

SOFTWARE

Open Access



An automated calculation pipeline for differential pair interaction energies with molecular force fields using the Tinker Molecular Modeling Package

Felix Bänsch¹ , Mirco Daniel¹ , Harald Lanig² , Christoph Steinbeck³ and Achim Zielesny^{1*}

Abstract

An automated pipeline for comprehensive calculation of intermolecular interaction energies based on molecular force-fields using the Tinker molecular modelling package is presented. Starting with non-optimized chemically intuitive monomer structures, the pipeline allows the approximation of global minimum energy monomers and dimers, configuration sampling for various monomer–monomer distances, estimation of coordination numbers by molecular dynamics simulations, and the evaluation of differential pair interaction energies. The latter are used to derive Flory–Huggins parameters and isotropic particle–particle repulsions for Dissipative Particle Dynamics (DPD). The computational results for force fields MM3, MMFF94, OPLS-AA and AMOEBA09 are analyzed with Density Functional Theory (DFT) calculations and DPD simulations for a mixture of the non-ionic polyoxyethylene alkyl ether surfactant $C_{10}E_4$ with water to demonstrate the usefulness of the approach.

Scientific Contribution

To our knowledge, there is currently no open computational pipeline for differential pair interaction energies at all. This work aims to contribute an (at least academically available, open) approach based on molecular force fields that provides a robust and efficient computational scheme for their automated calculation for small to medium-sized (organic) molecular dimers. The usefulness of the proposed new calculation scheme is demonstrated for the generation of mesoscopic particles with their mutual repulsive interactions.

Keywords Intermolecular interaction, Nonbonding interaction, Molecular force field, Molecular modeling, Molecular dynamics, Monomer, Dimer, Flory–Huggins parameter, Isotropic repulsion, Dissipative particle dynamics, DPD, Non-ionic polyoxyethylene alkyl ether surfactant, $C_{10}E_4$

*Correspondence:

Achim Zielesny

achim.zielesny@w-hs.de

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

The quantitative description of non-bonding interactions between molecules is fundamental to understanding and designing chemical processes in materials and life sciences [1, 2]. In contrast to covalent bonding within molecules, non-bonding intermolecular interactions comprise dispersed variations of electromagnetic interactions like dipole/dipole, dipole/induced dipole, induced dipole/induced dipole (van der Waals) interactions, hydrogen bonding, (partial) charge interactions, π - π , cation/anion- π or polar π -effects. Each different spatial configuration of two molecules can be assigned a corresponding nonbonding intermolecular interaction energy. Its determination is challenging because the intermolecular interactions are generally small compared to covalent bonding and especially to the total energy of a molecular configuration.

For an accurate quantitative description, a quantum chemical treatment with a suitable model chemistry is commonly advised, e.g., application of Density Functional Theory (DFT) with an appropriate combination of functional and basis set: the complexation energy (i.e., the energy difference between a specific dimer configuration and the two monomer molecules that form it) can then quantify the intermolecular interaction. In particular, Symmetry-Adapted Perturbation Theory (SAPT) allows direct computation of non-bonding intermolecular interactions (i.e., without the need to calculate the total energies of monomers and dimer) and provides a physically meaningful decomposition of its contributing (electrostatics, induction, dispersion, short-range repulsion) terms [2]. Recent DFT-SAPT approaches have demonstrated a comparatively fast calculation in combination with remarkable accuracy for small- to medium-sized dimers up to the adenine-thymine base pair [3, 4].

When multiple spatial dimer configurations need to be sampled, quantum chemical approaches become increasingly expensive due to their considerable computational complexity, as a single calculation can easily take minutes or even longer. As a purely classical alternative, molecular force fields may be employed instead: they allow intermolecular energy calculations within fractions of a second for a specific spatial dimer configuration, i.e., an acceleration by several orders of magnitude. This comes at the expense of accuracy, as a given force field can easily lead to deceptively erroneous results.

This work aims at providing a robust automated molecular force-field based calculation pipeline for comprehensive estimation of mutual intermolecular energies of a set of small to medium-sized monomer molecules. For a chosen force field, the monomer molecules must be provided with an initial, chemically intuitive spatial geometry with associated atom types (where for small

molecules a planar 2D geometry provided by any 2D structure editor seems to be sufficient). Then, each monomer start geometry is globally optimized to its minimum energy conformer. Several monomer-monomer distances with multiple configurations at each distance are sampled to obtain a near-global minimum energy dimer. This dimer is then locally optimized towards its global minimum. The sampled configurations can be averaged by Boltzmann weights to get mean non-bonded dimer interactions at different distances. The calculations can be extended to obtain differential molecule pair interaction energies, which describe the excess intermolecular interaction of two different molecules in comparison to the average interaction of the molecules themselves (see Eqs. 1 and 2 in “Methods” section below), where molecular dynamics (MD) simulations are used to estimate the coordination numbers of neighboring molecules.

