

On the Transferability of the SPC/L Water Model to Biomolecular Simulation

Alice Glättli[†], Chris Oostenbrink[†], Xavier Daura[‡], Daan P. Geerke[†],
Haibo Yu[†], and Wilfred F. van Gunsteren[†]

[†]Laboratory of Physical Chemistry, Swiss Federal Institute of Technology Zürich, ETH-Hönggerberg, CH-8093 Zürich, Switzerland

[‡]Institució Catalana de Recerca i Estudis Avancats (ICREA) and Institute of Biotechnology and Biomedicine,
Universitat Autònoma de Barcelona, E-08193 Bellaterra, Spain

Received on 08 September, 2003.

We investigated the performance of the recently developed SPC/L model for liquid water, as a pure liquid, in binary mixtures with DMSO, and as a solvent model in a peptide folding simulation. Additionally, in order to test the compatibility with the GROMOS biomolecular force field, free energies of hydration of a set of representative compounds were computed. The results are compared to those for the well established SPC water model, which is generally used as a solvent model in conjunction with the GROMOS force field already for more than two decades. It turns out that as a pure liquid and in binary mixtures with DMSO the SPC/L model outperforms SPC, whereas as solvent in combination with the GROMOS force field both models perform equally well.

1 Introduction

Water is the most abundant compound on the surface of the Earth and the principal constituent of all living organisms. Consequently, it plays a key role in environmental science, industrial technology and biology. Most of the living species contain a large amount of water, varying from 97 % to less than 50 %. The role of water is to act as the most important constituent of body fluids and as a solvent for all kinds of biomolecules. It plays an important role in many biological processes such as molecular transport, folding and aggregation of biopolymers, enzyme-substrate binding and catalysis [1]. On the other hand, water is also one of the substances with most peculiar properties: Water contracts on melting, shows a density maximum in the liquid phase at 4 °C and 1 atm pressure, has high melting, boiling and critical temperatures, and a large dielectric constant [2]. It is, therefore, not surprising that water is the liquid by far most studied by computer simulation. Since the first molecular dynamics (MD) simulation of water in 1971 by Rahman and Stillinger [3] a large number of water models have been developed. However, none of the model potentials proposed in the last thirty years is able to reproduce all the properties of water (for a recent review see reference [4]). While there are only a few *ab initio* models, which explicitly include a quantum-mechanical description of the valence electrons, described in the literature [5, 6, 7, 8], most of the models are empirical. They are based on interaction functions and force fields, which are obtained by fitting a suitable functional form to mostly experimental data, but in some cases also to data extracted from *ab initio* calculations [9]. These models are simpler and computationally less expensive than *ab initio*

models and therefore more commonly applied. The group of empirical models can be divided into flexible and rigid water models, depending on whether the vibrational degrees of freedom are included or not, and into polarizable or non-polarizable models, depending whether the model explicitly accounts for polarization or not. Early effective pair potentials for water such as the rigid SPC (simple point charge, with three interaction sites) [10], SPC/E (extended SPC) [11], TIP3P (transferable intermolecular potential with three interaction sites) or TIP4P (TIP with four interaction sites) [12] models describe the thermodynamic, structural and dynamical properties under near-ambient conditions quite well [13], but fail to reproduce the liquid-vapor coexistence curve [14, 15] and the temperature of the density maximum [16, 17]. It turned out that refining the SPC/E model to improve the agreement with experimental values for saturated densities and vapor pressure resulted in a model that fails to reproduce the second peak of the oxygen-oxygen radial distribution function $g_{OO}(r)$ indicating a loosening of the hydrogen-bond network [18, 19]. The recently developed TIP5P model [20, 21], a rigid nonpolarizable model with five interaction sites, shows an overall very good agreement with experimental data at room temperature and also accurately reproduces the density of liquid water from 238 to 336 K, and thus reproduces the density maximum of water at 277 K. However, it predicts the critical temperature of water significantly below the experimental value. Similarly, the DEC model [22], which employs diffuse charges in addition to point charges, reproduces the thermodynamic and structural properties under various conditions very well, but underestimates the critical temperature as well as the temperature of maximum density by roughly the same amount

as most of the other models. Explicit introduction of polarizability did up to now not show the desired improvement. Polarizable models generally overestimate the dielectric permittivity [4, 23] and do not better describe the coexistence curve, in some case even worsen it [24, 25, 26, 27]. Thus, the connection between the microscopic structure of water and its macroscopic properties is still not fully elucidated and the development of the “ultimate” water model remains a challenge for computational physics and chemistry.

The investigation of the stability and dynamics of a biomolecule clearly requires to explicitly account for solvent degrees of freedom [28]. In a system consisting of a protein in solution a large fraction of the system is represented by solvent molecules. Since in general the protein is of primary interest, the model used to represent the solvent should be computationally as cheap as possible, yet reproduce the basic features of the real solvent well enough. Modeling solvent effects by an implicit representation of the solvent is computationally most efficient, but has shown in some cases to stabilize conformers other than the native structure [29] and to lead to a less realistic dynamic behaviour than an explicit treatment of the solvent degrees of freedom [30]. Suitable water models for use as a solvent in a biomolecular simulation are models that are not only computationally efficient and reproduce the properties of bulk water sufficiently well, but should also be compatible with the force field employed to describe the solute interactions. Although simple effective pair potentials such as SPC, TIP3P and TIP4P are not able to accurately describe the whole range of water properties, they have proven to successfully model water as a solvent in biomolecular simulation. Their major weakness is the overestimation of the diffusion coefficient and the incomplete description of the dielectric properties. Introducing a correction for the self-polarization energy in the effective pair-potential improved this situation. The extended SPC model (SPC/E) shows both a dielectric permittivity and a diffusion coefficient close to the experimentally measured values [11, 13]. However, it turned out to be less suited for combination with current biomolecular force fields, since these do not include the proper environment-dependent self-polarization correction. [31, 32].

We recently investigated different approaches to refine the SPC water model in order to obtain a model with overall improved properties at room temperature and ambient pressure [33, 34]. A simultaneous improvement of diffusion and dielectric properties appeared to be difficult within the class of rigid, nonpolarizable models and without increasing the number of interaction sites [33]. Among the collection of SPC-like water models presented, one model, SPC/L, represents a clear improvement over SPC in the properties of bulk water. However, the question of its range of application is still open. Compared to SPC, SPC/L is characterized by a longer bond length, a smaller bond angle, a smaller partial charge and an attractive van der Waals interaction term for the hydrogen atoms (see Table 1). It is especially the much lower partial charges that might make it incompatible with the GROMOS force field. Here, we wish to investigate the performance of the SPC/L water model as a solvent in

biomolecular simulation.

