New functionalities in the GROMOS biomolecular simulation software

ARTICLE *in* JOURNAL OF COMPUTATIONAL CHEMISTRY · JANUARY 2012

Impact Factor: 3.59 · DOI: 10.1002/jcc.21954 · Source: PubMed

8 AUTHORS, INCLUDING:

Bruno Horta
Federal University of Rio de Janeiro
29 PUBLICATIONS **315** CITATIONS
SEE PROFILE

Wilfred F van Gunsteren
ETH Zurich
563 PUBLICATIONS **42,630** CITATIONS
SEE PROFILE

DOI 10.1002/jcc.21954

New Functionalities in the GROMOS Biomolecular Simulation Software

Anna-Pitschna E. Kunz,[a] Jane R. Allison,[a] Daan P. Geerke,†[a] Bruno A. C. Horta,[a] Philippe H. Hünenberger,[a] Sereina Riniker,[a] Nathan Schmid,[a] and Wilfred F. van Gunsteren*[a]

Since the most recent description of the functionalities of the GROMOS software for biomolecular simulation in 2005 many new functions have been implemented. In this article, the new functionalities that involve modified forces in a molecular dynamics (MD) simulation are described: the treatment of electronic polarizability, an implicit surface area and internal volume solvation term to calculate interatomic forces, functions for the GROMOS coarse-grained supramolecular force field, a multiplicative switching function for nonbonded interactions, adiabatic decoupling of a number of degrees of freedom with temperature or force scaling to enhance sampling, and nonequilibrium MD to calculate the dielectric permittivity or viscosity. Examples that illustrate the use of these functionalities are given. © 2011 Wiley Periodicals, Inc. J Comput Chem 33: 340–353, 2012

Keywords: GROMOS • polarizability • switching function • SASA/VOL • coarse-graining • nonequilibrium dynamics

Introduction

During the past three decades, the GROMOS software for biomolecular simulation has been developed in the research group of van Gunsteren, first at the University of Groningen and since 1990 at the ETH.[1,2] The functionalities of this molecular simulation software were summarized in three papers[3–5] for the GROMOS87 and GROMOS96 FORTRAN releases and for the new C++ version GROMOS05. The latter version has been under continuous development and a release to the scientific community occurred in 2011. A description of its functionalities can be found in the nine GROMOS manuals (http://www.gromos.net). Here, we briefly describe some of the new functionalities that have been added since 2005: electronic polarizability,[6,7] an implicit surface area and internal volume solvation term[8] for the calculation of interatomic forces, functions required to use the GROMOS coarse-grained (CG) supramolecular force field,[9] a multiplicative switching function for nonbonded interactions, which replaces the one described in Ref. 5, adiabatic decoupling as a means to enhance the sampling of a not too large subset of degrees of freedom of a system,[10] and nonequilibrium molecular dynamics (MD) methodology to compute the dielectric permittivity and relaxation[11] or viscosity. The architecture and implementation of the GROMOS C++ code is described elsewhere,[12] as are features with regard to biomolecular structure refinement based on experimental NMR or X-ray data,[13] to the computation of relative free energies,[14] and the analysis of MD trajectories using GROMOS.[15] The new simulation functionalities are illustrated using examples.

Methods

Polarizability

Electronic polarizability allows for a more accurate description of the nonbonded interactions in classical atomistic simulations and must be implemented in the software when using polarizable force fields.[6]

Several methods have been described in the literature that explicitly treat electronic polarization.[6,16] GROMOS uses the charge-on-spring (COS),[17] also called Drude-oscillator[18] or shell,[19] model where an inducible dipole $\boldsymbol{\mu}_i$ is modelled by attaching a massless, virtual site with a point-charge q_i^v to the polarizable center i, via a spring with harmonic force constant k_i^{ho},

$$k_i^{\mathrm{ho}} = \frac{\left(q_i^v\right)^2}{\alpha_i}. \tag{1}$$

The inducible dipole $\boldsymbol{\mu}_i$ adapts its size and direction according to its polarizability α_i and the electric field $\mathbf{E}_i'$ at the COS of the polarizable center i, assuming isotropic α_i and linear dependence of $\boldsymbol{\mu}_i$ on $\mathbf{E}_i'$, according to

$$\boldsymbol{\mu}_i = \alpha_i \mathbf{E}_i'. \tag{2}$$

The charge at the polarizable center with position $\mathbf{r}_i$, which may have a permanent charge q_i, is then $(q_i - q_i^v)$. Thus, the induced dipoles $\boldsymbol{\mu}_i$ are represented by

[a] *A. P. E. Kunz, J. R. Allison, D. P. Geerke, B. A. C. Horta, P. H. Hünenberger, S. Riniker, N. Schmid, W. F. van Gunsteren*
Laboratory of Physical Chemistry, Swiss Federal Institute of Technology ETH, 8093 Zürich, Switzerland
E-mail: wfvgn@igc.phys.chem.ethz.ch

†Present address: Division of Molecular and Computational Toxicology, VU University Amsterdam, 1081 HV Amsterdam, The Netherlands.

Contract/grant sponsors: Swiss National Science Foundation, National Center of Competence in Reserch in structural biology; Contract/grant number: 200020-121913; Contract/grant sponsor: European Research Council; Contract/grant number: 228076; Contract/grant sponsor: Sino-Swiss Science and Technology Cooperation Program; Contract/grant numbers: IZLCZ2-123884 (Swiss) and GJHZ0906 (Chinese)

$$\boldsymbol{\mu}_i = q_i^{\nu}\left(\mathbf{r}_i' - \mathbf{r}_i\right) \tag{3}$$

where $\mathbf{r}_i'$ is the position of the COS. In COS-based schemes in which the COS are not explicitly treated as additional degrees of freedom, the sum of the electrostatic and spring forces acting on any COS should be zero, and the virtual charge q_i^{ν} must be positioned such that

$$\mathbf{f}_i'^{\text{ho}} + \mathbf{f}_i'^{\text{coul}} = 0 \tag{4}$$

with the force $\mathbf{f}_i'^{\text{ho}}$ due to the spring given by

$$\mathbf{f}_i'^{\text{ho}} = -k_i^{\text{ho}}\left(\mathbf{r}_i' - \mathbf{r}_i\right) = -\frac{\left(q_i^{\nu}\right)^2}{\alpha_i}\left(\mathbf{r}_i' - \mathbf{r}_i\right) \tag{5}$$

and $\mathbf{f}_i'^{\text{coul}}$ due to the (Coulombic) electric field at the COS ($\mathbf{E}_i'$) given by

$$\mathbf{f}_i'^{\text{coul}} = q_i^{\nu}\mathbf{E}_i'. \tag{6}$$

To satisfy eq. (4), the $\boldsymbol{\mu}_i$, that is, the $\mathbf{r}_i'$, must be determined from the $\mathbf{E}_i'$, where the field $\mathbf{E}_i'$ does not contain a contribution from the charge $(q_i - q_i^{\nu})$ at $\mathbf{r}_i$. However, as the displacement $|\mathbf{r}_i' - \mathbf{r}_i|$ of the COS from the polarizable center is nonzero on polarization, a better approximation of the ideal inducible dipole $\boldsymbol{\mu}_i$ at site i would require $\mathbf{r}_i'$ to be determined from the electric field $\mathbf{E}_i$ at the polarizable center itself,[7]

$$\mathbf{r}_i' = \mathbf{r}_i + \frac{\alpha_i \mathbf{E}_i}{q_i^{\nu}}. \tag{7}$$

Using this approximation, the total force acting on the COS is only zero if

$$\mathbf{E}_i = \mathbf{E}_i', \tag{8}$$

which is usually not the case for the induced COS dipole due to the nonzero values for $|\mathbf{r}_i' - \mathbf{r}_i|$. By choosing q_i^{ν} large enough, $|\mathbf{r}_i' - \mathbf{r}_i|$ adopts relatively small values, resulting in small differences between $\mathbf{E}_i$ and $\mathbf{E}_i'$. On the other hand, the size of q_i^{ν} is limited to values for which $|\mathbf{r}_i' - \mathbf{r}_i|$ is significant with respect to interatomic distances such that numerical precision is ensured when calculating, for example, interaction energies involving induced dipoles.[7,17] The COS method has been used in combination with iterative procedures to minimize the energy of the COS with respect to its position $\mathbf{r}_i'$, that is, solving eq. (4). In GROMOS q_i^{ν} is generally set to $-8e$.

The COS method treats the induced dipole moments via additional point charges only, which allows for an easy introduction of polarizability into schemes to compute long-range electrostatic forces, such as the reaction-field, Ewald-summation, particle-particle-particle-mesh (P3M), and Particle-Mesh-Ewald (PME) techniques or into a quantum-mechanical Hamiltonian for the electronic degrees of freedom of a (solute) molecule. The electrostatic potential ϕ_i at the polarizable centers i due to the monopoles and dipoles in the system can be expressed using Coulombic terms only

$$\phi_i(\mathbf{r}^N, \mathbf{r}'^N) = \frac{1}{4\pi\varepsilon_0\varepsilon_{\text{cs}}}\sum_{j\neq i}^{N}\left[\frac{\left(q_j - q_j^{\nu}\right)}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{q_j^{\nu}}{\left|\mathbf{r}_i - \mathbf{r}_j'\right|}\right], \tag{9}$$

where the positions of the N atoms and corresponding virtual charges are denoted by $\mathbf{r}^N = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ and $\mathbf{r}'^N = (\mathbf{r}_1'\mathbf{r}_2', \ldots, \mathbf{r}_N')$, and ε_{cs} is the relative dielectric permittivity used for the model interactions. Because of the dependence of the $\mathbf{r}_i'$ on $\mathbf{r}^N$ and $\mathbf{r}'^N$ via $\mathbf{E}_i$ in eq. (7), the relation between ϕ_i and the electric field $\mathbf{E}_i$ is given by

$$\mathbf{E}_i = -\nabla_i\phi_i(\mathbf{r}^N, \mathbf{r}'^N) = -\left(\frac{\partial\phi_i}{\partial\mathbf{r}_i} + \sum_{k\neq i}^{N}\frac{\partial\phi_i}{\partial\mathbf{r}_k'} \bullet \frac{\partial\mathbf{r}_k'}{\partial\mathbf{r}_i}\right). \tag{10}$$

when applying a Born–Oppenheimer-like iterative self-consistent field procedure, however, all $\mathbf{r}'^N$ are at every iteration step determined in the electric field due to all other q_j and q_j^{ν}. When using a convergence criterion which minimizes ϕ_i with respect to the positions $\mathbf{r}_k'$, the second term in eq. (10) is zero at convergence because $\partial\phi_i/\partial\mathbf{r}_k' = 0$. Thus,

$$\mathbf{E}_i = -\frac{\partial\phi_i}{\partial\mathbf{r}_i} = \frac{1}{4\pi\varepsilon_0\varepsilon_{\text{cs}}}\sum_{j\neq i}^{N}\left[\frac{\left(q_j - q_j^{\nu}\right)(\mathbf{r}_i - \mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|^3} + \frac{q_j^{\nu}\left(\mathbf{r}_i - \mathbf{r}_j'\right)}{\left|\mathbf{r}_i - \mathbf{r}_j'\right|^3}\right]. \tag{11}$$

The electrostatic part V^{ele} of the potential energy can also be expressed in terms of Coulomb interactions. The only non-Coulombic term to be added to V^{ele} is the self-polarization energy V^{self}, which in the COS model can be expressed in terms involving point charges as well

$$V^{\text{ele}} = V^{\text{coul}} + V^{\text{self}} \tag{12}$$

with

$$V^{\text{coul}}(\mathbf{r}^N, \mathbf{r}'^N) = \frac{1}{4\pi\varepsilon_0\varepsilon_{\text{cs}}}\sum_{i=1}^{N-1}\sum_{j>i}^{N}\left[\frac{\left(q_i - q_i^{\nu}\right)\left(q_j - q_j^{\nu}\right)}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{\left(q_i - q_i^{\nu}\right)q_j^{\nu}}{\left|\mathbf{r}_i - \mathbf{r}_j'\right|} + \frac{\left(q_j - q_j^{\nu}\right)q_i^{\nu}}{\left|\mathbf{r}_i' - \mathbf{r}_j\right|} + \frac{q_i^{\nu}q_j^{\nu}}{\left|\mathbf{r}_i' - \mathbf{r}_j'\right|}\right] \tag{13}$$

and

$$V^{\text{self}}(\mathbf{r}^N, \mathbf{r}'^N) = \frac{1}{2}\sum_{i=1}^{N}\frac{\left(q_i^{\nu}\right)^2}{\alpha_i}\left|\mathbf{r}_i' - \mathbf{r}_i\right|^2. \tag{14}$$