The resulting interaction energies may be useful for different purposes. Based on the comprehensive sampling, suitable spatial start configurations can be obtained for more elaborate (e.g., quantum chemical) refinement procedures. Pairwise non-bonded intermolecular interactions can be considered as molecule-pair descriptors for Cheminformatics tasks like molecular similarity estimation. Differential molecule pair interaction energies play an important role in statistical thermodynamics, e.g., for the quantitative estimation of excess quantities that determine the properties of mixtures [5]. In particular, they can be related to Flory-Huggins interaction parameters for polymer models (Eq. 9), which in turn can be used to describe isotropic particle-particle repulsions for mesoscopic simulation approaches (Eq. 10) [6]. The latter application is particularly studied in this work.

The backbone of the calculation pipeline is implemented using the Java programming language. All force-field-based calculations are performed with the Tinker Molecular Modelling package [7]. The Tinker package is modularized into stand-alone, task-based executables (marked in italics in “Methods” section below), which fit well into the Java backbone-driven parallelized computational scheme that fully exploits the computational capabilities of multicore workstations. All force fields provided by Tinker can be used for the calculation pipeline.

Methods

The force-field-based intermolecular energy $E_{ij}^C(r)$ between two molecules i and j is determined by different non-covalent-bonding contributions (van der Waals and partial charge interactions, hydrogen bonding etc.) and depends on the intermolecular distance r between the centers of the molecules as well as their relative spatial configurations C . For each specific distance r_{fix}

there is a minimum energy configuration C_{min} with $E_{ij}^{C_{min}}(r_{fix}) \leq E_{ij}^C(r_{fix})$. The global minimum energy dimer is characterized by a distinct distance r_{min} so that $E_{ij}^{min} = E_{ij}^{C_{min}}(r_{min}) \leq E_{ij}^C(r)$, i.e. r_{min} is the distance between the centers of two molecules i and j when the dimer geometry corresponds to the global energy minimum. If different spatial configurations with a fixed intermolecular distance r_{fix} are averaged, an averaged intermolecular energy for this distance $\langle E_{ij} \rangle(r_{fix})$ is obtained: The corresponding minimum energy distance $r_{\langle E_{ij} \rangle, min}$ with $\langle E_{ij} \rangle^{min} = \langle E_{ij} \rangle(r_{\langle E_{ij} \rangle, min}) \leq \langle E_{ij} \rangle(r)$ does not necessarily coincide with minimum distance r_{min} of the global minimum energy dimer. Among the concrete averaged configurations with a fixed intermolecular distance r_{fix} the configuration C^* with the minimal intermolecular energy is denoted $E_{ij}^{C^*}(r_{fix}) \geq E_{ij}^{C_{min}}(r_{fix})$.

A differential pair interaction energy describes the excess intermolecular interaction of molecules i and j in comparison to the average interaction of the molecules themselves. This may be defined with respect to the global minimum energy dimers

$$\Delta E_{ij} = E_{ij}^{min} - \frac{1}{2}(E_{ii}^{min} + E_{jj}^{min}) \quad (1)$$

or corresponding averages (see Eqs. 5 and 6 below for calculation details)

$$\Delta \langle E_{ij} \rangle = \langle E_{ij} \rangle^{min} - \frac{1}{2}(\langle E_{ii} \rangle^{min} + \langle E_{jj} \rangle^{min}). \quad (2)$$