After a brief description of the SPC/L model and its pure liquid properties, the performance of SPC/L is tested and compared to that of SPC when combined with dimethylsulfoxide (DMSO) to a binary mixture and as a solvent in calculations of the free energy of hydration of a number of different compounds. Co-solvents are in practice often used to steer processes such as (un)folding of proteins, membrane permeation, or cell fusion. Consequently, a water model for use in biomolecular simulation should also perform well in binary mixtures with commonly used co-solvents such as DMSO. Finally, we compare the two water models, SPC and SPC/L, in a simulation of the folding/unfolding of a β -heptapeptide with the sequence H- β -HVal- β -HGlU- β -HLys* - β -HVal-HLys* - β -HGlU- β -HVal. (see Fig. 3, panels A and B). Note that Lys* stands for a lysine residue with one carbon atom less in the side-chain. This peptide is shown by NMR to form a left-handed 3_{14} -helix in aqueous solution [35] (see Fig. 3, panels A and B). The totality of these tests should give an impression of the performance of SPC/L as a solvent in biomolecular simulation and of its compatibility with the current version of the GROMOS force field [36].

Table 1. Parameters of the SPC and SPC/L water models. r_{OH} : OH bond length, \angle HOH: HOH bond angle, q_H : partial charge on the hydrogen atom $2q_H = -q_O$, C_6 : attractive Lennard-Jones coefficient, C_{12} : repulsive Lennard-Jones coefficient.

| Parameters of the SPC and the SPC/L water models | | |
|--|---------|---------|
| Parameter | SPC | SPC/L |
| r_{OH} [nm] | 0.1 | 0.11 |
| \angle HOH [deg] | 109.47 | 104.50 |
| q_H [e] | 0.41 | 0.34425 |
| μ [e nm] | 0.0473 | 0.0464 |
| $(C_6(O))^{\frac{1}{2}} \left[\frac{kJnm^6}{mol} \right]^{\frac{1}{2}}$ | 0.05116 | 0.05116 |
| $(C_{12}(O))^{\frac{1}{2}} 10^{-3} \left[\frac{kJnm^{12}}{mol} \right]^{\frac{1}{2}}$ | 1.623 | 1.623 |
| $(C_6(H))^{\frac{1}{2}} 10^{-3} \left[\frac{kJnm^6}{mol} \right]^{\frac{1}{2}}$ | 0.0 | 0.53 |
| $(C_{12}(H))^{\frac{1}{2}} 10^{-6} \left[\frac{kJnm^{12}}{mol} \right]^{\frac{1}{2}}$ | 0.0 | 0.0 |

2 Methods

2.1 Pure liquid and binary mixture simulations

For both water models, SPC and SPC/L, MD simulations of pure water and of DMSO-water mixtures at six compositions, with DMSO fractions x_{DMSO} of 1, 0.814, 0.478, 0.349, and 0.188, were performed using the GROMOS96 biomolecular simulation package [37]. The DMSO model recently developed by Geerke *et al.* [38] was employed. For each simulation a cubic box was filled with 1000 molecules such that its size corresponded to the experimental density at room temperature and ambient pressure. Periodic boundary conditions were applied. The water and DMSO geometries were constrained using the SHAKE algorithm [39] with a relative geometric tolerance of 10^{-4} . Temperature and

pressure were kept constant by weakly coupling them to a temperature bath at 298 K and to a pressure bath at 1 atm, respectively [40]. The relaxation time for the temperature coupling was set to 0.1 ps, for the pressure coupling to 0.5 ps. The isothermal compressibilities for pure water and pure DMSO were set equal to the experimental values of $7.513 \cdot 10^{-4} \text{ [(kJ mol}^{-1} \text{ nm}^{-3})^{-1}]$ [41] and $8.718 \cdot 10^{-4} \text{ [(kJ mol}^{-1} \text{ nm}^{-3})^{-1}]$ [42]. In the case of mixtures, a linear combination of these two values was taken.

The equations of motion were integrated using the leap frog algorithm and a time step of 2 fs. The pair-wise interaction energy function is expressed by a sum of a Lennard-Jones and a Coulomb term:

$$U(r_{ij}) = \frac{C_{12}(i,j)}{r_{ij}^{12}} - \frac{C_6(i,j)}{r_{ij}^6} + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (1)$$

where r_{ij} is the distance between atoms i and j , q_i is the charge of atom i , ϵ_0 is the dielectric permittivity of vacuum, and $C_6(i,j)$ and $C_{12}(i,j)$ are, respectively, the dispersion and repulsion Lennard-Jones coefficients for the interaction between atoms i and j . The interaction between atoms of pairs of molecules was calculated according to a spherical triple-range cutoff scheme: Short-range van der Waals and electrostatic interactions were evaluated at every time step by using a molecular-pair list that was generated with a short-range cutoff radius of 0.8 nm. Longer range van der Waals and electrostatic interactions, between molecules with the oxygen atoms at a distance longer than 0.8 nm and shorter than a long-range cutoff of 1.4 nm, were evaluated every fifth time step, at which point the pair list was also updated, and were kept unchanged between these updates. To approximate the electrostatic interactions beyond the long-range cutoff ($R_{cl} = 1.4 \text{ nm}$), a reaction field term [43, 44], U_{RF} , was included in (1):

$$U_{RF}(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0} \left[\frac{(\epsilon_{RF} - 1)r_{ij}^2}{(2\epsilon_{RF} + 1)R_{cl}^3} - \frac{3\epsilon_{RF}}{(2\epsilon_{RF} + 1)R_{cl}} \right] \quad (2)$$

For the simulations of pure water and DMSO, the dielectric constant of the continuum outside R_{cl} , ϵ_{RF} , was set equal to the experimental values of 78.5 [45] and 46 [42], respectively. For the mixtures, a linear combination of these values was taken. At the beginning of the simulation the velocities of the atoms were assigned from a Maxwell-Boltzmann distribution at 298 K.

For pure water, 100 ps equilibration were followed by 3 ns simulation from which a range of bulk properties (see Table 2) were calculated as described in reference [33]. For the DMSO-water mixtures and pure DMSO, after an equilibration time of 500 ps, data were collected for 2 ns to evaluate the heat of mixing ΔH_{mix} , the excess molar volume of mixing ΔV_{mix} , the diffusion coefficient of water and DMSO molecules, and the shear viscosity of the system [38].

Table 2. Bulk water properties of the SPC/L model compared to those of the SPC model and experimentally derived data. T : average temperature of the run, ρ : density, ΔH_{vap} : heat of vaporization, ΔF_{exs} : excess Helmholtz energy. ^aThis was recomputed

using equation 3, which is slightly different from the one used in reference [33]. C_p : heat capacity at constant pressure, α : thermal expansion coefficient, κ_T : isothermal compressibility, $r_1, g(r_1)$: position and height of the first maximum of the oxygen-oxygen radial distribution function in the liquid phase, $r_2, g(r_2)$: position and height of the second maximum, r_{min} : minimum energy oxygen-oxygen distance of the water dimer in the gas phase, E_{min} : energy minimum of the water dimer, D : self-diffusion coefficient, τ_i^α : rotational relaxation times of different axes, η : mean value of the shear viscosity, obtained by averaging over the values for η that have been calculated from the three independent off-diagonal components of the pressure tensor, $\epsilon(0)$: dielectric permittivity, G_k : finite-system Kirkwood factor, g_k : infinite-system Kirkwood factor, τ_D : Debye dielectric relaxation time.

