between the center of the charged site of side chain i and the center of the nonpolar site of side chain j , and r_{ij}''' is the distance between the center of the charged site of side chain j and the center of the nonpolar site of side chain i .

Figure 3.

Partial atomic charges (in electron charge units) of the propionate (a); butyrate (b); n-pentylamine cation (c); and 1-butylguanidnie cation (d) molecules calculated by using the RESP method⁴² based on HF/6-31G* calculations carried out with GAMESS⁴³ used in the calculations with the AMBER force field. The atoms are labeled with standard AMBER atom-type symbols, i.e., CT for an sp^3 carbon atom, and HC for a hydrogen atom connected to a carbon atom, CA for an aromatic or an olefinic carbon atom, C for an sp^2 carboxyl or carbonyl carbon, H for hydrogen bonded to nitrogen atoms, N3 for an sp^3 charged amino-group nitrogen, N2 for an sp^2 in amino-group nitrogen, O2 for a carboxyl group oxygen.

Figure 4.

Illustration of the (a) side-to-side, (b) head-to-head, (c) head-to-side, and (d) side-to-head orientation of two charged and spheroidal particles. The lines represent the long axes of the spheroids. The orientation variables (see Figure 1 for definition) are:

$$\theta_{ij}^{(1)} = 90^\circ,$$

$$\theta_{ij}^{(2)} = 90^\circ,$$

and ϕ_{ij}

$$= 0^\circ \text{ (a); } \theta_{ij}^{(1)}$$

$$= 0^\circ,$$

$$\theta_{ij}^{(2)} = 180^\circ,$$

and ϕ_{ij} undefined (b); $\theta_{ij}^{(1)} = 0^\circ$, $\theta_{ij}^{(2)} = 90^\circ$, and $\phi_{ij} = 0^\circ$ (c); $\theta_{ij}^{(1)}$

$$= 90^\circ,$$

$$\theta_{ij}^{(2)} = 180^\circ, \text{ and } \phi_{ij} \text{ undefined (d)}$$

. Filled circles at one of the ends of each particle refer to the charged “head” of the particle.

Figure 5.

The PMF curves for (a) propionate - propionate, (b) butyrate - butyrate, (c) propionate - butyrate, (d) n-pentylamine cation - n-pentylamine cation, (e) 1-butylguanidine cation - 1-butylguanidine cation, and (f) n-pentylamine cation - 1-butylguanidine cation dimers. The dashed black, red, blue, and green lines correspond to PMFs determined for the side-to-side (Figure 4a), head-to-head (Figure 4b), head-to-side (Figure 4c), and side-to-head (Figure 4d) orientation, respectively, obtained by MD simulations. The solid lines correspond to the analytical approximation to the PMFs, with coefficients determined by least-squares fitting (eq 21) of the analytical expression to the PMF determined by MD simulations, with the Gay-Berne potential (eq 2) to represent the van der Waals interactions between the solute

molecules, the sum of electrostatic (eq 10) and generalized Born potential (eq 11) to represent the electrostatic interactions between charged parts of the side chains, polarization term (eq 13) to represent interactions between charged and nonpolar parts of the side chains, isotropic cavity potential (eq 14), eq 16 to represent the cavity potential, and isotropic Lennard-Jones eq 19 to describe interactions between charged headgroups.

Figure 6.

Plots of the energy components of the analytical approximations to the PMFs for the n-pentylamine cation - 1-butylguanidine cation: (a) E_{GBerne} (eq 2), (b) $E_{el} + E_{pol}^{GB}$ (eq 10 and eq 11, respectively), (c) E_{pol} (eq 13), (d) ΔF_{cav}^{iso} (eq 14), (e) ΔF_{cav} (eq 16), and (f) E_{LJ} (eq 19). The solid black, red, blue, and green lines correspond to the side-to-side (Figure 4a), head-to-head (Figure 4b), head-to-side (Figure 4c), and side-to-head (Figure 4d) orientations.

Table 1

Box dimensions and number of water molecules in molecular dynamics simulations using the AMBER 9.0 force field for the systems studied in this work.

system ^a	box dimension [$\text{\AA}^3$]	number of water molecules
Pr p ⁻ ...Pr p ⁻	36×36×36	1519
But ⁻ ...But ⁻	36×36×36	1545
Pr p ⁻ ...But ⁻	39×39×39	1933
Pna ⁺ ...Pna ⁺	41×41×41	2256
BtG ⁺ ...BtG ⁺	43×43×43	2548
Pna ⁺ ...BtG ⁺	44×44×44	2736

^a Abbreviations: Pr p⁻ - propionate; But⁻ - butyrate; Pna⁺ - n-pentylamine cation; BtG⁺ - 1-butylguanidine cation.

Table 2

Parameters of E_{GB}^{GB} , E_{pol}^{iso} , E_{pol} , ΔF_{cav}^{iso} , ΔF_{cav} , E_{LJ} , $d_1^{(1)}$, $d_2^{(1)}$, $d_1^{(2)}$, and $d_2^{(2)}$ determined by minimization of the function defined by eq 21.^a