A positive (negative) differential pair interaction energy indicates a less (more) favorable intermolecular interaction compared to the individual ones. When differential pair interaction energies are applied to lattice models, all vertices of the lattice have a fixed number of surrounding neighbours (e.g. $Z = 4$ in two dimensions or $Z = 6$ in three dimensions). In contrast, continuum models can (and usually do) have different coordination numbers Z_{ij} , where the number Z_{ij} of molecules j surrounding molecule i can (and usually does) differ from the number Z_{ji} of molecules i that surround molecule j . Equations 1 and 2 can be extended accordingly to

$$\Delta E_{ij}^Z = \frac{1}{2}(Z_{ij}E_{ij}^{min} + Z_{ji}E_{ji}^{min}) - \frac{1}{2}(Z_{ii}E_{ii}^{min} + Z_{jj}E_{jj}^{min}) \quad (3)$$

and

$$\Delta \langle E_{ij} \rangle^Z = \frac{1}{2}(Z_{ij}\langle E_{ij} \rangle^{min} + Z_{ji}\langle E_{ji} \rangle^{min}) - \frac{1}{2}(Z_{ii}\langle E_{ii} \rangle^{min} + Z_{jj}\langle E_{jj} \rangle^{min}) \quad (4)$$

respectively, where superscript Z denotes the coordination number extension in contrast to Eqs. 1 and 2, $E_{ij}^{min} = E_{ji}^{min}$ and $\langle E_{ij} \rangle^{min} = \langle E_{ji} \rangle^{min}$ but Z_{ij} may be different from Z_{ji} , i.e. $Z_{ij} \neq Z_{ji}$.

Thus, the estimation of differential pair interaction energies requires two steps: (I) A calculation scheme to obtain the different (averaged) molecule pair interaction energies E_{ij}^{min} ($\langle E_{ij} \rangle^{min}$), and (II) a procedure to estimate the different coordination numbers Z_{ij} . In the following, the concrete implementation of a corresponding automated calculation pipeline for a selected molecular force field using the Tinker Molecular Modelling package is described. All dimer-related calculations are started with the global minimum energy conformers of the two constituent monomer molecules (see following “[Global minimum energy monomers](#)” section), where multiple dimer configurations are analyzed with these global minimum energy monomers (see following “[Global minimum energy dimers](#)” section) to obtain an approximate configuration for the global minimum energy dimer. The latter is achieved by optimizing this final approximate dimer configuration without constraints using all atomic degrees of freedom, so that the monomers are no longer confined to their individual global minimum energy conformers.

Global minimum energy monomers

The global minimum energy conformers are derived in advance from conformer search procedures: STARTING from an initially defined chemically intuitive geometry of a monomer molecule, a first geometry improvement is achieved with Tinker *optimize* using the Optimally Conditioned Variable Metric (OCVM) optimization technique [by default with a root mean square (RMS) gradient of 0.01 kcal/mole/Å] [8]. The resulting locally optimized geometry is then transferred to a low-mode conformational search (LMO) with Tinker *scan* to find the minimum energy conformer (where by default, the RMS gradient for the LMO procedure is set to 0.0001 kcal/mole/Å, the energy threshold for local minima is set to 100 kcal/mole, torsion angles are automatically selected, and five eigenvectors are used for the local search) [9]. For small molecules with O(10) number of atoms, the detected minimum energy conformers usually coincide with the global minimum energy conformers.

Global minimum energy dimers

To approximate the global minimum energy dimer of a pair of molecules i and j , the centers of the molecules are positioned at different distances ranging (by default) from 3 to 16 Å in steps of 0.5 Å where the initial relative configuration of the two molecules is arbitrary. For each distance, a configuration sampling procedure is performed, which is sketched in Fig. 1. A (unit) sphere around each center is constructed with a number of N_{sphere} evenly spaced points being generated on each sphere using a Fibonacci lattice as an adequate approximation (in comparison to a latitude–longitude lattice the surface points generated by a Fibonacci lattice are more evenly spaced with a smaller axial anisotropy) [10]. By rotating the molecules around their centers so that two adjacent spherical surface points and the centers of both molecules lie on a straight axis, $N_{sphere} \times N_{sphere}$ configurations are generated for which the corresponding interaction energies are

determined by Tinker *analyze* (with settings to only compute the non-bonding interactions for a significantly accelerated performance). In addition, for each single configuration the second molecule is rotated for a number N_{rot} of angles around the axis with corresponding interaction energy calculations, so that in total $N_{sphere} \times N_{sphere} \times N_{rot}$ spatial configurations are sampled for a single fixed distance between the monomer molecules (with each monomer being constrained to its individual global minimum energy conformer). The distance resolution is then refined around the distance with the detected minimum interaction energy dimer by reducing the step size from 0.5 to 0.1 Å, and then again from 0.1 to 0.01 Å (compare Fig. 2), so that a final near minimum energy dimer configuration is evaluated for the latter resolution. This resulting configuration C^* is then optimized with Tinker *optimize* without constraints using all atomic degrees of freedom (i.e. the monomers are no longer confined to their individual