| Properties of the SPC and SPC/L water models | | | |
|---|--------|-------------------|-----------------------|
| Property | SPC | SPC/L | Exp. |
| T [K] | 300.7 | 299.1 | 298.15 |
| ρ [g cm ⁻³] | 0.972 | 0.997 | 0.997 ⁴¹ |
| ΔH_{vap} [kJ mol ⁻¹] | 43.7 | 43.95 | 44.0 ⁶⁰ |
| ΔF_{exs} [kJ mol ⁻¹] | 22.7 | 22.3 ^a | 24.0 ⁵⁰ |
| C_p [J mol ⁻¹ K ⁻¹] | 75.6 | 75.3 | 75.3 ⁴⁵ |
| $\alpha \cdot 10^{-4}$ [K ⁻¹] | 7.3 | 7.5 | 2.0 ⁴¹ |
| $\kappa_T \cdot 10^{-6}$ [atm ⁻¹] | 47.4 | 50.2 | 45.8 ⁴¹ |
| r_1 [nm] | 0.278 | 0.273 | 0.273 ⁶¹ |
| $g(r_1)$ | 2.837 | 2.875 | 2.750 ⁶¹ |
| r_2 [nm] | 0.452 | 0.431 | 0.450 ⁶¹ |
| $g(r_2)$ | 1.070 | 1.076 | 1.161 ⁶¹ |
| r_{min} [nm] | 0.275 | 0.273 | - |
| E_{min} [kJ mol] | -27.57 | -27.37 | - |
| $D \cdot 10^{-9}$ [m ² s ⁻¹] | 4.2 | 3.9 | 2.3 ⁶² |
| τ_1^{HH} [ps] | 2.9 | 2.9 | |
| τ_2^{HH} [ps] | 1.7 | 1.9 | 2.0 ⁶³ |
| τ_1^{OH} [ps] | 3.0 | 3.3 | |
| τ_2^{OH} [ps] | 1.6 | 1.7 | 1.95 ⁶⁴⁻⁶⁷ |
| τ_1^μ [ps] | 3.2 | 3.5 | |
| τ_2^μ [ps] | 1.4 | 1.7 | 1.92 ⁶⁸ |
| η [cp] | 0.49 | 0.59 | 0.85 ⁴⁵ |
| μ [e nm] | 0.0473 | 0.0464 | |
| $\epsilon(0)$ | 66.6 | 73.8 | 78.5 ⁴⁵ |
| G_k | 2.64 | 2.96 | |
| g_k | 2.51 | 2.91 | 2.90 ⁶⁹ |
| τ_D [ps] | 6.2 | 7.9 | 8.3 ⁷⁰ |

2.2 Free energies of hydration

Free energies of hydration were calculated for eight small solutes representative for a range of functional groups occurring in biomolecular systems. In addition, the free energy of forming a neopentane shaped cavity was calculated by setting the C_6 van der Waals parameters for a neopentane molecule to zero. The thermodynamic integration (TI) approach was used to calculate the free energy differences between a state A ($\lambda = 0$) with the solute fully interacting with the solvent and a state B ($\lambda = 1$) where these interactions have been turned off. For every solute, 50 ps of equilibration were followed by 150 ps of data collection at 21 evenly spaced λ -points connecting states A and B. Free energy estimates were obtained by numerical integration from $\lambda = 0$

to $\lambda = 1$ of $\langle \partial H / \partial \lambda \rangle_\lambda$ at every value of λ . At intermediate λ -values all nonbonded interactions were made soft in order to remove singularities at the origin ($r_{ij} = 0$), which

would have led to extremely high energies for partially interacting particles[46]. Effectively, the nonbonded interaction functions of equations (1) and (2) can be written as

$$U(r_{ij}) = (1 - \lambda) \left\{ \frac{C_{12}(i, j)}{(\alpha_{LJ}\lambda^2 C_{126} + r_{ij}^6)^2} - \frac{C_6(i, j)}{(\alpha_{LJ}\lambda^2 C_{126} + r_{ij}^6)} + \frac{q_i q_j}{4\pi\epsilon_0} \left[\frac{1}{(\alpha_C\lambda^2 + r_{ij}^2)^{\frac{1}{2}}} - \frac{(\epsilon_{RF} - 1)r_{ij}^2}{(2\epsilon_{RF} + 1)(\alpha_C\lambda^2 + R_{cl}^2)^{\frac{3}{2}}} - \frac{3\epsilon_{RF}}{(2\epsilon_{RF} + 1)R_{cl}} \right] \right\} \quad (3)$$

where both the van der Waals and Coulombic soft-core-parameters were set to $\alpha_{LJ} = 0.5$ and $\alpha_C = 0.5 \text{ nm}^2$, respectively, and C_{126} was defined as C_{12}/C_6 . For interactions involving the SPC/L hydrogen (for which $C_{12} = 0$) this parameter was set to

$$C_{126} = 1000 \left(\frac{\alpha_{LJ}\lambda^2 C_6}{k_B T (1 - \alpha_{LJ}\lambda^2)} \right) \quad (4)$$

avoiding the occurrence of spurious minima in the solute-solvent interaction energy. In the calculations involving the neopentane shaped cavity, where $C_6 = 0$, C_{126} was set to

$$C_{126} = \sqrt{\frac{C_{12}}{k_B T (1 - \alpha_{LJ}\lambda^2)}} \quad (5)$$

The reaction field permittivity ϵ_{RF} was set to 66.6 for these simulations.[33]

The solute models were described by standard GROMOS parameters (parameter set 45A3) [37, 47]. Bond lengths were constrained using the SHAKE algorithm[39], but bond angles and torsional and improper dihedral angles were described by standard potential energy functions[37]. The solutes were placed in periodic truncated octahedral boxes containing 691 to 951 solvent molecules. The box sizes were chosen such that no solvent molecule could simultaneously interact with two periodic copies of the solute at any time during the simulation.

2.3 β -peptide in aqueous solution

The β -heptapeptide was modeled using the GROMOS96 biomolecular force field, parameter set 45A3 [37, 47], along the lines described by Daura *et al.* [48]. Starting from a fully extended structure (all backbone dihedral angles were set to 180°) the peptide was solvated in a truncated octahedron with 3919 SPC and in one with 3877 SPC/L water molecules, respectively. The initial minimum distance between peptide atoms and the square walls of the truncated octahedron was chosen to be 1.4 nm.

After relaxation of the systems using steepest descent energy minimization the MD simulations were started by

taking the initial velocities from a Maxwellian distribution at 298 K. The isothermal compressibility was set to $7.768 \cdot 10^{-4} (\text{kJ mol}^{-1} \text{ nm}^{-3})^{-1}$ for the simulation in SPC and to $8.227 \cdot 10^{-4} (\text{kJ mol}^{-1} \text{ nm}^{-3})^{-1}$ for the simulation in SPC/L, according to the values for the isothermal compressibility of the bulk liquid calculated for both models[33]. While the geometry of the solvent was kept rigid, for the solute only the bond lengths were constrained using the SHAKE algorithm [39]. Otherwise the above described simulation protocol was followed. Data were collected for 50 ns for analysis.

A conformational cluster analysis was performed using the structures at every 0.01 ns as described by Daura *et al.* [49]. To that end the atom-positional root-mean-square deviation (rmsd) using the backbone atoms of residues 2 to 6 was calculated for every pair of structures. For each trajectory structure the number of structures (neighbours) with a $\text{rmsd} \leq 0.1 \text{ nm}$ was determined. The structure with the highest number of (structural) neighbours was then taken as the central member of the cluster of similar structures forming a conformation. After removing the structures belonging to this first, most populated cluster from the pool of structures the procedure was repeated to find the second cluster and so on.