Next, we consider the expression for the forces $\mathbf{f}_i$ that act on (polarizable) atomic centers i

$$\mathbf{f}_i = -\nabla_i V_i^{\text{ele}}(\mathbf{r}^N, \mathbf{r}'^N) = -\left(\frac{\partial V^{\text{ele}}}{\partial\mathbf{r}_i} + \sum_{k\neq i}^{N}\frac{\partial V^{\text{ele}}}{\partial\mathbf{r}_k'} \bullet \frac{\partial\mathbf{r}_k'}{\partial\mathbf{r}_i}\right). \tag{15}$$

Note again the dependence of the $\mathbf{r}'_k$ on the $\mathbf{r}_i$ which appears in the second term on the right in eq. (15), which might adopt nonzero values because V^{ele} not only contains terms due to the ϕ_i, the first two terms on the right in eq. (13), but also due to the ϕ_j, the last two terms on the right in eq. (13), and V^{self}, whereas when using eq. (7) only ϕ_i is minimized with respect to all $\mathbf{r}'_i$. When nevertheless using assumptions in eqs. (4) and (8), eq. (15) reduces to

$$\mathbf{f}_i^{\text{red}} = -\frac{\partial V^{\text{ele}}}{\partial \mathbf{r}_i} = -\left(\frac{\partial V^{\text{coul}}}{\partial \mathbf{r}_i} + \frac{\partial V^{\text{self}}}{\partial \mathbf{r}_i}\right) \tag{16}$$

From the assumptions in eqs. (4) and (8), we have

$$-\frac{\partial V^{\text{self}}}{\partial \mathbf{r}_i} = \mathbf{f}_i^{\text{ho}} = -\mathbf{f}_i'^{\text{ho}} = \mathbf{f}_i'^{\text{coul}} = -\frac{\partial V^{\text{coul}}}{\partial \mathbf{r}'_i} \tag{17}$$

and the reduced expression $\mathbf{f}_i^{\text{red}}$ for the atomic forces becomes

$$\mathbf{f}_i^{\text{red}} = \frac{1}{4\pi\varepsilon_0\varepsilon_{\text{cs}}}\sum_{i=1}^{N-1}\sum_{j>i}^{N}\left[\frac{(q_i - q_i^{\nu})(q_j - q_j^{\nu})(\mathbf{r}_i - \mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|^3} + \frac{(q_i - q_i^{\nu})q_j^{\nu}(\mathbf{r}_i - \mathbf{r}'_j)}{|\mathbf{r}_i - \mathbf{r}'_j|^3} + \frac{(q_j - q_j^{\nu})q_i^{\nu}(\mathbf{r}'_i - \mathbf{r}_j)}{|\mathbf{r}'_i - \mathbf{r}_j|^3} + \frac{q_i^{\nu}q_j^{\nu}(\mathbf{r}'_i - \mathbf{r}'_j)}{|\mathbf{r}'_i - \mathbf{r}'_j|^3}\right]. \tag{18}$$

If the COS model is used with a reaction-field, a reaction-field correction term[11] is to be added to V^{coul}

$$\begin{aligned} V^{\text{rf}}(\mathbf{r}^N, \mathbf{r}'^N,) = \frac{1}{4\pi\varepsilon_0\varepsilon_{\text{cs}}}\sum_{i=1}^{N-1}\sum_{\substack{j>i \\ j \text{ inside cut-off } i}}^{N} & \times\left[(q_i - q_i^{\nu})(q_j - q_j^{\nu})\left(-\frac{\frac{1}{2}C_{\text{rf}}|\mathbf{r}_i - \mathbf{r}_j|^2}{R_{\text{rf}}^3} - \frac{1 - \frac{1}{2}C_{\text{rf}}}{R_{\text{rf}}}\right)\right. \\ & + (q_i - q_i^{\nu})q_j^{\nu}\left(-\frac{\frac{1}{2}C_{\text{rf}}|\mathbf{r}_i - \mathbf{r}'_j|^2}{R_{\text{rf}}^3} - \frac{1 - \frac{1}{2}C_{\text{rf}}}{R_{\text{rf}}}\right) \\ & + q_i^{\nu}(q_j - q_j^{\nu})\left(-\frac{\frac{1}{2}C_{\text{rf}}|\mathbf{r}'_i - \mathbf{r}_j|^2}{R_{\text{rf}}^3} - \frac{1 - \frac{1}{2}C_{\text{rf}}}{R_{\text{rf}}}\right) \\ & \left. + q_i^{\nu}q_j^{\nu}\left(-\frac{\frac{1}{2}C_{\text{rf}}|\mathbf{r}'_i - \mathbf{r}'_j|^2}{R_{\text{rf}}^3} - \frac{1 - \frac{1}{2}C_{\text{rf}}}{R_{\text{rf}}}\right)\right] \\ & - \sum_{i=1}^{N}\frac{q_i^2}{4\pi\varepsilon_0\varepsilon_{\text{cs}}}\frac{1}{2}\frac{\left(1 - \frac{1}{2}C_{\text{rf}}\right)}{R_{\text{rf}}}, \end{aligned} \tag{19}$$

where R_{rf} is the reaction-field cut-off and C_{rf} is the reaction-field constant defined as[11]

$$C_{\text{rf}} = \frac{2(\varepsilon_{\text{cs}} - \varepsilon_{\text{rf}})(1 + \kappa R_{\text{rf}}) - \varepsilon_{\text{rf}}(\kappa R_{\text{rf}})^2}{(\varepsilon_{\text{cs}} + 2\varepsilon_{\text{rf}})(1 + \kappa R_{\text{rf}}) + \varepsilon_{\text{rf}}(\kappa R_{\text{rf}})^2} \tag{20}$$

and κ is the inverse Debye screening length and ε_{rf} is the reaction-field dielectric permittivity outside the reaction-field cut-off radius R_{rf}. The reaction-field term makes the electrostatic force and interaction zero at the reaction-field cut-off distance. Using a cut-off radius plus reaction-field, the summation over j in eqs. (13) and (18) is over sites j inside the cut-off of site i, and omits the covalently bound nearest neighbors j of atom i that are excluded from the nonbonded interaction. These neighbors j of i must not be excluded in the summation over j in eq. (19) because they contribute to the reaction-field.[11] The last term in eq. (19) is a constant that is added to represent the self-interaction of the permanent, that is, non-COS, charges.

Off-atom site. For certain polarizable models,[20,21] it is necessary to add an additional virtual atomic interaction center M. The position of this virtual site M is defined in the plane of the three nonvirtual atoms i, j, and k,

$$\mathbf{r}_M = \mathbf{r}_i + \frac{\gamma}{2}(\mathbf{r}_{ji} + \mathbf{r}_{ki}), \tag{21}$$

where γ is a constant, which determines $d_{iM} = r_{iM}$ as a function of the distances $d_{ij} = r_{ij}$ and $d_{ik} = r_{ik}$. For example, in a model for H_2O, i would denote the oxygen atom and j and k the hydrogen atoms. The addition of the massless site M defined by eq. (21) does not introduce any extra degrees of freedom into the molecule in the calculation of the kinetic energy of the system. The "pseudo-force" $\mathbf{f}_M$ that acts on the virtual-atom site M is redistributed to the nonvirtual atoms i, j, and k according to

$$\begin{aligned} &\mathbf{f}_i + (1 - \gamma)\mathbf{f}_M \\ &\mathbf{f}_j + \frac{\gamma}{2}\mathbf{f}_M \\ &\mathbf{f}_k + \frac{\gamma}{2}\mathbf{f}_M. \end{aligned} \tag{22}$$

Damping of the polarization. A problematic feature of most polarizable models, apart from their larger demand for computing power than nonpolarizable ones, is their tendency to allow for overpolarization in the presence of strong local electric fields, leading to a polarization catastrophe, and a too large static dielectric permittivity. There are several approaches to resolve this problem. Using the GROMOS force fields, a polarization catastrophe is avoided by having a large enough repulsive Lennard–Jones interaction between nonhydrogen atoms to ensure that dipole–dipole distances are larger than $(4\alpha_i^2)^{1/6}$. In addition, the GROMOS simulation software allows the linear dependence of the induced dipole $\boldsymbol{\mu}_{\text{ind},i}$ on the electric field $\mathbf{E}_i$ to be substituted by a sublinear dependence for large field strengths,[21] which can be achieved by making the polarizability α_i electric field dependent. For example, α_i is replaced by $\alpha_{D,i} = \alpha_{D,i}(E)$ for large $\mathbf{E}_i$ using

$$\alpha_{D,i} = \begin{cases} \alpha_i & \text{for } E_i \le E_{0,i} \\ \dfrac{\alpha_i E_{0,i}}{p_i E_i}\left[p_i + 1 - \left(\dfrac{E_{0,i}}{E_i}\right)^{p_i}\right] & \text{for } E_i > E_{0,i}. \end{cases} \tag{23}$$

where p_i is a polarization damping parameter. The induced dipole is then defined as

$$\boldsymbol{\mu}_i^{\text{ind}} = \begin{cases} \alpha_i \mathbf{E}_i & \text{for } E_i \le E_{0,i} \\ \dfrac{\alpha_i E_{0,i}}{p_i}\left[p_i + 1 - \left(\dfrac{E_{0,i}}{E_i}\right)^{p_i}\right]\dfrac{\mathbf{E}_i}{E_i} & \text{for } E_i > E_{0,i} \end{cases}. \quad (24)$$

This change in the expression for $\boldsymbol{\mu}_i^{\text{ind}}$ also influences the self-polarization contribution to the potential energy

$$V^{\text{self},i} = \begin{cases} \dfrac{1}{2}\alpha_i E_i^2 & \text{for } E_i \le E_{0,i} \\ \dfrac{1}{2}\alpha_i E_{0,i}^2 + \dfrac{\alpha_i E_{0,i}^2}{p_i(p_i-1)} & \\ \times\left[-p_i^2 + (p_i^2-1)\left(\dfrac{E_i}{E_{0,i}}\right) + \left(\dfrac{E_{0,i}}{E_i}\right)^{p_i-1}\right] & \text{for } E_i > E_{0,i} \end{cases} \quad (25)$$

with $V^{\text{self}} = \sum_i V^{\text{self},i}$, where i runs over all polarizable centers.

Free energy calculation. For free energy calculations using the coupling parameter approach, the Hamiltonian of the system has to be made dependent on a coupling parameter λ. Free energy differences between two states A ($\lambda = 0$) and B ($\lambda = 1$) are then obtained from a thermodynamic integration over λ of the averages of the derivative of the Hamiltonian with respect to λ. Several terms in the λ-dependent Hamiltonian will change compared to the nonpolarizable case.