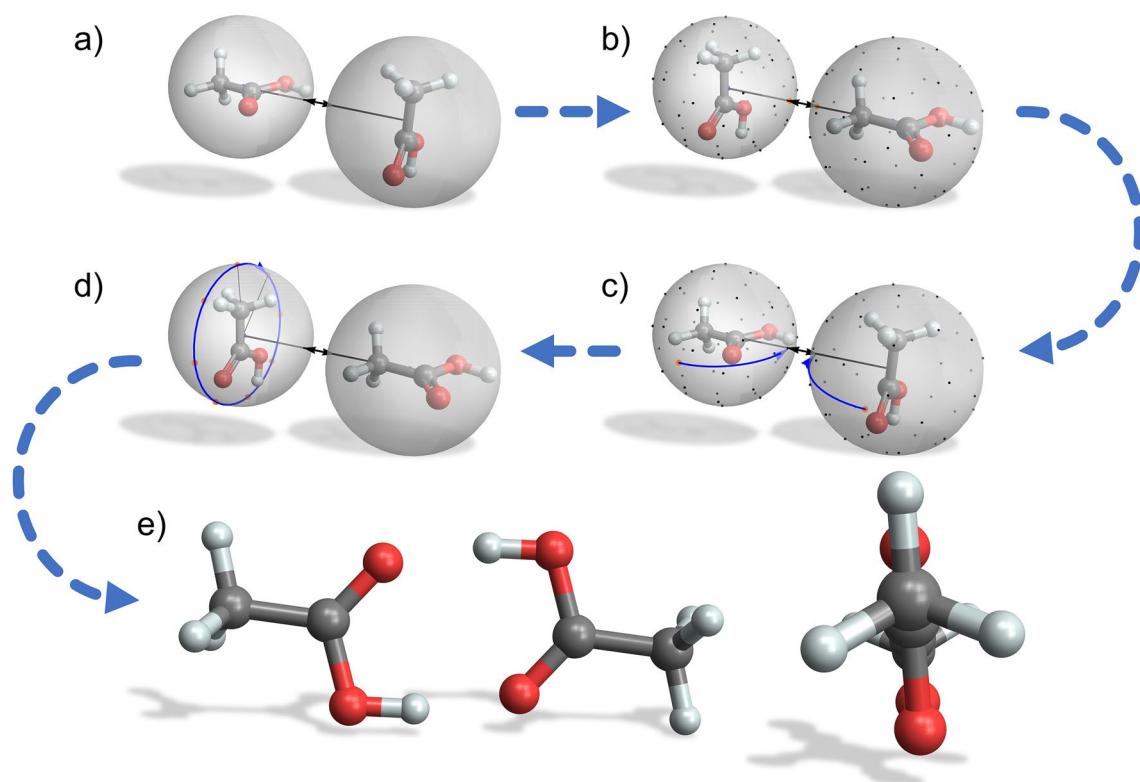


Fig. 1 Configuration sampling for the acetic acid dimer: **a** An initial dimer of two acetic acid molecules is constructed with a distinct distance between the centers of both molecules (with each monomer being constrained to its individual global minimum energy conformer). **b** Spheres around the centers of the molecules are populated with a number N_{sphere} of evenly distributed (equidistant) points on their surfaces. **c** The interaction energy is determined for every configuration where two adjacent spherical surface points and the centers of both molecules lie on a straight axis (which is achieved by corresponding rotations of the molecules around their centers). **d** For each single configuration in **c** the second molecule is also rotated for a number of angles N_{rot} around the axis with a corresponding interaction energy calculation. **e** By varying and refining the distance between the molecule centers, the minimum energy dimer is determined and finally optimized to approximate the global minimum energy dimer without constraints using all atomic degrees of freedom. For the acetic acid dimer two perspectives of the finally optimized global minimum energy dimer with E_{ij}^{min} are shown, using the MMFF94 force field [11]