3 Results and Discussion

3.1 Properties of SPC/L as pure liquid

In Table 1 the force field parameters of the SPC and SPC/L models are shown, while in Table 2 the pure liquid properties of the two models at room temperature are presented along with the experimentally derived values for these properties. Both models were fitted to reproduce the experimental density and the heat of vaporization at room temperature and therefore the values for these properties should deviate only a little from the experiment. However, as the SPC water

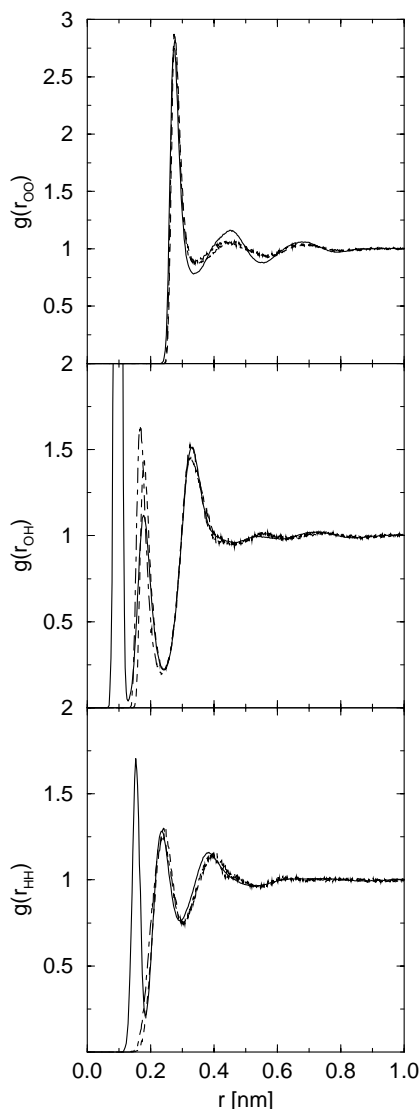


Figure 1. Liquid phase radial distribution function $g(r)$ for the oxygen-oxygen, oxygen-hydrogen and hydrogen-hydrogen pairs of the SPC (dashed line) and SPC/L (dot-dashed line) water models along with the curve (solid line) derived from experiment [61].

model was originally developed using a significantly different simulation protocol (smaller periodic box, smaller cut-off, no reaction field and very short simulation time)[10], it underestimates the experimentally measured value for the density by about 3% using the simulation protocol described in section 2.1. The SPC/L model shows similar thermodynamic properties as SPC, which compare well to the experimentally measured values: The heat capacities C_p of both models (SPC: $75.6 \text{ J mol}^{-1} \text{ K}^{-1}$, SPC/L: $75.3 \text{ J mol}^{-1} \text{ K}^{-1}$) accurately reproduce the experimental value of $75.3 \text{ J mol}^{-1} \text{ K}^{-1}$. However, the thermal expansion coefficient α is for both models much larger than the experimental value of $2.0 \cdot 10^{-4} \text{ K}^{-1}$, indicating that with increasing temperature the density of SPC and SPC/L decreases faster than experimentally observed. The isothermal compressibility κ_T of SPC/L is $50.2 \cdot 10^{-6} \text{ atm}^{-1}$ overestimating the experimental value by 9.6%, while SPC with a κ_T of $47.4 \cdot 10^{-6} \text{ atm}^{-1}$

shows only a 3.5% larger compressibility than experiment. The excess free energy of SPC/L (22.3 kJ mol^{-1}) calculated using equation (3) compares well to the value of 22.7 kJ mol^{-1} obtained for SPC, but is slightly lower than the experimentally derived value [50] of 24.0 kJ mol^{-1} . The structural properties of water reflected in the radial distribution functions for the oxygen-oxygen, oxygen-hydrogen and hydrogen-hydrogen pairs are equally well reproduced by both models (see Table 2 and Fig. 1). The second and third peak of the oxygen-oxygen radial distribution function are in both cases not as pronounced as suggested by experiment. For SPC/L, the first peak of the oxygen-hydrogen distribution function is observed to lie slightly too inward, which is probably due to the lengthening of the OH bond length, which for the same density implies a reduction of the distance between the hydrogen atoms and the nearest oxygen to which they are not covalently bound. In terms of diffusion properties such as self-diffusion and rotational relaxation times of HH, OH and water dipole axes, SPC/L shows a slight improvement over SPC, but still diffuses too fast compared to experiment. The shear viscosity for SPC/L was calculated to be 0.59 cp, which signifies an improvement over the SPC model. The latter model underestimates the experimental value for the shear viscosity by more than 40%. Even though the dipole moment of SPC/L is smaller than the one of SPC, its dielectric properties show a clear improvement. Its permittivity, infinite-system Kirkwood factor and Debye relaxation time compare very well to the experimental values. The smaller bond angle combined with the introduction of purely attractive van der Waals interaction sites are thought to cause this improvement in dielectric properties as discussed in reference [33]. In summary, while SPC/L has similar thermodynamic and structural properties as SPC, its diffusion properties are slightly closer to the experimental values and it shows a clear improvement in the dielectric properties and the viscosities.

3.2 Performance of SPC/L in DMSO-water mixtures

DMSO-water mixtures are non-ideal mixtures. It is experimentally known that the mixing enthalpy, ΔH_{mix} , and the excess volume of mixing, ΔV_{exc} , are negative over the complete range of mole fraction DMSO, x_{DMSO} [51, 52]. Adding DMSO to water slows down the self-diffusion of the water molecules and vice versa [53]. Accordingly, the shear viscosity for the mixtures is larger than for the pure liquids[51]. As shown in Figs. 2A-F, both water models, SPC and SPC/L, in combination with the DMSO model by Geerke *et al.* [38] reflect the non-ideal behaviour of DMSO-water mixtures over a wide range of mole fraction x_{DMSO} . Figs. 2A and 2B show that the DMSO-SPC/L model reproduces the mixing enthalpy and the excess volume of mixing slightly better than DMSO-SPC. The densities of the mixtures at various compositions (Fig. 2C) are correspondingly better reproduced by the DMSO-SPC/L. The latter is mainly due to the fact the SPC/L reproduces the density of bulk water more accurately than SPC. The better performance in the enthalpy of mixing and the excess volume demonstrates

that the DMSO-water interactions are slightly better modeled by the DMSO-SPC/L model than by the DMSO-SPC model. Figs. 2D and 2E reflect the mobility of the DMSO and the water molecules, respectively, as a function of the molar fraction, x_{DMSO} . The effect of slowing down the self-diffusion of the water and DMSO molecules is well reproduced for both water models. Similarly, the change in shear viscosity of the DMSO-water mixture as a function of its composition, displayed in Fig. 2F, qualitatively reproduces the experimental curve, but underestimates the viscosity of the system at any composition. This is probably due to the fact that the shear viscosity for the pure liquids is un-

derestimated by both water models as well as by the DMSO model used [38]. For the diffusion properties and the shear viscosity, the DMSO-SPC/L model again performs slightly better than DMSO-SPC. This has also been reported for the relaxation times of the DMSO dipole moment vector and the O-H bond of the water molecules in DMSO-water mixtures at various compositions and as well as for the static dielectric permittivity of the DMSO-water system [38]. The overall better performance of the DMSO-SPC/L model is only partly due to the improved pure liquid properties of the SPC/L water model compared to SPC. SPC/L seems also to be slightly more compatible with the DMSO model used.