Parameter ^a	System ^b							
	Pr ⁻ p ⁻ ... Pr ⁻ p ⁻	But ⁻ ... But ⁻	Pr ⁻ p ⁻ ... But ⁻	Pna ⁺ ... Pna ⁺	But ⁺ ... But ⁺	Pna ⁺ ... But ⁺	Pna ⁺ ... But ⁺	
ϵ_{ij}^0 [kcal/mol]	$2.3 \cdot 10^{-6}$	$2.1 \cdot 10^{-6d}$	0.917	$5.4 \cdot 10^{-4}$	$2.2 \cdot 10^{-5}$		1.46	
σ_{ij}^0 [Å]	5.55	5.57	3.97	5.58	7.58		3.67	
σ_i^{iso} [Å]	9.29	9.31	0.61	2.51	1.21		0.69	
σ_j^{iso} [Å]	<i>e</i>	<i>e</i>	0.69	<i>e</i>	<i>e</i>		0.78	
$\alpha_{ij}^{iso(1)}$ [kcal/mol]	13.23	13.61	79.66	87.59	87.59		79.66	
$\alpha_{ij}^{iso(2)}$ [kcal/mol]	-13.80	-13.78	51.15	63.96	63.95		51.16	
$\alpha_{ij}^{iso(3)}$ [kcal/mol]	1.30	0.52	34.46	36.73	36.73		34.46	
$\alpha_{ij}^{iso(4)}$ _c	33.28	33.26	4.56	21.18	21.19		4.55	
a_i [Å]	0.34	0.34	4.37	2.11	0.41		4.38	
a_j [Å]	<i>e</i>	<i>e</i>	0.68	<i>e</i>	<i>e</i>		0.70	
$\chi_{ij}^{(1)}$ _c	-0.279	-0.996	0.083	0.397	0.146		0.158	

Parameter ^a	System ^b						
	Pr p ⁻ ... Pr p ⁻	But ⁻ ... But ⁻	Pr p ⁻ ... But ⁻	Pna ⁺ ... Pna ⁺	BiG ⁺ ... BiG ⁺	Pna ⁺ ... BiG ⁺	
$\chi_{ij}^{(2)} c$	<i>e</i>	<i>e</i>	0.059	<i>e</i>	<i>e</i>	<i>e</i>	0.009
$\chi_{ij}^{(1)} c$	-0.990 ^d	-0.990 ^d	0.992 ^d	0.257	-0.107		0.381
$\chi_{ij}^{(2)} c$	<i>e</i>	<i>e</i>	0.395	<i>e</i>	<i>e</i>		-0.036
$\chi_{ij}^{(1)} c$	0.074	0.077	0.485	-0.158	0.165		0.011
$\chi_{ij}^{(2)} c$	<i>e</i>	<i>e</i>	-0.091	<i>e</i>	<i>e</i>		0.066
α_{ij}^{pol} [kcal/mol.Å ⁴]	0.936	0.843	0.036	0.013	0.010		0.015
α_{ji}^{pol} [kcal/mol.Å ⁴]	<i>e</i>	<i>e</i>	0.635	<i>e</i>	<i>e</i>		0.046
σ_i [Å]	4.44	4.91	3.50	5.13	6.47		4.14
σ_j [Å]	<i>e</i>	<i>e</i>	6.30	<i>e</i>	<i>e</i>		7.47
$\alpha_{ij}^{(1)}$ [kcal/mol]	0.07	0.08	6.44	0.76	1.10		7.21
$\alpha_{ij}^{(2)}$ [kcal/mol]	-62.38	-62.37	0.37	-0.19	2.58		0.19
$\alpha_{ij}^{(3)}$ [kcal/mol]	6.29	6.34	0.63	3.81	2.63		0.45
$\alpha_{ij}^{(4)} c$	12.97	12.95	13.71	33.01	32.98		13.61

System ^b									
Parameter ^a	Pr p ⁻ ... Pr p ⁻	But ⁻ ... But ⁻	Pr p ⁻ ... But ⁻	Pna ⁺ ... Pna ⁺	ButG ⁺ ... ButG ⁺	Pna ⁺ ... ButG ⁺			
ϵ_{ij} [kcal/mol]	9.2·10 ⁻⁴	7.8·10 ⁻³	1.0·10 ^{-4d}	1.0·10 ^{-4d}	0.01 ^d	0.007 ^d			
σ_{ij} [Å]	5.00	4.89	4.99	3.00	3.18	4.00			
$d_i^{(1)}$ [Å]	0.50 ^d	0.50 ^d	0.50 ^d	1.00 ^d	0.80 ^d	1.00 ^d			
$d_i^{(2)}$ [Å]	0.50 ^d	0.60 ^d	0.50 ^d	0.70 ^d	0.30 ^d	0.70 ^d			
$d_j^{(1)}$ [Å]	<i>e</i>	<i>e</i>	0.50 ^d	<i>e</i>	<i>e</i>	0.80 ^d			
$d_j^{(2)}$ [Å]	<i>e</i>	<i>e</i>	0.60 ^d	<i>e</i>	<i>e</i>	0.30 ^d			

^aThe calculated van der Waals energy was based on the Gay-Berne potential (eq 2), the generalized Born potential (eq 11), the polarization term (eq 13), and the isotropic cavity creation potential (eq 14). The cavity creation energy term was obtained from eq 16; isotropic Lennard-Jones from eq 19, $d_i^{(1)}$ and $d_j^{(1)}$ are the distances from the geometrical center to the center of the charge of particles *i* and *j*, respectively, and $d_i^{(2)}$ and $d_j^{(2)}$ are the distances from the geometrical center to the nonpolar center of particles *i* and *j*, respectively (Figure 2). The parameters are best fitted ones.

^bAbbreviations: Pr p⁻ - propionate; But⁻ - butyrate; Pna⁺ - n-pentylamine cation; ButG⁺ - 1-butylguanidine cation.

^cDimensionless.

^dThis value was fixed during minimization.

^eSolute *i* is the same as *j*.