The λ-dependent V^{coul} term in eq. (12) reads

$$\begin{aligned} V^{\text{coul}}(\mathbf{r}^N, \mathbf{r}'^N; \lambda) = \frac{1}{4\pi\varepsilon_0\varepsilon_{\text{cs}}} \sum_{i=1}^{N-1} \sum_{\substack{j>i \\ j \text{ inside cut-off } i \\ (i,j) \text{ not excluded}}}^{N} \\ \times \Bigg[(1-\lambda)^n \Bigg\{ \left(q_i^A - q_i^{A,v}\right)\left(q_j^A - q_j^{A,v}\right)\left(\frac{1}{(|\mathbf{r}_i - \mathbf{r}_j|^2 + \alpha_{\text{C}}\lambda^2)^{1/2}}\right) \\ + \left(q_i^A - q_i^{A,v}\right) q_j^{A,v}\left(\frac{1}{\left(|\mathbf{r}_i - \mathbf{r}'_j|^2 + \alpha_{\text{C}}\lambda^2\right)^{1/2}}\right) \\ + q_i^{A,v}\left(q_j^A - q_j^{A,v}\right)\left(\frac{1}{\left(|\mathbf{r}'_i - \mathbf{r}_j|^2 + \alpha_{\text{C}}\lambda^2\right)^{1/2}}\right) \\ + q_i^{A,v} q_j^{A,v}\left(\frac{1}{\left(|\mathbf{r}'_i - \mathbf{r}'_j|^2 + \alpha_{\text{C}}\lambda^2\right)^{1/2}}\right)\Bigg\} \\ + \lambda^n \Bigg\{ \left(q_i^B - q_i^{B,v}\right)\left(q_j^B - q_j^{B,v}\right)\left(\frac{1}{(|\mathbf{r}_i - \mathbf{r}_j|^2 + \alpha_{\text{C}}(1-\lambda)^2)^{1/2}}\right) \\ + \left(q_i^B - q_i^{B,v}\right) q_j^{B,v}\left(\frac{1}{\left(|\mathbf{r}_i - \mathbf{r}'_j|^2 + \alpha_{\text{C}}(1-\lambda)^2\right)^{1/2}}\right) \\ + q_i^{B,v}\left(q_j^B - q_j^{B,v}\right)\left(\frac{1}{\left(|\mathbf{r}'_i - \mathbf{r}_j|^2 + \alpha_{\text{C}}(1-\lambda)^2\right)^{1/2}}\right) \\ + q_i^{B,v} q_j^{B,v}\left(\frac{1}{\left(|\mathbf{r}'_i - \mathbf{r}'_j|^2 + \alpha_{\text{C}}(1-\lambda)^2\right)^{1/2}}\right)\Bigg\}\Bigg] \end{aligned} \quad (26)$$

and the λ-dependent, reaction-field correction term V^{rf} is

$$\begin{aligned} V^{\text{rf}}(\mathbf{r}^N, \mathbf{r}'^N; \lambda) = -\frac{1}{4\pi\varepsilon_0\varepsilon_{\text{cs}}} \sum_{i=1}^{N-1} \sum_{\substack{j>i \\ j \text{ inside cut-off } i}}^{N} \Bigg[(1-\lambda)^n \\ \times \Bigg\{ \left(q_i^A - q_i^{A,v}\right)\left(q_j^A - q_j^{A,v}\right)\left(C_{\text{rf,aux}}(\lambda)|\mathbf{r}_i - \mathbf{r}_j|^2 + \frac{1-\frac{1}{2}C_{\text{rf}}}{R_{\text{rf}}}\right) \\ + \left(q_i^A - q_i^{A,v}\right) q_j^{A,v}\left(C_{\text{rf,aux}}(\lambda)|\mathbf{r}_i - \mathbf{r}'_j|^2 + \frac{1-\frac{1}{2}C_{\text{rf}}}{R_{\text{rf}}}\right) \\ + q_i^{A,v}\left(q_j^A - q_j^{A,v}\right)\left(C_{\text{rf,aux}}(\lambda)|\mathbf{r}'_i - \mathbf{r}_j|^2 + \frac{1-\frac{1}{2}C_{\text{rf}}}{R_{\text{rf}}}\right) \\ + q_i^{A,v} q_j^{A,v}\left(C_{\text{rf,aux}}(\lambda)|\mathbf{r}'_i - \mathbf{r}'_j|^2 + \frac{1-\frac{1}{2}C_{\text{rf}}}{R_{\text{rf}}}\right)\Bigg\} \\ + \lambda^n \Bigg\{ \left(q_i^B - q_i^{B,v}\right)\left(q_j^B - q_j^{B,v}\right)\left(C_{\text{rf,aux}}(1-\lambda)|\mathbf{r}_i - \mathbf{r}_j|^2 + \frac{1-\frac{1}{2}C_{\text{rf}}}{R_{\text{rf}}}\right) \\ + \left(q_i^B - q_i^{B,v}\right) q_j^{B,v}\left(C_{\text{rf,aux}}(1-\lambda)|\mathbf{r}_i - \mathbf{r}'_j|^2 + \frac{1-\frac{1}{2}C_{\text{rf}}}{R_{\text{rf}}}\right) \\ + q_i^{B,v}\left(q_j^B - q_j^{B,v}\right)\left(C_{\text{rf,aux}}(1-\lambda)|\mathbf{r}'_i - \mathbf{r}_j|^2 + \frac{1-\frac{1}{2}C_{\text{rf}}}{R_{\text{rf}}}\right) \\ + q_i^{B,v} q_j^{B,v}\left(C_{\text{rf,aux}}(1-\lambda)|\mathbf{r}'_i - \mathbf{r}'_j|^2 + \frac{1-\frac{1}{2}C_{\text{rf}}}{R_{\text{rf}}}\right)\Bigg\}\Bigg] \\ - \frac{1}{4\pi\varepsilon_0\varepsilon_{\text{cs}}} \sum_{i=1}^{N} \left[(1-\lambda)q_i^A + \lambda q_i^B\right]^2 \frac{1}{2}\frac{\left(1-\frac{1}{2}C_{\text{rf}}\right)}{R_{\text{rf}}}, \end{aligned} \quad (27)$$

where

$$C_{\text{rf,aux}}(\lambda) = \frac{\frac{1}{2}C_{\text{rf}}}{\left(R_{\text{rf}}^2 + \alpha_{\text{C}}\lambda^2\right)^{3/2}} \quad (28)$$

is an auxiliary function to simplify the formula, $q_i^{A(B)}$ is the charge on atom i in state $A(B)$, α_{C} is the Coulombic soft-core parameter, n is an integer larger than zero, and C_{rf} is defined in eq. (20). The derivatives with respect to λ are then

$$\begin{aligned} \frac{\partial V^{\text{coul}}}{\partial\lambda} = \frac{1}{4\pi\varepsilon_0\varepsilon_{\text{cs}}} \sum_{i=1}^{N-1} \sum_{\substack{j>i \\ j \text{ inside cut-off } i \\ (i,j) \text{ not excluded}}}^{N} \Bigg[-(1-\lambda)^n\lambda\alpha_{\text{C}} \\ \times \Bigg\{ \left(q_i^A - q_i^{A,v}\right)\left(q_j^A - q_j^{A,v}\right)\left(\frac{1}{(|\mathbf{r}_i - \mathbf{r}_j|^2 + \alpha_{\text{C}}\lambda^2)^{3/2}}\right) \\ + \left(q_i^A - q_i^{A,v}\right) q_j^{A,v}\left(\frac{1}{\left(|\mathbf{r}_i - \mathbf{r}'_j|^2 + \alpha_{\text{C}}\lambda^2\right)^{3/2}}\right) \\ + q_i^{A,v}\left(q_j^A - q_j^{A,v}\right)\left(\frac{1}{\left(|\mathbf{r}'_i - \mathbf{r}_j|^2 + \alpha_{\text{C}}\lambda^2\right)^{3/2}}\right) \\ + q_i^{A,v} q_j^{A,v}\left(\frac{1}{\left(|\mathbf{r}'_i - \mathbf{r}'_j|^2 + \alpha_{\text{C}}\lambda^2\right)^{3/2}}\right)\Bigg\} + \lambda^n(1-\lambda)\alpha_{\text{C}} \end{aligned}$$

$$
\begin{aligned}
&\times\left\{\left(q_i^B-q_i^{B,v}\right)\left(q_j^B-q_j^{B,v}\right)\left(\frac{1}{(|\mathbf{r}_i-\mathbf{r}_j|^2+\alpha_C(1-\lambda)^2)^{3/2}}\right)\right.\\
&+\left(q_i^B-q_i^{B,v}\right)q_j^{B,v}\left(\frac{1}{\left(\left|\mathbf{r}_i-\mathbf{r}_j'\right|^2+\alpha_C(1-\lambda)^2\right)^{3/2}}\right)\\
&+q_i^{B,v}\left(q_j^B-q_j^{B,v}\right)\left(\frac{1}{\left(\left|\mathbf{r}_i'-\mathbf{r}_j\right|^2+\alpha_C(1-\lambda)^2\right)^{3/2}}\right)\\
&\left.+q_i^{B,v}q_j^{B,v}\left(\frac{1}{\left(\left|\mathbf{r}_i'-\mathbf{r}_j'\right|^2+\alpha_C(1-\lambda)^2\right)^{3/2}}\right)\right\}-n(1-\lambda)^{n-1}\\
&\times\left\{\left(q_i^A-q_i^{A,v}\right)\left(q_j^A-q_j^{A,v}\right)\left(\frac{1}{(|\mathbf{r}_i-\mathbf{r}_j|^2+\alpha_C\lambda^2)^{1/2}}\right)\right.\\
&+\left(q_i^A-q_i^{A,v}\right)q_j^{A,v}\left(\frac{1}{\left(\left|\mathbf{r}_i-\mathbf{r}_j'\right|^2+\alpha_C\lambda^2\right)^{1/2}}\right)\\
&+q_i^{A,v}\left(q_j^A-q_j^{A,v}\right)\left(\frac{1}{\left(\left|\mathbf{r}_i'-\mathbf{r}_j\right|^2+\alpha_C\lambda^2\right)^{1/2}}\right)\\
&\left.+q_i^{A,v}q_j^{A,v}\left(\frac{1}{\left(\left|\mathbf{r}_i'-\mathbf{r}_j'\right|^2+\alpha_C\lambda^2\right)^{1/2}}\right)\right\}+n\lambda^{n-1}\\
&\times\left\{\left(q_i^B-q_i^{B,v}\right)\left(q_j^B-q_j^{B,v}\right)\left(\frac{1}{(|\mathbf{r}_i-\mathbf{r}_j|^2+\alpha_C(1-\lambda)^2)^{1/2}}\right)\right.\\
&+\left(q_i^B-q_i^{B,v}\right)q_j^{B,v}\left(\frac{1}{\left(\left|\mathbf{r}_i-\mathbf{r}_j'\right|^2+\alpha_C(1-\lambda)^2\right)^{1/2}}\right)\\
&+q_i^{B,v}\left(q_j^B-q_j^{B,v}\right)\left(\frac{1}{\left(\left|\mathbf{r}_i'-\mathbf{r}_j\right|^2+\alpha_C(1-\lambda)^2\right)^{1/2}}\right)\\
&\left.\left.+q_i^{B,v}q_j^{B,v}\left(\frac{1}{\left(\left|\mathbf{r}_i'-\mathbf{r}_j'\right|^2+\alpha_C(1-\lambda)^2\right)^{1/2}}\right)\right\}\right] \qquad (29)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial V^{\mathrm{rf}}}{\partial\lambda}=\frac{1}{4\pi\varepsilon_0\varepsilon_{\mathrm{cs}}}\sum_{i=1}^{N-1}\sum_{\substack{j>i\\ j\ \text{inside cut-off}\ i}}^{N}&\left[-(1-\lambda)^n\lambda\alpha_C\right.\\
&\times\left\{\left(q_i^A-q_i^{A,v}\right)\left(q_j^A-q_j^{A,v}\right)\left(-\frac{\frac{3}{2}C_{\mathrm{rf}}|\mathbf{r}_i-\mathbf{r}_j|^2}{\left(R_{\mathrm{rf}}^2+\alpha_C\lambda^2\right)^{5/2}}\right)\right.\\
&+\left(q_i^A-q_i^{A,v}\right)q_j^{A,v}\left(-\frac{\frac{3}{2}C_{\mathrm{rf}}\left|\mathbf{r}_i-\mathbf{r}_j'\right|^2}{\left(R_{\mathrm{rf}}^2+\alpha_C\lambda^2\right)^{5/2}}\right)\\
&+q_i^{A,v}\left(q_j^A-q_j^{A,v}\right)\left(-\frac{\frac{3}{2}C_{\mathrm{rf}}\left|\mathbf{r}_i'-\mathbf{r}_j\right|^2}{\left(R_{\mathrm{rf}}^2+\alpha_C\lambda^2\right)^{5/2}}\right)\\
&\left.+q_i^{A,v}q_j^{A,v}\left(-\frac{\frac{3}{2}C_{\mathrm{rf}}\left|\mathbf{r}_i'-\mathbf{r}_j'\right|^2}{\left(R_{\mathrm{rf}}^2+\alpha_C\lambda^2\right)^{5/2}}\right)\right\}+\lambda^n(1-\lambda)\alpha_C\\
&\times\left\{\left(q_i^B-q_i^{B,v}\right)\left(q_j^B-q_j^{B,v}\right)\left(-\frac{\frac{3}{2}C_{\mathrm{rf}}|\mathbf{r}_i-\mathbf{r}_j|^2}{(R_{\mathrm{rf}}^2+\alpha_C(1-\lambda)^2)^{5/2}}\right)\right.\\
&+\left(q_i^B-q_i^{B,v}\right)q_j^{B,v}\left(-\frac{\frac{3}{2}C_{\mathrm{rf}}\left|\mathbf{r}_i-\mathbf{r}_j'\right|^2}{\left(R_{\mathrm{rf}}^2+\alpha_C(1-\lambda)^2\right)^{5/2}}\right)\\
&+q_i^{B,v}\left(q_j^B-q_j^{B,v}\right)\left(-\frac{\frac{3}{2}C_{\mathrm{rf}}\left|\mathbf{r}_i'-\mathbf{r}_j\right|^2}{\left(R_{\mathrm{rf}}^2+\alpha_C(1-\lambda)^2\right)^{5/2}}\right)\\
&\left.\left.+q_i^{B,v}q_j^{B,v}\left(-\frac{\frac{3}{2}C_{\mathrm{rf}}\left|\mathbf{r}_i'-\mathbf{r}_j'\right|^2}{\left(R_{\mathrm{rf}}^2+\alpha_C(1-\lambda)^2\right)^{5/2}}\right)\right\}\right]\\
&-n(1-\lambda)^{n-1}\\
&\times\left\{\left(q_i^A-q_i^{A,v}\right)\left(q_j^A-q_j^{A,v}\right)\left(-C_{\mathrm{rf,aux}}(\lambda)|\mathbf{r}_i-\mathbf{r}_j|^2-\frac{1-\frac{1}{2}C_{\mathrm{rf}}}{R_{\mathrm{rf}}}\right)\right.\\
&+\left(q_i^A-q_i^{A,v}\right)q_j^{A,v}\left(-C_{\mathrm{rf,aux}}(\lambda)\left|\mathbf{r}_i-\mathbf{r}_j'\right|^2-\frac{1-\frac{1}{2}C_{\mathrm{rf}}}{R_{\mathrm{rf}}}\right)\\
&+q_i^{A,v}\left(q_j^A-q_j^{A,v}\right)\left(-C_{\mathrm{rf,aux}}(\lambda)\left|\mathbf{r}_i'-\mathbf{r}_j\right|^2-\frac{1-\frac{1}{2}C_{\mathrm{rf}}}{R_{\mathrm{rf}}}\right)\\
&\left.+q_i^{A,v}q_j^{A,v}\left(-C_{\mathrm{rf,aux}}(\lambda)\left|\mathbf{r}_i'-\mathbf{r}_j'\right|^2-\frac{1-\frac{1}{2}C_{\mathrm{rf}}}{R_{\mathrm{rf}}}\right)\right\}+n\lambda^{n-1}\\
&\times\left\{\left(q_i^B-q_i^{B,v}\right)\left(q_j^B-q_j^{B,v}\right)\left(-C_{\mathrm{rf,aux}}(1-\lambda)|\mathbf{r}_i-\mathbf{r}_j|^2-\frac{1-\frac{1}{2}C_{\mathrm{rf}}}{R_{\mathrm{rf}}}\right)\right.\\
&+\left(q_i^B-q_i^{B,v}\right)q_j^{B,v}\left(-C_{\mathrm{rf,aux}}(1-\lambda)\left|\mathbf{r}_i-\mathbf{r}_j'\right|^2-\frac{1-\frac{1}{2}C_{\mathrm{rf}}}{R_{\mathrm{rf}}}\right)\\
&+q_i^{B,v}\left(q_j^B-q_j^{B,v}\right)\left(-C_{\mathrm{rf,aux}}(1-\lambda)\left|\mathbf{r}_i'-\mathbf{r}_j\right|^2-\frac{1-\frac{1}{2}C_{\mathrm{rf}}}{R_{\mathrm{rf}}}\right)\\
&\left.\left.+q_i^{B,v}q_j^{B,v}\left(-C_{\mathrm{rf,aux}}(1-\lambda)\left|\mathbf{r}_i'-\mathbf{r}_j'\right|^2-\frac{1-\frac{1}{2}C_{\mathrm{rf}}}{R_{\mathrm{rf}}}\right)\right\}\right]\\
&-\frac{1}{4\pi\varepsilon_0\varepsilon_{\mathrm{cs}}}\sum_{i=1}^{N}\left(q_i^B-q_i^A\right)\left(\lambda q_i^B+(1-\lambda)q_i^A\right)\frac{\left(1-\frac{1}{2}C_{\mathrm{rf}}\right)}{R_{\mathrm{rf}}}. \qquad (30)
\end{aligned}
$$