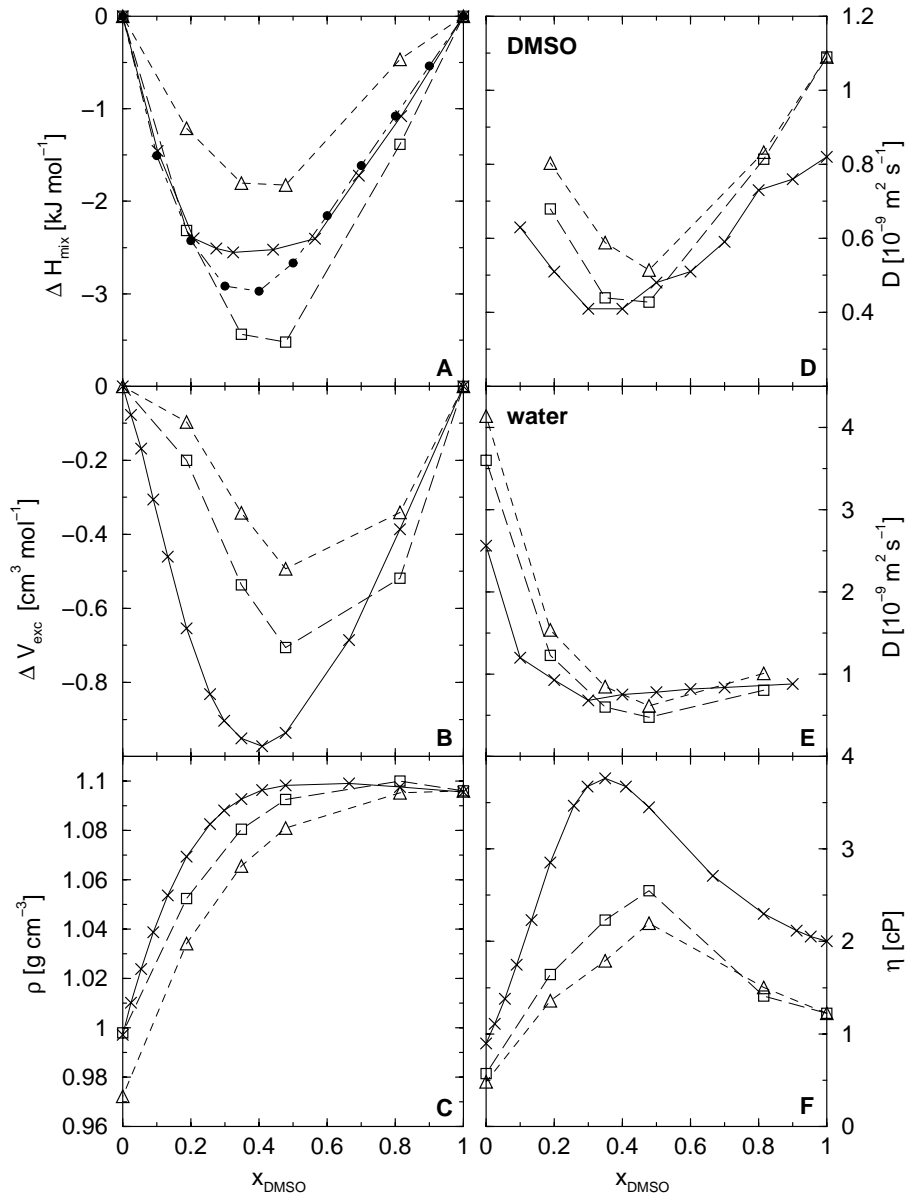


Figure 2. Thermodynamic and dynamic properties of DMSO-water mixtures at 298 K and 1 atm, as function of the mole fraction DMSO, x_{DMSO} , using the DMSO model by Geerke *et al.* [38] in combination with the SPC (\triangle) and SPC/L (\square) models: (A) enthalpies of mixing ΔH_{mix} ; (B) excess volume of mixing ΔV_{exc} ; (C) densities ρ ; (D) diffusion coefficient D of the DMSO molecules; (E) diffusion coefficient D of the water molecules; (F) shear viscosity η . Experimental values were taken from references [51] (\times), [52] (\bullet) and from reference [53] for the diffusion coefficients (\times). Lines are drawn to guide the eye.

3.3 Free energies of hydration in SPC and SPC/L

The calculated free energies of hydration in both SPC and SPC/L are compared in Table 3. Experimental values are also listed in this table. In all cases the value obtained for solvation in SPC/L is slightly larger than the one for SPC. When comparing to the experimental data, it is clear that for the three polar compounds the free energies of hydration are strongly underestimated. This was observed for virtually all polar groups that occur in amino acid side chains [54] and was the main driving force for an extensive recalibration of the nonbonded interaction parameters of the polar atoms in the GROMOS force field [55]. Changing the solvent model from SPC to SPC/L, which behaves so similar in its bulk properties, does not correct these large discrepancies with experiment, as expected. Rather, the only slightly different free energies of hydration indicate that the SPC/L model is generally compatible with the GROMOS force field, which has proven itself in many applications of biomolecular simulation.

Table 3. Free energy of hydration for selected compounds in SPC and SPC/L. All values in kJ mol^{-1} . The experimental values for the aliphatic compounds correspond to values averaged over experimental data from references [57] and [58]. The experimental values for the polar compounds were derived from equilibrium constants between their dilute aqueous solutions and the vapour phase presented in reference [59].

| Free energies of hydration in SPC and SPC/L | | | |
|---|-------|-------|-------|
| Compound | SPC | SPC/L | Exp. |
| Ethane | 8.5 | 9.8 | 7.5 |
| Propane | 9.3 | 10.2 | 8.2 |
| Butane | 9.6 | 10.6 | 8.8 |
| Neopentane | 11.4 | 13.4 | 10.8 |
| Toluene | -2.6 | -0.3 | -3.1 |
| Ethanol | -13.0 | -12.1 | -20.5 |
| 1-Butaneamine | -5.4 | -4.3 | -18.3 |
| Acetamide | -19.7 | -18.3 | -40.6 |
| Neopentane cavity | 69.1 | 72.6 | - |

A naive comparison of the free energies of solvation in both water models, might lead to the expectation that the values obtained for SPC/L should be rather lower than higher. Especially for those solutes that do not possess a dipole in the GROMOS force field such as ethane, propane, butane and neopentane, the solute-solvent interactions involve an additional, purely attractive interaction with the hydrogen atoms and an otherwise unchanged van der Waals interaction with the oxygen when going from SPC to SPC/L. The slightly increased density of SPC/L (3%) may hardly account for the observed differences in free energies.