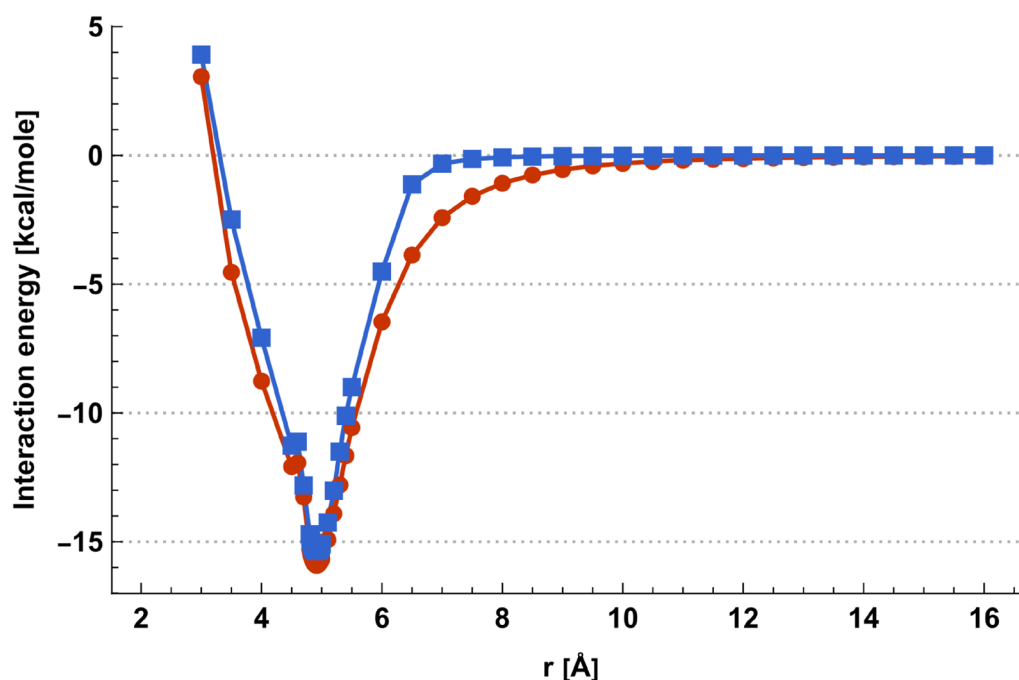


Fig. 2 Acetic acid dimer. Red: $E_{ij}^{C*}(r)$ for the minimum sampled energy configuration, Blue: Averaged interaction energy $\langle E_{ij} \rangle(r)$ for temperature $T = 298\text{K}$

global minimum energy conformers) to arrive at the force-field dependent approximation for E_{ij}^{min} of the two molecules i and j , the global minimum energy dimer using the specified force-field.

Configuration sampling

If configuration sampling is desired for a specific distance r_{fix} of the dimer molecules, the (already) obtained interaction energies of the $N_{sphere} \times N_{sphere} \times N_{rot}$ configurations for this distance may be averaged with Boltzmann weights $w_C(r_{fix})$ for a defined temperature (with k_B , the Boltzmann constant, and T , thermodynamic temperature) using E_{ij}^{min} :

$$w_C(r_{fix}) = e^{-\frac{E_{ij}^C(r_{fix}) - E_{ij}^{min}}{k_B T}} \quad (5)$$

$$\langle E_{ij} \rangle(r_{fix}) = \frac{\sum_C^{N_{sphere} \times N_{sphere} \times N_{rot}} E_{ij}^C(r_{fix}) w_C(r_{fix})}{\sum_C^{N_{sphere} \times N_{sphere} \times N_{rot}} w_C(r_{fix})} \quad (6)$$

Then $\langle E_{ij} \rangle$ can be evaluated from the different $\langle E_{ij} \rangle(r)$. Figure 2 shows the quantitative results for the acetic acid dimer using the Merck molecular force field (MMFF94) [11].

Coordination numbers

The coordination numbers Z_{ij} are estimated by MD simulations. The simulation box construction is based on a pure water box at 298 K to get consistent results. A water molecule has a van der Waals volume of $V_{vdW}^{H_2O} = 17.35 \text{ \AA}^3$, in a pure water box it occupies a volume of $V_{box}^{H_2O} = 30.00 \text{ \AA}^3$ at 298 K due to its density [12] and molar mass [13]. This relation is mapped to other molecules X with

$$\frac{V_{box}^{H_2O}}{V_{vdW}^{H_2O}} = \frac{V_{box}^X}{V_{vdW}^X} \quad (7)$$

so that the edge length a of a cubic simulation box of N molecules X is given by

$$a = \sqrt[3]{N \frac{V_{box}^{H_2O}}{V_{vdW}^{H_2O}} V_{vdW}^X} \quad (8)$$

The van der Waals volumes are approximated with the *VABCVolume* [14] descriptor of the Chemistry Development Kit (CDK) [15, 16]. A simulation box with a defined number N of molecules j (default is 400) and defined edge length a is created using Tinker *xyzedit*. Then a single molecule i is added to the box,