Using

$$\alpha_i(\lambda)=(1-\lambda)^n\alpha_i^A+\lambda^n\alpha_i^B \qquad (31)$$

$$E_i(\lambda)=(1-\lambda)^nE_i^A(\lambda)+\lambda^nE_i^B(\lambda) \qquad (32)$$

$$E_{0,i}(\lambda)=(1-\lambda)^nE_{0,i}^A+\lambda^nE_{0,i}^B \qquad (33)$$

one gets for the λ-dependent V^{self} term

$$
V_i^{\mathrm{self}}=\begin{cases}\frac{1}{2}\alpha_i(\lambda)E_i(\lambda)^2 & \text{for } E_i(\lambda)\le E_{0,i}(\lambda)\\ \frac{1}{2}\alpha_i(\lambda)E_{0,i}(\lambda)^2+\frac{\alpha_i(\lambda)E_{0,i}(\lambda)^2}{p_i(p_i-1)}\\ \left[-p_i^2+\left(p_i^2-1\right)\left(\frac{E_i(\lambda)}{E_{0,i}(\lambda)}\right)+\left(\frac{E_{0,i}(\lambda)}{E_i(\lambda)}\right)^{p_i-1}\right]. & \text{for } E_i(\lambda)>E_{0,i}(\lambda)\end{cases} \qquad (34)
$$

The λ-derivative of this perturbed self-polarization potential energy is

$$\frac{\partial V_i^{\text{self}}}{\partial \lambda} = \begin{cases} \frac{1}{2}\frac{\partial \alpha_i(\lambda)}{\partial \lambda} E_i(\lambda)^2 + \alpha_i(\lambda) E_i(\lambda) \frac{\partial E_i(\lambda)}{\partial \lambda} & \text{for } E_i(\lambda) \le E_{0,i}(\lambda) \\ \frac{1}{2}\frac{\partial \alpha_i(\lambda)}{\partial \lambda} E_{0,i}(\lambda)^2 + \alpha_i(\lambda) E_{0,i}(\lambda) \frac{\partial E_{0,i}(\lambda)}{\partial \lambda} & \text{for } E_i(\lambda) > E_{0,i}(\lambda) \\ + \frac{1}{p_i(p_i-1)} \left[\frac{\partial \alpha_i(\lambda)}{\partial \lambda} E_{0,i}(\lambda)^2 + 2\alpha_i(\lambda) E_{0,i}(\lambda) \frac{\partial E_{0,i}(\lambda)}{\partial \lambda} \right] & \\ \left[-p_i^2 + (p_i^2 - 1) \left(\frac{E_i(\lambda)}{E_{0,i}(\lambda)} \right) + \left(\frac{E_{0,i}(\lambda)}{E_i(\lambda)} \right)^{p_i - 1} \right] & \\ + \frac{\alpha_i(\lambda) E_{0,i}(\lambda)^2}{p_i(p_i-1)} \left[(p_i^2 - 1) \frac{E_{0,i}(\lambda) \frac{\partial E_i(\lambda)}{\partial \lambda} - E_i(\lambda) \frac{\partial E_{0,i}(\lambda)}{\partial \lambda}}{E_{0,i}(\lambda)^2} \right. & \\ \left. + (p_i - 1) \left(\frac{E_{0,i}(\lambda)}{E_i(\lambda)} \right)^{p_i - 2} \frac{E_{0,i}(\lambda) \frac{\partial E_i(\lambda)}{\partial \lambda} - E_i(\lambda) \frac{\partial E_{0,i}(\lambda)}{\partial \lambda}}{E_i(\lambda)^2} \right]. & \end{cases} \tag{35}$$

Implicit solute surface area and internal volume solvation term

To investigate the structure and dynamics of biological macromolecules in solution with MD simulation, typically an atomic-level representation is used. Inclusion of the solvent degrees of freedom then results in a huge number of degrees of freedom, greatly increasing the cost of the calculations. Complete neglect of the solvent effects, however, greatly limits the accuracy of the simulation.[22] For large-scale or computationally expensive applications such as drug screening, protein structure prediction, protein folding, computational engineering of peptides, and proteins, or aggregation of such molecules, a compromise between the speed of simulations done in vacuum and the accuracy of explicitly representing the solvent may be useful. Implicit solvent models, in which the effect of the solvent is modeled as a mean solvation term in the potential energy function of the solute, provide a means of including solvent effects at minimal cost.[23–26]

A widely used and simple class of implicit solvent models are the so-called solvent accessible surface area (SASA) models in which the local solute–solvent interactions are assumed to be proportional to the SASA of the solute atoms.[27–41] The SASA term should account for the free energy cost of forming a hard-sphere cavity in water. As it has been shown that the cavity creation work depends on both the solvent accessible volume and the SASA, a term proportional to the internal volume, defined as the sum of the volumes of all nonsolvent-exposed atoms, can be added that takes into account interactions between buried atoms and the solvent as well.[8] Another factor missing from existing SASA-based implicit solvent models is the inclusion of the favorable van der Waals interactions between the interior atoms of the solute and the solvent.[42–50] The GROMOS software contains functions that calculate the SASA of a solute according to the approximate expression of Hasel et al.[51] and the buried volume using an approximation[8] based on the SASA.

The solvation free enthalpy of a solute G_{solv} may be partitioned into a solvent–solvent cavity term G_{cav}, a solute–solvent van der Waals term G_{vdW}, and a solute–solvent electrostatic term G_{elec}

$$G_{\text{solv}} = G_{\text{cav}} + G_{\text{vdW}} + G_{\text{elec}}. \tag{36}$$

Here, as in the original SASA model built on the GROMOS force field,[37] rather than calculate G_{elec} explicitly, the vacuum parameterization for the intrasolute interaction function is used: the partial charges of groups of atoms with nonzero total charge are changed such that these groups have zero charge, but their hydrogen-bonding capacity is maintained.[2] G_{cav} and G_{vdW} are combined into a mean solvation term, such that the total interaction function for a protein solute containing N atoms with Cartesian coordinates $\mathbf{r}^N = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ consists of two parts

$$V(\mathbf{r}^N) = V_{\text{phys}}(\mathbf{r}^N) + V_{\text{solv}}(\mathbf{r}^N), \tag{37}$$

where $V_{\text{phys}}(\mathbf{r}^N)$ is determined by a standard GROMOS force field for vacuum simulations.[2] The mean solvation term may be further split

$$V_{\text{solv}}(\mathbf{r}^N) = V_{\text{solv}}^{\text{SASA}}(\mathbf{r}^N) + V_{\text{solv}}^{\text{VOL}}(\mathbf{r}^N). \tag{38}$$

Each term in eq. (38) is calculated as a sum of atomic contributions. The first term is that of the original GROMOS SASA model[37]

$$V_{\text{solv}}^{\text{SASA}}(\mathbf{r}^N) = \sum_{i=1}^{N} \sigma_i^{\text{SASA}} A_i(\mathbf{r}^N). \tag{39}$$

There are just two types of atom-specific solvation parameters σ_i, for hydrophilic ($\sigma_{\text{O,N}}^{\text{SASA}}$) and hydrophobic ($\sigma_{\text{C}}^{\text{SASA}}$) atoms.[8] The SASA of atom i, A_i, is defined using the approximate analytical expression of Hasel et al.[51]

$$A_i(\mathbf{r}^N) = S_i \prod_{\substack{j=1 \\ j \ne i}}^{N} \left[1 - p_i p_{ij} \frac{b_{ij}(r_{ij})}{S_i} \right]. \tag{40}$$

The total surface area of an isolated atom i with radius R_i accessible to a solvent probe atom with radius R_{solv} is given by

$$S_i = 4\pi [R_i + R_{\text{solv}}]^2 \tag{41}$$

and the overlap reduction factor[52] b_{ij} for atoms i and j at a distance $r_{ij} = ((\mathbf{r}_i - \mathbf{r}_j)^2)^{1/2}$ is given by

$$b_{ij}(r_{ij}) = \begin{cases} 0 & \text{if } r_{ij} \ge R_i + R_j + 2R_{\text{solv}} \\ \pi [R_i + R_{\text{solv}}][R_i + R_j + 2R_{\text{solv}} - r_{ij}] \cdot \left[1 + \frac{R_j - R_i}{r_{ij}} \right] & \text{if } 0 < r_{ij} < R_i + R_j + 2R_{\text{solv}}. \end{cases} \tag{42}$$

The parameters p_i and p_{ij} were optimized by Hasel et al.[51] for $R_{solv} = 0.14$ nm. Their mapping onto GROMOS atom types, along with the corresponding R_i values, are given in Ref. 8.