For a closer inspection, we have determined the free energy of the formation of a neopentane shaped cavity. Since the atoms that form this cavity only interact through the repulsive C_{12} term in equation (1), the interaction with a given water molecule is the same, regardless of whether SPC or SPC/L is used. The free energy difference of 3.5 kJ mol^{-1} can then be traced to two sources. The first effect is enthalpic; the solute-solvent interaction energy of the

cavity changes from 13.2 kJ mol^{-1} (SPC) to 14.7 kJ mol^{-1} (SPC/L), which can be explained from the increased density of the SPC/L model, resulting in a slightly larger number of (purely repulsive) solute-solvent interactions. The remaining 2 kJ mol^{-1} difference in the solvation of the cavity can then only be explained in terms of entropic effects. Apparently it is more difficult to break up the hydrogen bond network to create a cavity in SPC/L than in SPC. For real compounds, interacting fully with all atoms of the solvent models, the first term is offset by an increased attraction due to the hydrogens and attractive solute-solvent oxygen terms, but the entropic term involved in the cavity formation is still likely to give an unfavourable contribution to the free energy of solvation, when going from SPC to SPC/L.

3.4 Simulation of β -heptapeptide folding in SPC and SPC/L

The β -heptapeptide displayed in Fig. 3 was simulated with each of both water models as an attempt to compare their performance as a solvent model for water. In both simulations, the peptide folds within the first 10 ns into a left-handed 3_{14} -helix and thereafter repeatedly unfolds and folds into its helical structure, as illustrated (see Fig. 4, panel A) by the rmsd of the backbone atoms from the NMR model structure displayed in panel B of Fig. 3. Consequently, the most populated conformer corresponds in both cases to the 3_{14} -helix (see panel C of Fig. 3). In SPC, the helix is to 67% populated while in SPC/L to 80%. On the other hand the average lifetime of the most populated conformer is found to be shorter in SPC/L (0.44 ns) than in SPC (0.80 ns). These differences in lifetime and population might be rather due to insufficient sampling than due to the differences in the water models. Previous β -peptide folding studies have illustrated that extending the simulation from 50 ns to 200 ns for instance significantly changes the ratio between populations of folded and unfolded conformation [48, 49, 56]

A conformational cluster analysis of the combined trajectories shows that the peptide populates more or less the same conformational space in both water models. A total of 44 clusters was found. The two most populated clusters represent more than 85% of the total population. The population of the first 10 clusters (larger than 96% of the total population) and the portion of structures in these clusters originating from the trajectory in SPC or the one in SPC/L are shown in Panel B of Fig. 4. The first cluster, of which 54% of the members belong to the simulation in SPC/L, corresponds to the 3_{14} -helix, whereas cluster 2 represents a group of essentially unfolded structures. Clusters 3 and 7 show partly unfolded helices. While the first five clusters contain structures from both trajectories, the clusters 6, 7, 9 and 10 hold exclusively structures of one or the other trajectory. However, the latter group of clusters represent each only 1% or less of the total population.

Table 4 shows the intramolecular hydrogen bond occurrence and energetic properties of the peptide-water systems. In accordance with the atom-positional rmsd and the conformational clustering analysis, the most populated

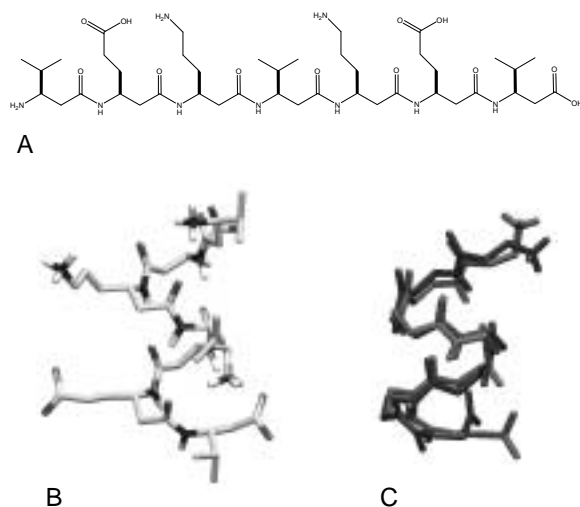


Figure 3. Molecular formula and structure of the β -heptapeptide studied in conjunction with the SPC and SPC/L water models. (A) Molecular formula of the β -heptapeptide. Note that in the simulations the N-terminus as well as the β -lysine type of side-chains were protonated, whereas the C-terminus and the β -glutamate side-chains were deprotonated in line with the experimental data. (B) NMR model structure of the β -heptapeptide in water[35]. The salt-bridges between the β -lysine* and the β -glutamate side-chains favour the formation of a left-handed 3_{14} -helix. (C) Central member structure of the most populated cluster of the β -heptapeptide in SPC (blue) and in SPC/L (red). In both water models the most populated conformer corresponds to the 3_{14} -helix. For clarity, only the backbone atoms are drawn.

intramolecular hydrogen bonds are of the type that forms 14-membered hydrogen bonded rings characteristic for the structure. Using the standard GROMOS definition, where a hydrogen bond is considered to exist when the donor-hydrogen-acceptor angle is larger than 135° and the hydrogen-acceptor distance is less than 0.25 nm, it appears that the C-terminal hydrogen bonds occur more often in SPC than in SPC/L, while the N-terminal hydrogen bonds are more frequently formed in SPC/L. The average number of all intramolecular hydrogen bonds per trajectory structure (averaged over the last 30 ns of the trajectory) is slightly larger in SPC, while the average number of intramolecular hydrogen bonds forming 14-membered hydrogen bonded rings is about the same in both solvents. Accordingly, the average number of intermolecular hydrogen bonds between the peptide and the water molecules per trajectory configuration is in SPC slightly lower than in SPC/L.

While the conformational and dynamic properties of the peptide are essentially the same in both water models, the energetics of the two systems directly reflect the differences between the two water models (see Table 4). Due to the additional attractive van der Waals interaction site on the water hydrogen atoms in SPC/L, the peptide-water and the water-water van der Waals interaction energies are more favourable. On the other hand, the total electrostatic contribution to the nonbonded energy is for SPC/L less negative,

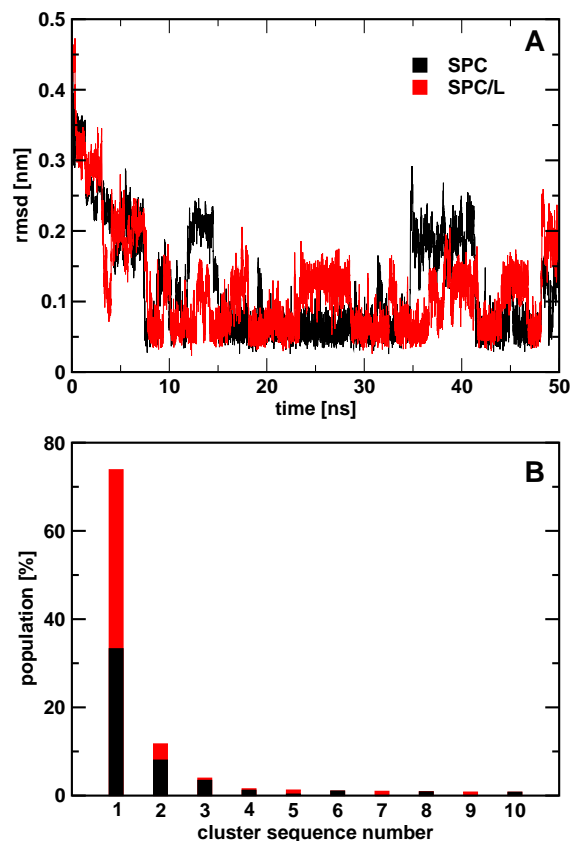


Figure 4. Comparison of the conformational preference of the β -heptapeptide when simulated with the SPC or SPC/L water model. (A) Root-mean-square deviation (rmsd) of the backbone atoms (N, CB, CA, C) of residues 2-6 from the (helical) NMR model structure. Black: Simulation in SPC. Red: Simulation in SPC/L. (B) Clustering of the combined trajectories of the β -heptapeptide in SPC and SPC/L. The plot shows the population in percentage per cluster and the portion of structures per cluster that belong to the trajectory of the peptide in SPC (black) and in SPC/L (red).