where Tinker *xyzedit* automatically removes molecules j to keep the defined density of the simulation box. The (possibly unfavorable) start configuration is optimized with Tinker *minimize* to avoid atomic contacts that could lead to instabilities. The following MD simulation uses Tinker *dynamics* with a step size of one femtosecond and an Andersen thermostat [17] for temperature equilibrium (default is 298 K). 10,000 (default) initial steps are used for box equilibration, followed by several hundred thousand steps with data recording (default is 400,000). For each recorded simulation step the number of molecules j surrounding the single molecule i is analyzed. This is done either by a “brute-force” counting approach or, alternatively, by the cell-index method with periodic boundary conditions [18]. The “brute-force” approach calculates the distances between each atom of the single molecule i and each atom of all molecules j . Alternatively, the cell-index method only considers the (drastically reduced) distances between each atom of single molecule i and the atoms of solvent molecules j in neighbouring cells. The criterion for including a molecule j as a relevant neighbor for the coordination number Z_{ij} is based on the distance between its atoms and those

of molecule i . If the distance of the respective atoms is less than or equal to the sum of their van der Waals radii plus an arbitrary “catch radius” (default is 1 Å), the molecules are considered neighbors. In Fig. 3, a snapshot of a simulation step is illustrated. For each simulation step, the number of neighboring molecules j is determined. The average over all recorded steps is used to estimate the coordination number Z_{ij} , see Fig. 3.

Flory–Huggins and mesoscopic repulsion parameters

Differential pair interaction energies may be directly utilized to estimate Flory–Huggins interaction parameters χ_{ij} by

$$\chi_{ij} = \frac{\Delta \langle E_{ij} \rangle^Z}{k_B T} \quad (9)$$

to describe polymer solutions [5], with $\Delta \langle E_{ij} \rangle^Z$ being defined in Eq. 4.

For “bridging the gap between atomistic and mesoscopic simulation” (Groot and Warren [6]), the interacting molecules can be identified with the particles of “bottom-up” mesoscopic Dissipative Particle Dynamics (DPD), where

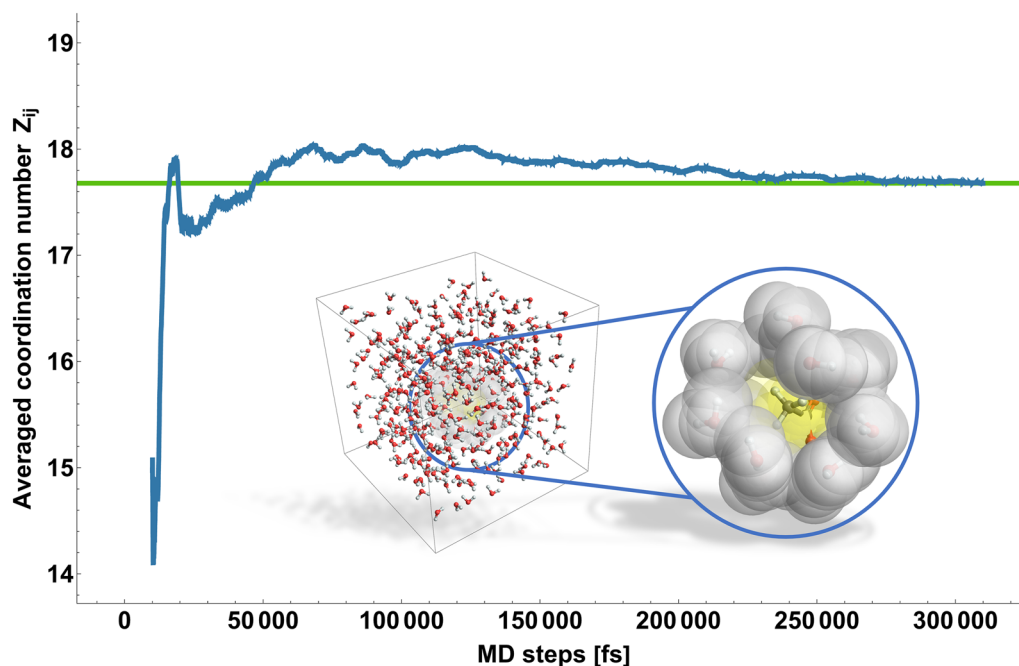


Fig. 3 Averaged coordination number Z_{ij} of a single acetic acid molecule in 400 water molecules starting after 10,000 initial steps (for box equilibration) with a “catch radius” of 1 Å using the MMFF94 force field. The averaged coordination number converges to a value of 17.7. Box graphics: Left: Snapshot of a simulation step of a single acetic acid molecule in 400 water molecules. Right: Magnification of the neighboring water molecules around the single acetic acid molecule. Yellow spheres: atoms of the single acetic acid molecule with their van der Waals radii. Grey spheres: Neighboring water molecules considered for the coordination number determination