The SASA values A_i obtained in eq. (40) are used in a switching function $g(A)$ to select the atoms of the solute that have $A_i = 0$ to contribute to the $V_{solv}^{VOL}(\mathbf{r}^N)$ term of eq. (38). This term accounts for the contribution of the interior solute atoms to the solvation free energy[45]

$$V_{solv}^{VOL}(\mathbf{r}^N) = \sum_{i=1}^{N} \sigma_i^{VOL} g\left(A_i(\mathbf{r}^N); A_i^{s1}, A_i^{s2}\right) \frac{4}{3}\pi R_i^3 \tag{43}$$

The parameters A_i^{s1} and A_i^{s2} switch g from 1 to 0 for $A_i^{s1} \le A_i \le A_i^{s2}$, so that only atoms with $A_i \le A_i^{s2}$ contribute to the buried volume. The conditions for the choice of switching function are (with $0 \le A_i^{s1} \le A_i^{s2}$)

$$\begin{aligned} g\left(A_i^{s1}\right) &= 1, \quad \frac{dg}{dA_i}\left(A_i^{s1}\right) = 0 \\ g\left(A_i^{s2}\right) &= 0, \quad \frac{dg}{dA_i}\left(A_i^{s2}\right) = 0. \end{aligned} \tag{44}$$

The following switching function was chosen

$$g\left(A_i; A_i^{s1}, A_i^{s2}\right) = \begin{cases} 1 & 0 \le A_i \le A_i^{s1} \\ \frac{2\left[A_i - A_i^{s1}\right]^3}{\left[A_i^{s2} - A_i^{s1}\right]^3} - \frac{3\left[A_i - A_i^{s1}\right]^2}{\left[A_i^{s2} - A_i^{s1}\right]^2} + 1 & A_i^{s1} \le A_i \le A_i^{s2}. \\ 0 & A_i^{s2} \le A_i \end{cases} \tag{45}$$

This differentiable approximation of a step function gives the value 1, if an atom has a very small or zero A_i, that is, the atom is in the middle of the solute, and a value of zero if an atom is at the surface, that is, A_i is large and thus the atom should not contribute to the buried volume term.

The choices of the parameters that can be freely adjusted in the SASA/VOL model, namely A_i^{s1}, A_i^{s2}, σ_i^{SASA}, and σ_i^{VOL}, are discussed in Ref. 8 where the spatial derivatives of the equations constituting the mean solvation energy term, which are required to compute the forces, are given in the Supporting Information.

Coarse-grained supramolecular force field terms

For very large systems, for example, protein complexes in aqueous solution, it may be efficient to represent the many solvent degrees of freedom at a supramolecular level. For such a CG solvent, several solvent molecules are put into one CG bead, particle or CG molecule, leading to faster simulations due to a reduction of the degrees of freedom compared to fine-grained molecular models and simulations. A CG model for liquid water should represent van der Waals as well as electrostatic interactions. A CG water model[9] that is compatible with the GROMOS force fields has one van der Waals interaction site CP and two electrostatic interaction sites, CP and dipole particle (DP), and one such a CG water molecule represents five real, fine-grained water molecules, thereby allowing orders of magnitude faster simulations. The first interaction site, called CP, interacts with other CG beads through a pairwise Lennard–Jones and an electrostatic potential energy function. The second interaction site is a so-called DP which carries a charge opposite to that of the central site and interacts purely electrostatically, that is, has no Lennard–Jones interaction, with the charged sites of other CG beads. The nonbonded potential energy between two sites i and j depends on their distance r_{ij} and has three contributions

$$V_{ij}^{LJ}(r_{ij}) = \left(\frac{C_{12}(i,j)}{r_{ij}^6} - C_6(i,j)\right)\frac{1}{r_{ij}^6} \tag{46}$$

$$V_{ij}^{coul}(r_{ij}) = \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_{cs}}\frac{1}{r_{ij}} \tag{47}$$

$$V_{ij}^{rf}(r_{ij}) = -\frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_{cs}}\left(\frac{\frac{1}{2}C_{rf}\cdot r_{ij}^2}{R_{rf}^3} + \frac{1-\frac{1}{2}C_{rf}}{R_{rf}}\right), \tag{48}$$

where $R_{rf} = R_c$ are the cut-off and reaction-field radii, $C_{12}(i,j)$ and $C_6(i,j)$ are the Lennard–Jones parameters, q_i is the charge of particle i, and C_{rf} is the Coulomb reaction-field constant defined in eq. (20).

The total nonbonded potential energy is obtained by summing over all ordered pairs (i,j), thereby taking into account the different nearest neighbor and cut-off radius exclusions,[11]

$$\begin{aligned} V^{nb}(\mathbf{r}^N) = &\sum_{i=1}^{N-1} \sum_{\substack{j>i \\ j \text{ inside cut-off } i \\ (i,j) \text{ not excluded}}}^{N} \left(V_{ij}^{LJ}(r_{ij}) + V_{ij}^{coul}(r_{ij})\right) \\ &+ \sum_{i=1}^{N-1} \sum_{\substack{j>i \\ j \text{ inside cut-off } i}}^{N} V_{ij}^{rf}(r_{ij}) - \sum_{i=1}^{N} \frac{q_i^2}{4\pi\varepsilon_0\varepsilon_{cs}}\frac{1}{2}\left(\frac{1-\frac{1}{2}C_{rf}}{R_{rf}}\right), \end{aligned} \tag{49}$$

where the last term is a constant added to represent the self-interaction of the charged particles. For atomistic models, the dielectric permittivity inside the cut-off sphere $\varepsilon_{cs} = 1$ whereas the CG model has $\varepsilon_{cs} > 1$, representing the dielectric response of the interior of a CG bead. The two interaction sites i and j of a bead that do not interact mutually through V_{ij}^{LJ} and V_{ij}^{coul} are connected by an unconstrained bond with a half attractive quartic potential energy function

$$V_{bond}^{CG}(r_{ij}) = \frac{1}{2}K_b(r_{ij} - r_{CP-DP})^4 \tag{50}$$

for $r_{ij} > r_{CP-DP}$, where r_{CP-DP} sets a bound on the distance r_{ij} between particles i and j belonging to one CG bead. The use of a quartic instead of a harmonic attractive bond-length potential energy function makes the polarizability nonlinear[21] and has the advantage that overpolarization due to close distances between charges of different molecules is avoided as a result of the superlinear increase of the forces for longer bond lengths. The polarizability of the model resides in the free rotation of the DP particle around the central CP site and the oscillation of the

bond connecting them, thus adjusting the dipole moment of a CG bead to the electric field of the environment.

Multiplicative switching function

When using a cut-off radius for nonbonded interactions, the error due to the discontinuity of the force at the cut-off can be masked by the use of a switching function that smoothly switches the force and energy to zero at the cut-off distance between two particles. To be able to switch or shift arbitrary functions of r, that is, not only negative powers of r, such as soft-core or reaction-field terms, one should use a multiplicative switching function that serves as a prefactor of the total nonbonded interaction function $V^{\mathrm{nb}}(r)$, and avoid the use of additive switching functions[5] that are added to each term of the nonbonded interaction function. Thus, in GROMOS the nonbonded interaction may be smoothly switched off at R_c according to

$$V(r) = S(r) \cdot V^{\mathrm{nb}}(r). \tag{51}$$

The function $S(r)$ equals 1 for $r \leq R_s$, where R_s is the switching radius, and equals 0 for $r \geq R_c$, where R_c is the cut-off radius. In the range $R_s \leq r \leq R_c$ it should satisfy the conditions

$$S(R_s) = 1, S'(R_s) = S''(R_s) = 0 \tag{52}$$
$$S(R_c) = S'(R_c) = S''(R_c) = 0, \tag{53}$$

where the notations $S' = dS/dr$ and $S'' = d^2S/dr^2$ are used for the derivatives. This is because we have

$$V'(r) = S'(r) \cdot V^{\mathrm{nb}}(r) + S(r) \cdot V^{\mathrm{nb}'}(r) \tag{54}$$

and

$$V''(r) = S''(r) \cdot V^{\mathrm{nb}}(r) + 2S'(r) \cdot V^{\mathrm{nb}'}(r) + S(r) \cdot V^{\mathrm{nb}''}(r). \tag{55}$$

A switching function that satisfies conditions in eq. (53) and has simple coefficients in its first derivative form, which is used to calculate forces, is

$$S(r) = -\frac{1}{3}A(R_c - r)^3 - \frac{1}{4}B(R_c - r)^4 - \frac{1}{5}C(R_c - r)^5, \tag{56}$$

with

$$S'(r) = A(R_c - r)^2 + B(R_c - r)^3 + C(R_c - r)^4 \tag{57}$$

and

$$S''(r) = -2A(R_c - r) - 3B(R_c - r)^2 - 4C(R_c - r)^3. \tag{58}$$

The constants A, B, and C can be chosen such that conditions in eq. (52) are satisfied and will only depend on R_s and R_c. We find

$$A = \frac{-30}{(R_c - R_s)^3}, \quad B = \frac{60}{(R_c - R_s)^4}, \quad C = \frac{-30}{(R_c - R_s)^5}, \tag{59}$$

which leads to

$$S(r) = \frac{(R_c - r)^3}{(R_c - R_s)^3}\left[10 - 15\frac{(R_c - r)}{(R_c - R_s)} + 6\frac{(R_c - r)^2}{(R_c - R_s)^2}\right] \tag{60}$$

$$S'(r) = -30\frac{(R_c - r)^2}{(R_c - R_s)^3}\left[1 - 2\frac{(R_c - r)}{(R_c - R_s)} + \frac{(R_c - r)^2}{(R_c - R_s)^2}\right] \tag{61}$$

$$S''(r) = 60\frac{(R_c - r)}{(R_c - R_s)^3}\left[1 - 3\frac{(R_c - r)}{(R_c - R_s)} + 2\frac{(R_c - r)^2}{(R_c - R_s)^2}\right] \tag{62}$$

for the range $R_s \leq r \leq R_c$,

$$S(r) = 1, S'(r) = S''(r) = 0 \quad \text{for } 0 \leq r \leq R_s, \tag{63}$$
$$S(r) = S'(r) = S''(r) = 0 \quad \text{for } r \geq R_c. \tag{64}$$

This switching function replaces the one described in Ref. 5.

Adiabatic decoupling of a subset of degrees of freedom to enhance sampling

If one is interested in artificially enhancing the sampling of a particular, small subset of the degrees of freedom of a system, one may adiabatically decouple their motion from the other degrees of freedom in the system and then use high temperature or force scaling to better sample the subset of degrees of freedom.[10]

A particular set of $3N^h$ degrees of freedom can be adiabatically decoupled from the other $3N^l$ degrees of freedom of the system by increasing the mass of the N^h h type particles,

$$m^h = s_m m^l \tag{65}$$

with $s_m \gg 1$ in such a way that the transfer of kinetic energy between the decoupled degrees of freedom through momentum conserving collisions becomes very small. s_m has to be chosen in such a way that the lightest mass of the N^h atoms, usually a hydrogen, is made significantly larger than the heaviest mass of the N^l atoms. It is therefore recommended to choose $s_m \geq 100$. Here, we have denoted the masses of the N^h h type particles collectively by m^h and the ones of the N^l remaining atoms collectively by m^l. This decoupling is more easily achieved[10] if $N^h \ll N^l$. The nondynamic equilibrium properties of the system remain unaltered by this change of mass of a subset of degrees of freedom.