which is due to its lower partial charges. Nevertheless, the electrostatic interaction energy between the peptide and the solvent is more favourable for the SPC/L system. This can be explained by the larger density of SPC/L, which results in a larger number of water molecules inside the cut-off radius.

4 Conclusion

We compared the performance of the recently developed SPC/L water model to the much used SPC model and to the available experimental data when using it as a model for pure liquid water, in binary mixtures with DMSO, and as a solvent model in combination with the GROMOS biomolecular force field (version 45A3). Overall, the bulk liquid properties of the SPC/L water model compare very well to the experimentally derived values. Especially its dielectric properties and viscosity represent a clear improvement over

Table 4. Hydrogen bonding and energetics of the β -heptapeptide in SPC and SPC/L. Occurrence of hydrogen bonds in %, the residue

sequence number is indicated between parentheses. NH(i)-O(i+2): occurrence of hydrogen bonds characteristic for a 3_{14} -helix. No other significantly populated hydrogen bonds ($\geq 5\%$) were observed. HB(p-p): average number of intramolecular hydrogen bonds per trajectory structure (peptide-peptide). HB₁₄(p-p): average number of hydrogen bonds forming 14-membered rings. HB(p-w): average number of intermolecular hydrogen bonds per trajectory configuration (peptide-water). Both numbers were obtained by averaging over the last 30 ns of each trajectory. U_{tot} : total energy of the system. U_{kin} : total kinetic energy. U_{pot} : total potential energy. U_{bon} : total covalent energy. U_{vdW} : total van der Waals interaction energy. U_{ele} : total electrostatic interaction energy including the constant and distance-dependent reaction-field contributions. The subscripts p-p, p-w and w-w denote the interaction energy between peptide atoms, peptide and solvent atoms and solvent atoms respectively. All energies are in kJ mol^{-1} .

| Peptide hydrogen bonds and energetics | | |
|---------------------------------------|----------------|------------------|
| | Peptide in SPC | Peptide in SPC/L |
| NH(1)-O(3) | 9.5 | 8.7 |
| NH(2)-O(4) | 58.6 | 67.7 |
| NH(3)-O(5) | 58.4 | 69.0 |
| NH(4)-O(6) | 61.5 | 48.2 |
| NH(5)-O(7) | 38.2 | 25.0 |
| HB(p-p) | 3.5 | 3.1 |
| HB ₁₄ (p-p) | 2.3 | 2.2 |
| HB(p-w) | 30.5 | 31.7 |
| U_{tot} | -134640.6 | -134381.5 |
| U_{kin} | 29581.7 | 29135.9 |
| U_{pot} | -164222.3 | -163517.4 |
| U_{bon} | 201.4 | 202.1 |
| U_{vdW} | 27247.4 | 23191.5 |
| $U_{vdW-p-p}$ | -143.8 | -142.4 |
| $U_{vdW-p-w}$ | -61.5 | -80.0 |
| $U_{vdW-w-w}$ | 27452.7 | 23413.9 |
| U_{ele} | -191678.2 | -186911.0 |
| $U_{ele-p-p}$ | -684.9 | -651.5 |
| $U_{ele-p-w}$ | -2343.3 | -2436.3 |
| $U_{ele-w-w}$ | -188643.0 | -183823.2 |

the SPC model. Likewise, in binary mixtures with DMSO the SPC/L model performs slightly better than SPC. It seems that SPC/L is more compatible with the DMSO model by Geerke *et al.* [38]. The free energies of hydration of a set of compounds in SPC/L compare well to their corresponding values in SPC. The systematic higher free energies in SPC/L can be explained on the one hand by its higher density and on the other hand by its slightly stronger hydrogen bond network. The latter is also reflected in the somewhat lower self-diffusion coefficient, the larger viscosity and the slightly more pronounced local structure of the pure liquid. These findings together with the observation that there is no essential difference in the conformational and dynamic behaviour of a peptide in SPC or SPC/L water lead to the conclusion that SPC/L is as compatible with the GROMOS force field as SPC, while it shows an overall improvement in the bulk water properties as well as in the mixing properties with DMSO.

Acknowledgments

Financial support was obtained from the Schweizer Nationalfonds, project number 2000-063590.00, which is gratefully acknowledged.

References

- [1] J. Israelachvili and H. Wennerström, *Nature* **379**, 219 (1996).
- [2] *Water, a Comprehensive Treatise*, edited by F. Franks (Plenum Press, New York, 1972).
- [3] A. Rahman and F. H. Stillinger, *J. Chem. Phys.* **55**, 3336 (1971).
- [4] B. Guillot, *J. Mol. Liq.* **101**, 219 (2002).
- [5] B. D. Bursulaya and H. J. Kim, *J. Chem. Phys.* **108**, 3277 (1998).
- [6] B. D. Bursulaya and H. J. Kim, *J. Chem. Phys.* **108**, 3286 (1998).
- [7] K. Laasonen, M. Sprik, M. Parrinello, and R. Car, *J. Chem. Phys.* **99**, 9080 (1993).
- [8] P.L. Silvestrelli and M. Parrinello, *J. Chem. Phys.* **111**, 3572 (1999).
- [9] O. Matsuoka, E. Clementi, and M. Yoshimine, *J. Chem. Phys.* **64**, 1351 (1976).
- [10] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, in *Intermolecular forces*, edited by B. Pullman (Reidel, Dordrecht, 1981), p. 331.
- [11] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, *J. Phys. Chem.* **91**, 6269 (1987).
- [12] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).
- [13] D. van der Spoel, P. J. van Maaren, and H. J. C. Berendsen, *J. Chem. Phys.* **108**, 10220 (1998).
- [14] Y. Guissani and B. Guillot, *J. Chem. Phys.* **98**, 8221 (1998).
- [15] J. Alejandro, D. J. Tildesley, and G. A. Chapela, *J. Chem. Phys.* **102**, 4574 (1995).
- [16] S.R. Billeter, P.M King, and W.F. van Gunsteren, *J. Chem. Phys.* **100**, 6692 (1994).
- [17] W. L. Jorgensen and C. Jenson, *J. Comp. Chem.* **19**, 1179 (1998).
- [18] G. C. Boulougouris, I. G. Economou, and D. N. Theodorou, *J. Phys. Chem. B* **102**, 1029 (1998).
- [19] J.R. Errington and A.Z. Panagiotopoulos, *J. Phys. Chem. B* **102**, 7470 (1998).
- [20] M.W. Mahoney and W.L. Jorgensen, *J. Chem. Phys.* **112**, 8910 (2000).
- [21] M.W. Mahoney and W.L. Jorgensen, *J. Chem. Phys.* **114**, 363 (2001).
- [22] B. Guillot and Y. Guissani, *J. Chem. Phys.* **114**, 6720 (2001).
- [23] H. Yu, T. Hansson, and W.F. van Gunsteren, *J. Chem. Phys.* **118**, 221 (2003).
- [24] A. Wallqvist and P.O. Astrand, *J. Chem. Phys.* **102**, 6559 (1995).
- [25] N. Yoshii, H. Yoshie, S. Miura, and S. Okazaki, *J. Chem. Phys.* **109**, 4873 (1999).
- [26] P. Jedlovsky and J. Richardi, *J. Chem. Phys.* **110**, 8019 (1999).