the microscopic Flory–Huggins interaction parameters χ_{ij} can be directly related to mesoscopic isotropic particle–particle repulsions $a_{ij}(T)$ (expressed in units of $k_B T$, with ρ_{DPD} being the dimensionless DPD density, refer to [6] for details)

$$a_{ij}(T) = 75 \frac{k_B T}{\rho_{DPD}} + 3.4965 k_B T \chi_{ij} \quad (10)$$

which determine the conservative DPD forces $\underline{F}_{ij}^{C,DPD}$:

$$\underline{F}_{ij}^C = \underline{F}_{ij}^{C,DPD} + \underline{F}_{ij}^{C,Bond} \quad (11)$$

$$\underline{F}_{ij}^{C,DPD}(\underline{r}_{ij}) = \begin{cases} a_{ij}(1 - r_{ij})\underline{r}_{ij}^0 & \text{for } r_{ij} < 1 \\ 0 & \text{for } r_{ij} \geq 1 \end{cases} \quad (12)$$

with $\underline{F}_{ij}^{C,DPD}$, $\underline{F}_{ij}^{C,Bond}$, soft repulsive DPD force and harmonic bond force on particle i exerted by particle j ; a_{ij} , maximum isotropic repulsion between particles i and j ; $\underline{r}_{ij} = \underline{r}_i - \underline{r}_j = r_{ij}\underline{r}_{ij}^0$; \underline{r}_{ij}^0 , unit vector. The numerical factor (3.4965) in Eq. 10 is derived from Eq. 24 in reference [6] where the inverse value (0.286) is given.

In interplay with the dissipative (frictional) forces \underline{F}_{ij}^D and random forces \underline{F}_{ij}^R the conservative forces \underline{F}_{ij}^C guide the trajectories $\underline{r}_i(t)$ of the DPD particles according to Newton's equation of motion:

$$m_i \frac{d^2 \underline{r}_i}{dt^2} = \underline{F}_i = \sum_{j=1, j \neq i}^N (\underline{F}_{ij}^C + \underline{F}_{ij}^D + \underline{F}_{ij}^R) \quad (13)$$

with m_i , \underline{r}_i , mass and position vector of particle i ; t , time; \underline{F}_i , total force on particle i exerted by other particles j ; N , total number of particles in simulation; \underline{F}_{ij}^C , \underline{F}_{ij}^D , \underline{F}_{ij}^R , conservative, dissipative, and random force on particle i exerted by particle j .

Thus, the described calculation pipeline may be applied to construct a force-field-based particle set for DPD simulations.

Pipeline code availability

The pipeline code is written in Java and openly available at [19]. A dedicated installer executable for the Java pipeline code, which comprises a full Java runtime environment, is available at [20] for the Windows operating system. For Linux operating systems a zip file is available at [20]. According to its licensing terms the Tinker executables for *optimize*, *scan*, *analyse* etc. have to be downloaded from its website [21] into a specified pipeline directory: a detailed instruction how to perform this is provided at [22]. A set of stand-alone Mathematica notebooks [23] for result visualizations is provided at [24]. A test pipeline is available at [19] to ensure proper installation.

Pipeline calculation performance

Calculation of a full single differential pair interaction energy for the force fields MM3, MMFF94 and OPLS-AA with default settings ($N_{sphere} \times N_{sphere} \times N_{rot} = 144 \times 144 \times 16 = 331,776$ dimer configurations for each fixed distance to approximate the intermolecular interaction energies, 10,000 equilibration steps and 400,000 simulation steps for the MD simulations to estimate the coordination number with 400 molecules in the box) takes several hours, where the AMOEBA09 force field requires a multiple. Since the pipeline supports comprehensive calculation parallelization for a set of monomer molecules, a single differential pair interaction energy can be obtained on average in less than an hour on a 64-core AMD Ryzen™ Threadripper™ PRO 5995 CPU workstation [25].