To enhance the sampling of the h type degrees of freedom, their temperature, potential energy, or force can be scaled.[10] Applying temperature scaling, the temperature T^h of the h type degrees of freedom is set to

$$T^h = s_T T^l \text{ or } \beta^h = \beta^l / s_T \tag{66}$$

where $\beta = (k_B T)^{-1}$ and k_B is the Boltzmann constant. The Hamiltonian for a system of N^h h type particles and N^l l type particles is then

$$H(p^h, p^l, x^h, x^l) = K^h(p^h) + V^{hh}(x^h) + V^{hl}(x^h, x^l) + K^l(p^l) + V^{ll}(x^l), \tag{67}$$

where $K^h(p^h)$ and $K^l(p^l)$ are the kinetic energy of the h type and the l type particles, respectively, $V^{hh}(x^h)$ is the potential energy between h type particles, $V^{ll}(x^l)$ is the potential energy between l type particles and $V^{hl}(x^h, x^l)$ that between h and l type particles. In the adiabatically decoupled limit, the partition function of the l type particles is

$$Z^l(x^h;\beta^l) = \int e^{-\beta^l K^l(p^l)} dp^l \int e^{-\beta^l V^{ll}(x^l)} e^{-\beta^l V^{hl}(x^h,x^l)} dx^l, \quad (68)$$

and the Hamiltonian of the h type degrees of freedom can therefore be written as

$$H^h(p^h, x^h;\beta^l) = K^h(p^h) + V^{hh}(x^h) - (\beta^l)^{-1} \ln Z^l(x^h;\beta^l). \quad (69)$$

The average value of a property $Q(x^h)$ at temperature T^l, $\langle Q\rangle_{\beta^l}$, can then be calculated from $\langle Q\rangle_{\beta^h}$ of $Q(x^h)$ at T^h using the standard unbiasing formula,

$$\langle Q\rangle_{\beta^l} = \frac{\langle Q e^{-(\beta^l-\beta^h)[V^{hh}-(\beta^l)^{-1}\ln Z^l]}\rangle_{\beta^h}}{\langle e^{-(\beta^l-\beta^h)[V^{hh}-(\beta^l)^{-1}\ln Z^l]}\rangle_{\beta^h}}, \quad (70)$$

where the ensemble averages are over the h type degrees of freedom, and

$$Z^l(x^h;\beta^l) = \frac{\int e^{-\beta^l K^l(p^l)} dp^l \int e^{-\beta^l V^{ll}(x^l)} dx^l}{\langle e^{+\beta^l V^{hl}(x^h)}\rangle}, \quad (71)$$

where the ensemble average in the denominator is over the l type degrees of freedom[10] at T^l.

Applying force scaling,[10] the force acting on a h type degree of freedom is scaled by a factor s_V,

$$\mathbf{f}_h = (\mathbf{f}_{hh} + \mathbf{f}_{hl}) s_V, \quad (72)$$

where $\mathbf{f}_{hh}$ is the force exerted by the other h type degrees of freedom and $\mathbf{f}_{hl}$ is the force exerted by the l type degrees of freedom.

The Hamiltonian, in analogy to eq. (69), can be written as

$$H^h(p^h, x^h) = K^h(p^h) + s_V V^{hh}(x^h) - s_V(\beta^l)^{-1} \ln Z^l(x^h;\beta^l), \quad (73)$$

with $Z^l(x^h;\beta^l)$ defined as in eq. (71).

The average value of a property $Q(x^h)$ at T^l, $\langle Q\rangle_{\beta^l}$, can then be calculated from $\langle Q\rangle_{\beta^h}$ of $Q(x^h)$ at $\beta^h = \beta^l s_V$ using the standard unbiasing formula eq. (70).

External forces and nonequilibrium simulation

External electric field. The application of an external electric field enables a simple and fast way to probe the static dielectric properties of a system like the dielectric permittivity $\varepsilon(0)$ or the Debye relaxation time τ_D.[11]

If a constant homogeneous external field $\mathbf{E}^{\text{ext}}$ is applied to a system, for examples along the z-axis,

$$\mathbf{E}^{\text{ext}} = E^{\text{ext}} \mathbf{e}_z, \quad (74)$$

an additional force on particle i is present,

$$\mathbf{f}_i^{\text{ext,cs}} = \frac{q_i}{4\pi\varepsilon_0} \mathbf{E}^{\text{ext,cs}} = \frac{q_i}{4\pi\varepsilon_0} (\mathbf{E}^{\text{ext}} + \mathbf{E}^{\text{ext,rf}}), \quad (75)$$

where $\mathbf{E}^{\text{ext,rf}}$ describes the contribution to the electric field $\mathbf{E}^{\text{ext,cs}}$ inside the cut-off sphere due to the polarization $\mathbf{P}^{\text{ext,rf}}$ in the dielectric continuum outside the cut-off sphere that is induced by the difference in dielectric permittivity inside, ε_{cs}, and outside, ε_{rf}, the cut-off sphere. This contribution depends on the shape of the cut-off region.[53] For a spherical region, we have

$$\begin{aligned} \mathbf{E}^{\text{ext,cs}} &= \mathbf{E}^{\text{ext}} - \frac{\varepsilon_{\text{cs}} - \varepsilon_{\text{rf}}}{\varepsilon_{\text{cs}} + 2\varepsilon_{\text{rf}}} \mathbf{E}^{\text{ext}} \\ &= \mathbf{E}^{\text{ext}} + \frac{4\pi}{\varepsilon_{\text{cs}} + 2\varepsilon_{\text{rf}}} \mathbf{P}^{\text{ext,rf}} \\ &= \frac{3\varepsilon_{\text{rf}}}{\varepsilon_{\text{cs}} + 2\varepsilon_{\text{rf}}} \mathbf{E}^{\text{ext}}, \end{aligned} \quad (76)$$

where the polarization induced around the cut-off cavity in the dielectric continuum by the external field is

$$\mathbf{P}^{\text{ext,rf}} = \frac{\varepsilon_{\text{rf}} - \varepsilon_{\text{cs}}}{4\pi} \mathbf{E}^{\text{ext}}. \quad (77)$$

The electric field $\mathbf{E}^{\text{ext}}$ that is applied to the system will induce a polarization $\mathbf{P}$ in the system,

$$\mathbf{P} = \frac{1}{4\pi} (\varepsilon(0) - 1) \mathbf{E}^{\text{ext}}, \quad (78)$$

where $\varepsilon(0)$ is the zero-frequency or static dielectric permittivity of the molecular model. The polarization $\mathbf{P}$ can be calculated from the total dipole moment $\mathbf{M}$ of the system and its volume V,

$$\mathbf{P} = \frac{\mathbf{M}}{V} = \frac{1}{V} \sum_{i=1}^{N} q_i \mathbf{r}_i \quad (79)$$

if $\sum_{i=1}^{N} q_i = 0$. If the latter condition is not satisfied, $\mathbf{M}$ and thus $\mathbf{P}$ will be origin dependent. For a homogeneous external field, eq. (74) along the z-axis we then find

$$\varepsilon(0) = 1 + 4\pi \frac{\langle M_z\rangle}{V E_z^{\text{ext}}}, \quad (80)$$

where $\langle M_z\rangle$ is the average dipole moment of the volume V in the z-direction in the simulation.

It is also possible to obtain a value of τ_D from averaging over a set of nonequilibrium MD simulations that start from a Boltzmann distributed set of initial configurations and velocities and in which a homogeneous static external electric field $\mathbf{E}^{\text{ext}}$ is

switched on at $t = t_0$. On switching on $\mathbf{E}^{\text{ext}}$ along the z-axis at $t = t_0$, the z-component M_z of $\mathbf{M}$ will increase from its initial value $M_z(t_0)$, which values are Gaussian distributed around $M_z = 0$, to a steady state value $M_z(t = \infty)$. For a Debye dielectric medium, this build up will be exponential,

$$\langle M_z(t)\rangle_{t_0} = \langle M_z(t = \infty)\rangle_{t_0}[1 - e^{-(t-t_0)/\tau_M}]. \tag{81}$$

The value of $\langle M_z(t = \infty)\rangle_{t_0}$ will be larger for larger E_z^{ext}, but different field strengths E_z^{ext} should yield the same τ_M, as long as E_z^{ext} is not too small and not too large. The relation between τ_M and τ_D is

$$\tau_D = \frac{\varepsilon(0) + 2 + C_{\text{rf}}(\varepsilon(0) - 1)}{3}\tau_M, \tag{82}$$

which is a generalization for $\kappa \geq 0$ of the relation for $\kappa = 0$ given by Neumann.[54] The value of $\varepsilon(0)$ can be calculated using eq. (80), in which $\langle M_z\rangle = \langle M_z(t = \infty)\rangle_{t_0}$.

Sinusoidal applied acceleration field. In the GROMOS software, a sinusoidal, periodic applied acceleration field[55] can be used to calculate the viscosity of a liquid.

The Navier–Stokes equation provides a continuum description of the motion of a fluid with mass density ρ,

$$\rho\frac{\partial \mathbf{u}}{\partial t} + \rho(\mathbf{u}\cdot\nabla)\mathbf{u} = -\nabla p + \eta\nabla^2\mathbf{u} + \mathbf{f} \tag{83}$$

in which on the left side, the material derivative of the velocity field describes the unsteady and convective accelerations per volume, and on the right side the first term corresponds to the pressure gradient, the second to the viscosity, and $\mathbf{f}$ represents additional forces that can be imposed on the system. If an external shear-stress acceleration field $\mathbf{a}$ is applied, the equation becomes

$$\rho\frac{\partial \mathbf{u}}{\partial t} + \rho(\mathbf{u}\cdot\nabla)\mathbf{u} = -\nabla p + \eta\nabla^2\mathbf{u} + \rho\mathbf{a}. \tag{84}$$

Assume that the acceleration field $\mathbf{a}$ is applied in the x-direction and has a sinusoidal dependence on the z-coordinate of the particle i such that

$$a_{i,x}(z) = A\cos(nkz_i), \tag{85}$$

where $k = 2\pi/l_z$, A is the applied acceleration amplitude, l_z is the box length in the z direction, z_i is the z-coordinate of particle i and n an integer with $n > 0$. As the Navier–Stokes equation constitutes a continuum description, the approach is only valid in the long wave length limit, where the wavelength of the external field is large compared to the molecular dimensions.[56] Thus, the natural choice is to take n equal to 1. The box size should also be chosen in such a way that l_z is big compared to the size of the individual molecules. Given the nature of the perturbation, the average velocities along y- and z-directions (u_y and u_z) vanish, and therefore, no pressure gradient in the x-direction is present, which simplifies eq. (84), leading to an equation for $u_x(z)$ of the form

$$\rho\frac{\partial u_x(z)}{\partial t} = \rho a_x(z) + \eta\frac{\partial^2 u_x(z)}{\partial z^2}. \tag{86}$$

At the steady-state

$$\frac{\partial u_x(z)}{\partial t} = 0, \tag{87}$$

and the steady-state solution of eq. (86) is given by

$$a_x(z) + \frac{\eta}{\rho}\frac{\partial^2 u_x(z)}{\partial z^2} = 0. \tag{88}$$

By applying an acceleration field of the form of eq. (85) with $n = 1$ in eq. (88) and assuming that $u_x(z) = 0$ at $t = 0$, the resulting velocity field has the form

$$u_x(z, t) = V(1 - e^{-t/\tau_r})\cos(kz) \tag{89}$$

where

$$V = \frac{A}{\eta}\frac{\rho}{k^2}, \tag{90}$$

and where the relaxation time τ_r is given by

$$\tau_r = \frac{\rho}{\eta k^2}. \tag{91}$$

For t sufficiently larger than τ_r one has

$$u_x(z) = V\cos(kz). \tag{92}$$

The calculation of V can be easily performed on the fly, for example, at every few time steps, using

$$V(t) = \frac{2}{N}\sum_i^N v_{i,x}(t)\cos(kz_i(t)). \tag{93}$$

where $v_{i,x}(t)$ is the velocity of particle i in the acceleration direction x. By monitoring the time series of $V(t)$, an equilibration part is observed, which has to be excluded from the averaging. From the average of $V(t)$, one can calculate the viscosity using eq. (90). Error estimates can be obtained from the fluctuations of $V(t)$ and applying block averaging. The amplitude A of the applied acceleration should not be chosen too small, which would lead to a low signal-to-noise ratio, and not too big, which would lead to a disturbance of the equilibrium properties of the liquid, temperature drift, and nonlinear responses.[56]

Examples of the use of the simulation software

Mixtures of polarizable water and polarizable ethylene glycol.