- [27] I. M. Svishchev and T.M. Hayward, *J. Chem. Phys.* **111**, 9034 (1999).
- [28] H. Schäfer, X. Daura, A.E. Mark, and W.F. van Gunsteren, *Proteins* **43**, 45 (2001).
- [29] R. Zhou and B.J. Berne, *Proc. Natl. Acad. Sci.* **99**, 12777 (2002).
- [30] X. Daura, A.E. Mark, and W.F. van Gunsteren, *Comput. Phys. Commun.* **123**, 97 (1999).
- [31] D. P. Tieleman and H. J. C. Berendsen, *J. Chem. Phys.* **105**, 4871 (1996).
- [32] A. R. van Buuren, S.-J. Marrink, and H. J. C. Berendsen, *J. Phys. Chem.* **97**, 9206 (1993).
- [33] A. Glättli, X. Daura, and W.F. van Gunsteren, *J. Chem. Phys.* **116**, 9811 (2002).
- [34] A. Glättli, X. Daura, and W.F. van Gunsteren, *J. Comput. Chem.* **24**, 1087 (2003).
- [35] M.A. Rueping, Ph.D. thesis, Diss. ETH No. 14677, ETH Zürich, Switzerland, 2002.
- [36] L.D. Schuler, X. Daura, and W.F. van Gunsteren, *J. Comput. Chem.* **22**, 1205 (2001).
- [37] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, and I. G. Tironi, *Biomolecular Simulation: The GROMOS96 Manual and User Guide* (vdf Hochschulverlag, ETH Zürich, Switzerland, 1996).
- [38] D.P. Geerke, C. Oostenbrink, N.F.A. van der Vegt, and W.F. van Gunsteren, *J. Phys. Chem. B*, *in press* (2003).
- [39] J.-P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, *J. Comput. Phys.* **23**, 327 (1977).
- [40] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *J. Chem. Phys.* **81**, 3684 (1984).
- [41] G. S. Kell, *J. Chem. Eng. Data* **12**, 66 (1967).
- [42] J. A. Riddick, W. B. Bunger, and T. K. Sakano, *Organic Solvents: Physical Properties and Methods of Purification* (Wiley, New York, USA, 1986).
- [43] M. Neumann, *J. Chem. Phys.* **82**, 5663 (1985).
- [44] I. G. Tironi, R. Sperb, P. E. Smith, and W. F. van Gunsteren, *J. Chem. Phys.* **102**, 5451 (1995).
- [45] *Handbook of Chemistry and Physics*, 56 ed., edited by R. C. Weast (CRC press, Boca Raton, USA, 1976).
- [46] T. C. Beutler, A. E. Mark, R. C. van Schaik, P. R. Gerber, and W. F. van Gunsteren, *Chem. Phys. Lett.* **222**, 529 (1994).
- [47] W. F. van Gunsteren, X. Daura, and A. E. Mark, in *Encyclopaedia of Computational Chemistry*, edited by P. von Ragué Schleyer (Wiley-VCH, Chichester, 1998), Vol. 2, pp. 1211–1216.
- [48] X. Daura, B. Jaun, D. Seebach, W.F. van Gunsteren, and A.E. Mark, *J. Mol. Biol.* **280**, 925 (1998).
- [49] X. Daura, W.F. van Gunsteren, and A.E. Mark, *Proteins: Struct. Funct. & Genet.* **34**, 269 (1999).
- [50] J. Hermans, A. Pathiaseril, and A. Anderson, *J. Am. Chem. Soc.* **110**, 5982 (1988).
- [51] J.M.G. Cowie and P.M. Toporowski, *Can. J. Chem.* **39**, 2240 (1961).
- [52] J. Kenttamaa and J.J. Lindberg, *Suomen Kemist* **33**, 98 (1960).
- [53] K.J. Packer and D.J. Tomlinson, *Trans. Faraday Soc.* **67**, 1302 (1971).
- [54] A. Villa and A.E. Mark, *J. Comp. Chem.* **23**, 548 (2002).
- [55] C. Oostenbrink, A. Villa, A.E. Mark, and W.F. van Gunsteren, *for submission to J. Comp. Chem.* (2003).
- [56] W. F. van Gunsteren, R. Bürge, C. Peter, and X. Daura, *Angew. Chemie Intl. Ed.* **40**, 351 (2001).
- [57] S. Cabani, P. Gianni, V. Mollica, and L. Lepori, *J. Solut. Chem.* **10**, 563 (1981).
- [58] A. Ben-Naim and Y. Marcus, *J. Chem. Phys.* **81**, 2016 (1984).
- [59] R. Wolfenden, L. Andersson, P.M Cullis, and C.C.B. Southgate, *Biochemistry* **20**, 849 (1981).
- [60] J. P. M. Postma, Ph.D. thesis, Rijksuniversiteit, Groningen, The Netherlands, 1985.
- [61] A. K. Soper, *Chem. Phys.* **258**, 121 (2000).
- [62] K. Krynicki, C. D. Green, and D. W. Sawyer, *Faraday Discuss. Chem. Soc.* **66**, 199 (1978).
- [63] B. Halle and H. Wennerström, *J. Chem. Phys.* **75**, 1928 (1981).
- [64] J. R. C. van der Maarel, D. Lankhorst, J. de Bleijser, and J. C. Leyte, *Chem. Phys. Lett.* **122**, 541 (1985).
- [65] R. P. W. J. Struis and J. C. Leyte, *J. Phys. Chem.* **91**, 1639 (1987).
- [66] T. W. N. Bieze, J. R. C. van der Maarel, and J. C. Leyte, *Chem. Phys. Lett.* **216**, 56 (1993).
- [67] R. Ludwig, *Chem. Phys.* **195**, 329 (1995).
- [68] M. S. P. Sansom, I. D. Kerr, J. Breed, and R. Sankaranarayanan, *Biophys. J.* **70**, 693 (1996).
- [69] M. Neumann, *Mol. Phys.* **57**, 97 (1986).
- [70] U. Kaatz, *J. Chem. Eng. Data* **34**, 371 (1989).