The necessity to use polarizability is illustrated with a water–ethylene glycol mixture, described in Ref. 57. As shown in Figure 1, using different nonpolarizable force fields for ethylene glycol (blue lines) in water the maximum in the experimental solvation

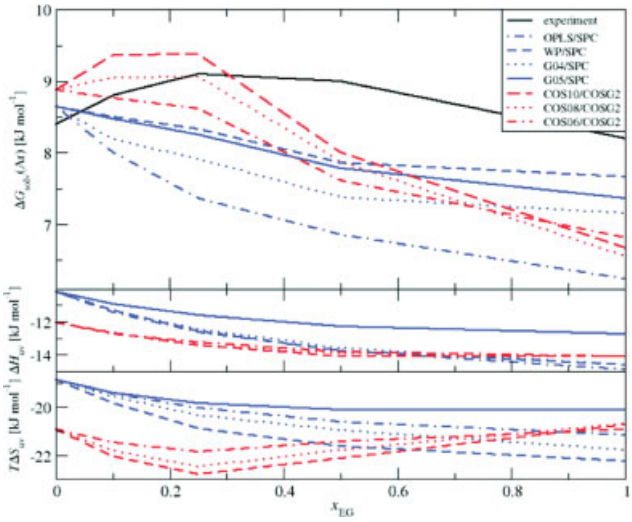

Figure 1. Solvation free enthalpies ΔG_{solv} (upper panel), solute–solvent interaction enthalpies ΔH_{uv} (middle panel) and solute–solvent interaction entropies $T\Delta S_{uv}$ (lower panel) of argon in ethylene glycol–water mixtures at different mole fractions of ethylene glycol (x_{EG}).[57]

Table 1. Parameters of the coarse-grained models of DMSO and MeOH.

Model	m_{CP} u	m_{DP} u	q_{CP} e	r_0 nm	K_b 10^6 mol^{-1} nm^{-4}	σ_{LJ} nm	ε_{LJ} kJ mol^{-1}	ε_{CS}
DMSO	78.2588	78	0.642	0.2	20	0.542	2.44	2.5
MeOH	64.1696	64	0.56	0.2	2.0	0.56	2.1	2.5

m_{CP}, mass of the CP particle of the bead; m_{DP}, mass of the DP particle of the bead; q_{CP}, charge of the CP particle of the bead; r_0, ideal distance between CP and DP; K_b, force constant between CP and DP; σ_{LJ}, Lennard–Jones parameter; ε_{LJ}, Lennard–Jones parameter; and ε_{CS}, dielectric permittivity inside the cut-off sphere.

free enthalpies ΔG_{solv} of argon as a function of the ethylene glycol mole fraction x_{EG} is not reproduced. The introduction of polarizability in the ethylene glycol and water models (red lines) leads to a qualitative reproduction of the experimental data. The separation into solute–solvent enthalpic contributions ΔH_{uv} and entropic contributions $T\Delta S_{uv}$ reveals that the occurrence of a maximum in ΔG_{solv} is entropic in nature.

Mixtures of coarse-grained DMSO or MeOH and coarse-grained water

The application of coarse-graining is demonstrated with mixtures of water[9] and methanol (MeOH), and water[9] and dimethylsulfoxide (DMSO). For the CG DMSO systems a cubic box with edge length of 6.82 nm was filled with 2100 DMSO beads leading to

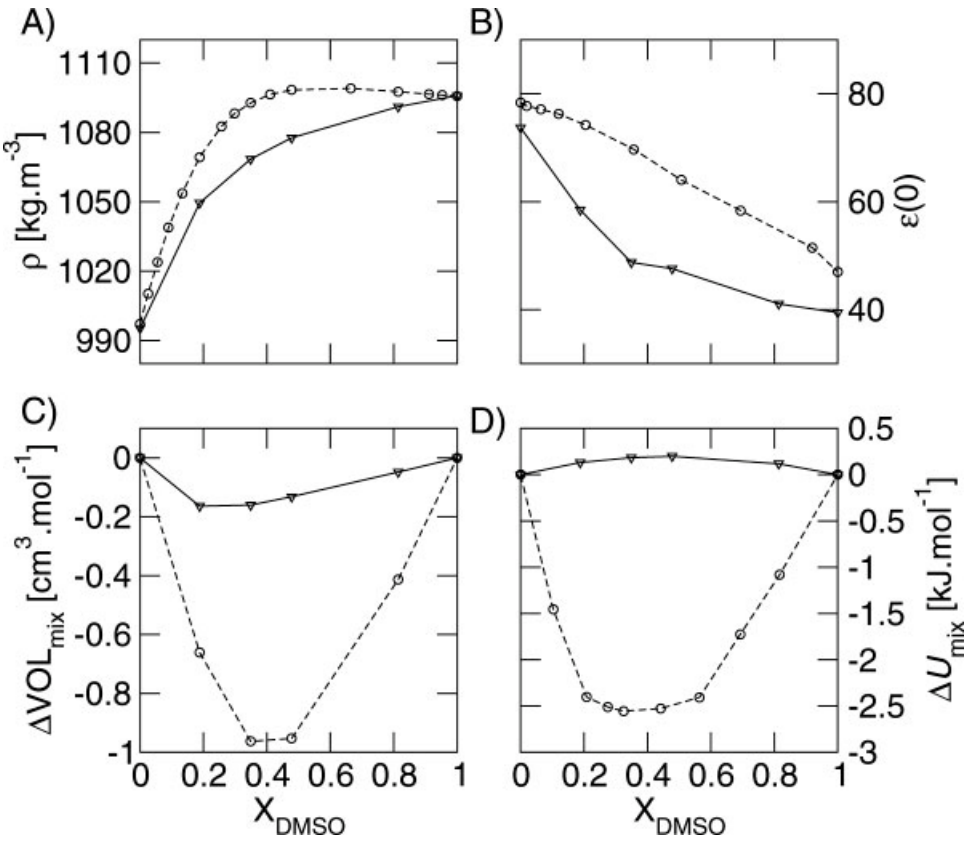

Figure 2. Thermodynamic properties of DMSO : H_2O mixtures at 298 K and 1 atm, as a function of the mole fraction of DMSO, X_{DMSO}, from MD simulations using the CG model of DMSO described in Table 1 and the CG model of H_2O from[9]: (A) densities ρ, (B) dielectric permittivities $\varepsilon(0)$, (C) excess volume of mixing ΔVOL_{mix}, and (D) excess potential energy of mixing ΔU_{mix}. Experimental data are shown with circles and a dashed line, and results of the CG simulations are shown with triangles and a solid line.

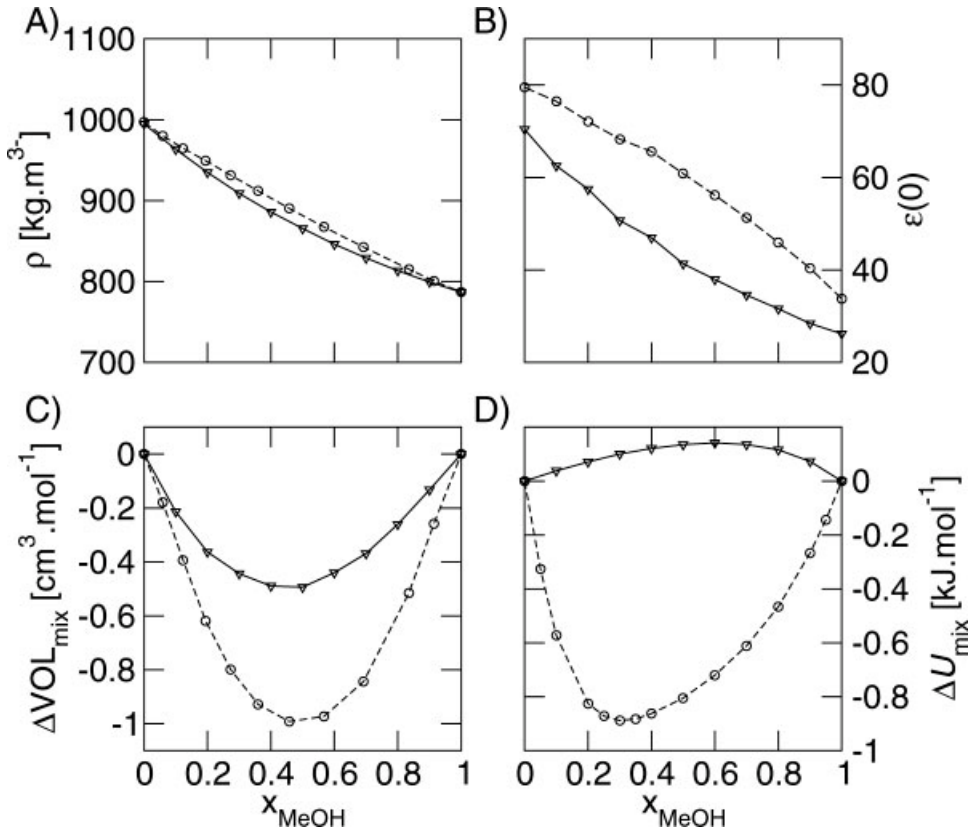

Figure 3. Thermodynamic properties of MeOH : H_2O mixtures at 298 K and 1 atm, as a function of the mole fraction of MeOH, X_{MeOH}, from MD simulations using the CG model of MeOH described in Table 1 and the CG model of H_2O from[9]: (A) densities ρ, (B) dielectric permittivities $\varepsilon(0)$, (C) excess volume of mixing ΔVOL_{mix}, and (D) excess potential energy of mixing ΔU_{mix}. Experimental data are shown with circles and a dashed line, and results of the CG simulations are shown with triangles and a solid line.

a density of 1096 kg/m^3, corresponding to the density of liquid DMSO at 298 K and 1 atm.[58] For the CG MeOH systems a cubic box with edge length of 6.82 nm was filled 1875 MeOH beads leading to a density of 787 kg/m^3, corresponding to the density of liquid MeOH at 298 K and 1 atm.[58]

For all systems, the temperature was weakly coupled[59] to a bath of 298 K with a relaxation time of 0.1 ps. The pressure was weakly coupled[59] to a bath of 1 atm with a relaxation time of 0.5 ps. The isothermal compressibility was set to the experimental value[60] of 7.513×10^{-4} (kJ mol^{-1} nm^{-3})$^{-1}$ for water, 8.718×10^{-4}(kJ mol^{-1} nm^{-3})$^{-1}$ for pure DMSO, 7.5×10^{-4}(kJ mol^{-1} nm^{-3})$^{-1}$ for pure MeOH, and linear combinations for the mixtures. The nonbonded van der Waals and electrostatic interactions were calculated using a cutoff radius of 2 nm. The long range electrostatic interactions beyond the cut-off of 2 nm were represented by a reaction field[54,61] with $\varepsilon_{RF} = 78.5$ for water, 46 for DMSO, 32.6 for MeOH, and linear combinations for the mixtures.

The equations of motion were integrated using the leap-frog algorithm with a time step of 2 fs. The velocities of the atoms at the beginning of the simulation were assigned from a Maxwell distribution at 298 K. During the runs, configurations of the system were saved every 2 ps.

The parameters of the CG DMSO and the MeOH models are given in Table 1, the ones for the water model in Ref. 9. For the mixtures the Lennard–Jones parameters were calculated using $C_{12}^{AB} = \sqrt{C_{12}^{A} \cdot C_{12}^{B}}$ and $C_{6}^{AB} = \sqrt{C_{6}^{A} \cdot C_{6}^{B}}$. Starting configurations for the mixtures were obtained by converting a fraction of the DMSO and the MeOH beads, respectively, into water beads.[9] The mixtures were simulated for 4.5 ns after 0.1 ns of equilibration.

The behavior of mixtures of CG water with CG DMSO and MeOH are shown in Figure 2 and Figure 3, respectively. The CG models are able to reproduce the general trends of the experimental data for the density, the excess volume, and the dielectric permittivity. For the excess potential energy of mixing the CG model predicts the opposite trend of the experimental data, which indicates that the geometric-mean combination rule for obtaining the Lennard–Jones parameters between unlike beads should be modified.[63]

Properties from nonequilibrium simulations of polarizable water

The nonequilibrium methods were tested with a box of polarizable COS/D water.[21]

For the calculation of the dielectric permittivity, a cubic box with an edge length of 3.1057 nm was filled with 1000 COS/D[21] water molecules, resulting in a density of 997 kg/m^3, corresponding to the density of liquid water at 298 K and 1 atm.[60] For the

calculation of the viscosity, the box was extended in the *z*-direction by a factor 5, containing 5000 water molecules. The geometry of the water molecules was constrained by applying the SHAKE algorithm[62] with a relative geometric tolerance of 10^{-4} on the OH bond length and on the intramolecular HH distance.

The temperature was weakly coupled[59] to a bath of 298.15 K with a relaxation time of 0.1 ps. The box to calculate the viscosity was simulated at constant volume, for the other, the pressure was weakly coupled[59] to a bath of 1 atm with a relaxation time of 0.5 ps. The isothermal compressibility was set to the experimental value[60] of $7.513 \times 10^{-4}(\mathrm{kJ\,mol^{-1}\,nm^{-3}})^{-1}$. The nonbonded van der Waals and electrostatic interactions were calculated using triple-range cutoff radii of 0.8/1.4 nm. The short range interactions were calculated every time step by updating the molecular pair list for distances smaller than the first cutoff radius of 0.8 nm. For the intermediate range of distances between 0.8 and 1.4 nm, the pairlist was only updated every fifth timestep and at the same time the interaction was calculated and kept unchanged between these updates. The long range electrostatic interactions beyond the outer cutoff of 1.4 nm were represented by a reaction field[54,61] with $\varepsilon_{RF} = 78.5$ for water. The equations of motion were integrated using the leap-frog algorithm with a time step of 2 fs. The velocities of the atoms at the beginning of the simulation were assigned from a Maxwell distribution at 298 K. During the runs, configurations of the system were saved every 0.2 ps for the dielectric permittivity and every 1 ps for the viscosity.

The electric field in the *z*-direction was set to $E_z^{ext} = 0$, 0.01, 0.025, 0.0375, 0.5, 0.625, 0.75, 0.1, 0.15, and 0.2 e nm^{-2} . For each value, the box was simulated for 5.8 ns after an equilibration phase of 0.2 ns.

The amplitude of the acceleration field was set to $A =$ 0.001, 0.002, 0.005, 0.01, 0.02, 0.04, 0.05, 0.06, 0.07, 0.08, and 0.1 kJ mol^{-1} nm^{-1} u^{-1}, and for each value, a simulation of 2.9 ns was performed after 0.1 ns of equilibration.

The dependence of the polarization in *z*-direction for different field strength in the same direction is shown in Figure 4. The static dielectric permittivity $\varepsilon(0)$ of the polarizable COS/D water model derived from the slope, ignoring large field strengths, is 77.3 $\pm$ 0.6 from a total of 34.8 ns of simulation. This is closer to the experimental value[58] of 78.5 than the previously reported[21] $\varepsilon(0) = 69.8 \pm 6$ derived from a total of 50 ns of simulation using the box dipole moment fluctuation formula.

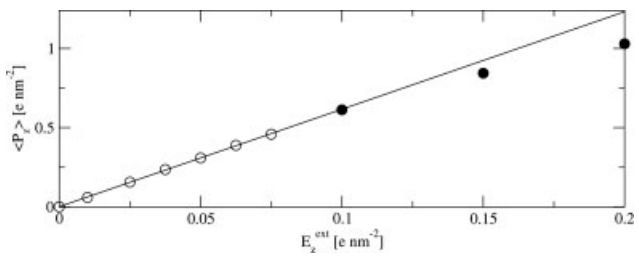

Figure 4. Dependence of the polarization averaged over time $\langle P_z \rangle$ on the applied electric field E_z^{ext} for COS/D[21] water. The circles are the results from the simulation, the line from linear regression for which the filled cicles were ignored since the system is saturated for larger field strengths.

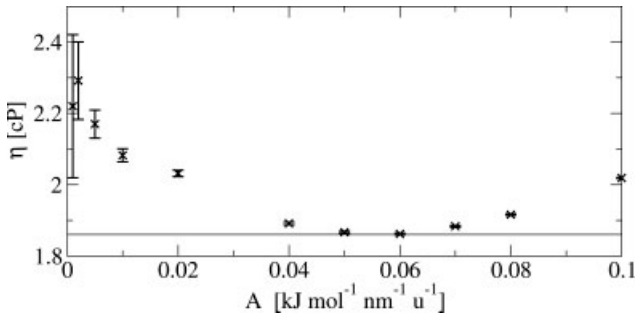

Figure 5. Viscosity η as a function of the amplitude *A* of the applied external sinusoidal acceleration. The solid line indicates the assumed minimum of the curve used to determined the viscosity η of COS/D[21] water.

The viscosity was determined to be 1.86 cP (see Fig. 5) for the region in which the perturbation is small enough not to disturb the system but strong enough to converge within reasonable time. The value found is about twice as high as the experimental value[58] of 0.89 cP.

Acknowledgments

The authors thank Markus Christen, Haibo Yu, Chris Oostenbrink, Moritz Winger, and Merijn Schenk for fruitful discussions.

[1] W. F. van Gunsteren, H. J. C. Berendsen, GRoningen MOlecular Simulation (GROMOS) Library Manual; Biomos: Groningen, The Netherlands, 1987.

[2] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, I. G. Tironi, Biomolecular Simulation: The GROMOS96 Manual and User Guide; Vdf Hochschulverlag AG an der ETH Zürich: Switzerland, 1996.

[3] W. R. P. Scott, W. F. van Gunsteren, Methods and Techniques in Computational Chemistry: METECC-95; STEF: Cagliari, Italy, 1995; pp. 397–434.

[4] W. R. P. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger, van Gunsteren, W. F. J Phys Chem A 1999, 103, 3596.

[5] M. Christen, P. H. Hünenberger, D. Bakowies, R. Baron, R. Bürgi, D. P. Geerke, T. N. Heinz, M. A. Kastenholz, V. Kräutler, C. Oostenbrink, C. Peter, D. Trzesniak, W. F. van Gunsteren, J Comput Chem 2005, 26, 1719.

[6] H. B. Yu, W. F. van Gunsteren, Comput Phys Commun 2005, 172, 69.

[7] D. P. Geerke, W. F. van Gunsteren, J Phys Chem B 2007, 111, 6425.

[8] J. R. Allison, K. Boguslawski, F. Fraternali, W. F. van Gunsteren, J Phys Chem B 2011, 115, 4547.

[9] S. Riniker, W. F. van Gunsteren, J Chem Phys 2011, 124, 084110.

[10] A. P. E. Kunz, H. Liu, W. F. van Gunsteren, J Chem Phys 2011, 135, 104206.

[11] S. Riniker, A. P. E. Kunz, W. F. van Gunsteren, J Chem Theory Comput 2011, 7, 1469.

[12] N. Schmid, C. D. Christ, M. Christen, A. P. Eichenberger, W. F. van Gunsteren, Comput Phys Commun, Submitted.

[13] N. Schmid, J. R. Allison, J. Dolenc, A. P. Eichenberger, A. P. E. Kunz, W. F. van Gunsteren, J Biomol NMR online.

[14] S. Riniker, C. D. Christ, H. Hansen, P. H. Hünenberger, C. Oostenbrink, D. Steiner, W. F. van Gunsteren, J Phys Chem 2011, Submitted.

[15] A. P. Eichenberger, J. R. Allison, J. Dolenc, D. P. Geerke, B. A. C. Horta, K. Meier, N. Schmid, D. Steiner, D. Wang, W. F. van Gunsteren, J Chem Theory Comput 2011, accepted.

[16] S. W. Rick, S. J. Stuart, In Reviews in Computational Chemistry, Editors: K.B. Lipkowitz and D.B. Boyd, Vol. 18; Wiley-Vch: New York, 2002; pp. 89–146.
[17] T. P. Straatsma, J. A. McCammon, Mol Simul 1990, 5, 181.
[18] P. Drude, The Theory of Optics; Longmans, Green, and Co.: New York, 1902.
[19] M. Born, K. Huang, Dynamic Theory of Crystal Lattices; Oxford University Press: Oxford, UK, 1954.
[20] H. B. Yu, W. F. van Gunsteren, J Chem Phys 2004, 121, 9549.
[21] A. P. E. Kunz, W. F. van Gunsteren, J Phys Chem A 2009, 113, 11570.
[22] X. Daura, A. E. Mark, W. F. van Gunsteren, Comput Phys Commun 1999, 123, 97.
[23] W. F. van Gunsteren, F. J. Luque, D. Timms, A. E. Torda, Annu Rev Biophys Biomol Struct 1994, 23, 847.
[24] B. Roux, T. Simonson, Biophys Chem 1999, 78, 1.
[25] M. Feig, C. L. Brooks, Curr Opin Struct Biol 2004, 14, 217.
[26] J. Chen, C. L. Brooks, J. Khandogin, Curr Opin Struct Biol 2008, 18, 140.
[27] C. Chothia, Nature 1974, 248, 338.
[28] D. Eisenberg, A. D. McLachlan, Nature 1986, 319, 199.
[29] T. Ooi, M. Oobatake, G. Némethy, H. A. Scheraga, Proc Natl Acad Sci USA 1987, 84, 3086.
[30] Y. K. Kang, K. D. Gibson, G. Némethy, H. A. Scheraga, J Phys Chem 1988, 92, 4739.
[31] W. C. Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, J Am Chem Soc 1990, 112, 6127.
[32] K. Sharp, A. Nicholls, R. Fine, B. Honig, Science 1991, 252, 106.
[33] J. Vila, R. L. Williams, M. Vásquez, H. A. Scheraga, Proteins: Struct Funct Genet 1991, 10, 199.
[34] C. A. Schiffer, J. W. Caldwell, R. M. Stroud, P. A. Kollman, Protein Sci 1992, 1, 396.
[35] L. Wesson, D. Eisenberg, Protein Sci 1992, 1, 227.
[36] P. Stouten, C. Frömmel, H. Nakamurah, C. Sander, Mol Simul 1993, 10, 97.
[37] F. Fraternali, W. F. van Gunsteren, J Mol Biol 1996, 256, 939.
[38] T. Lazaridis, M. Karplus, Proteins: Struct Funct Genet 1999, 35, 133.
[39] P. Ferrara, J. Apostolakis, A. Caflisch, Proteins: Struct Funct Genet 2002, 46, 24.
[40] F. Fraternali, L. Cavallo, Nucleic Acids Res 2002, 30, 2950.
[41] D. S. Cerutti, L. F. Ten Eyck, J. A. McCammon, J Chem Theory Comput 2004, 1, 143.
[42] F. Floris, J. Tomasi, J Comput Chem 1989, 10, 616.
[43] F. M. Floris, J. Tomasi, J. L. P. Ahuir, J Comput Chem 1991, 12, 784.
[44] E. Gallicchio, M. M. Kubo, R. M. Levy, J Phys Chem B 2000, 104, 6271.
[45] J. W. Pitera, W. F. van Gunsteren, J Am Chem Soc 2001, 123, 3163.
[46] M. Zacharias, J Phys Chem A 2003, 107, 3000.
[47] E. Gallicchio, R. M. Levy, J Comput Chem 2004, 25, 479.
[48] N. Choudhury, B. M. Pettitt, J Am Chem Soc 2005, 127, 3556.
[49] R. Malham, S. Johnstone, R. J. Bingham, E. Barratt, S. E. V. Phillips, C. A. Laughton, S. W. Homans, J Am Chem Soc 2005, 127, 17061.
[50] J. A. Wagoner, N. A. Baker, Proc Natl Acad Sci USA 2006, 103, 8331.
[51] W. Hasel, T. F. Hendrickson, W. C. Still, Tetrahedron Comput Methodol 1988, 1, 103.
[52] S. J. Wodak, J. Janin, Proc Natl Acad Sci USA 1980, 77, 1736.
[53] W. K. H. Panofsky, M. Phillips, Classical Electricity and Magnetism; Addison-Wesley: Reading, MA, 1964.
[54] M. Neumann, J Chem Phys 1985, 82, 5663.
[55] K. A. Feenstra, B. Hess, H. J. C. Berendsen, J Comput Chem 1999, 20, 786.
[56] V. R. Vasquez, E. A. Macedo, M. S. Zabaloy, Int J Thermophy 2004, 25, 1799.
[57] D. P. Geerke, W. F. van Gunsteren, Mol Phys 2007, 105, 1861.
[58] D. R. Lide, Handbook of Chemistry and Physics, 88th ed.; CRC Press/Taylor and Francis: Boca Raton, FL, 2007–2008.
[59] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, J Chem Phys 1984, 81, 3684.
[60] G. S. Kell, J Chem Eng Data 1967, 12, 66.
[61] I. G. Tironi, R. Sperb, P. E. Smith, W. F. van Gunsteren, J Chem Phys 1995, 102, 5451.
[62] J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, J Comput Phys 1977, 23, 327.
[63] J. R. Allison, S. Riniker, W. F. van Gunsteren, J. Chem. Phys., Submitted.

Received: 10 May 2011
Revised: 26 August 2011
Accepted: 3 September 2011
Published online on 11 November 2011