DFT calculations for result evaluation

DFT calculations are performed with Gaussian 16 [26] and analyzed with GaussView 6 [27]. All molecular geometries are optimized using the dispersion-corrected wB97XD functional [28] with the 6–311++G(d,p) basis set where counterpoise calculations are used to obtain basis set superposition error (BSSE) corrected interaction energies. All Gaussian jobs files used are openly available at [29].

DPD simulations for result evaluation

All DPD simulations of this study are performed with the MFsim simulation system [30] using the Jdpd simulation kernel [31]. All constructed particle sets and MFsim simulation jobs are openly available at [32].

Results and discussion

To demonstrate the applicability of the different steps of the calculation pipeline, several small molecules are selected: Water (abbreviated H₂O), methane (Me), ethane (Et), methanol (MeOH), dimethyl ether (Me₂O) and the acetic acid dimer (HAc). The calculation results for this molecule set are evaluated and compared with alternative approaches and experimental results. All averaged energies and MD simulations refer to a temperature of $T = 298K$. All intermolecular energy calculations were performed with $N_{sphere} \times N_{sphere} \times N_{rot} = 144 \times 144 \times 16 = 331,776$ dimer configurations for each fixed distance of the molecule centers.

Acetic acid dimer

The acetic acid dimer is stabilized by two hydrogen bonds and has a planar geometry. Calculation results with the MMFF94 force field are shown in Figs. 2 and

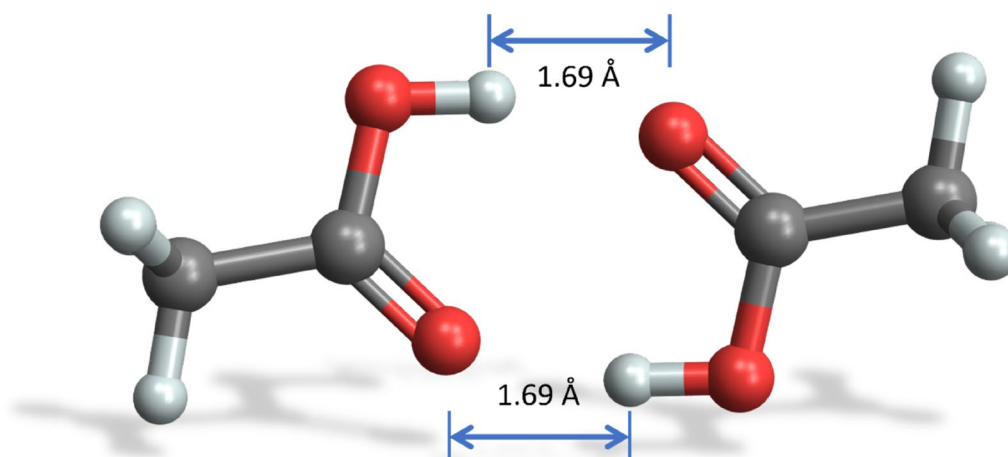


Fig. 4 Minimal energy configuration C^* of the acetic acid dimer with two hydrogen bonds

4. Figure 2 depicts the minimal energy configuration C^* energies $E_{ij}^{C^*}(r_{fix})$ for each specific distance r_{fix} in red, exhibiting a single minimum at $r_{fix}=4.91$ Å with $E_{ij}^{C^*}(4.91$ Å) = -15.9 kcal/mole. The corresponding averaged intermolecular energies $\langle E_{ij} \rangle(r_{fix})$ are shown in blue, where the single minimum distance coincides at $r_{fix}=4.91$ Å with $\langle E_{ij} \rangle(4.91$ Å) = -15.4 kcal/mole in this case. The two hydrogen bonds of the minimal energy configuration C^* have a length of 1.69 Å and a distance of 2.63 Å between the corresponding donor and acceptor

oxygen atoms (see Fig. 4), which is close to the experimental values of 2.68 Å [33].

With the final optimization of the near minimal energy configuration C^* the global MMFF94 force-field-based minimum energy configuration C_{min} with $E_{ij}^{min} = E_{ij}^{C_{min}}(r_{min}) = -17.6$ kcal/mole is obtained, which is 1.7 kcal/mole below the sampled minimal energy C^* configuration: the finally optimized dimer keeps the distance of $r_{min}=4.91$ Å but shows a planar geometry with a slightly reduced hydrogen bond length

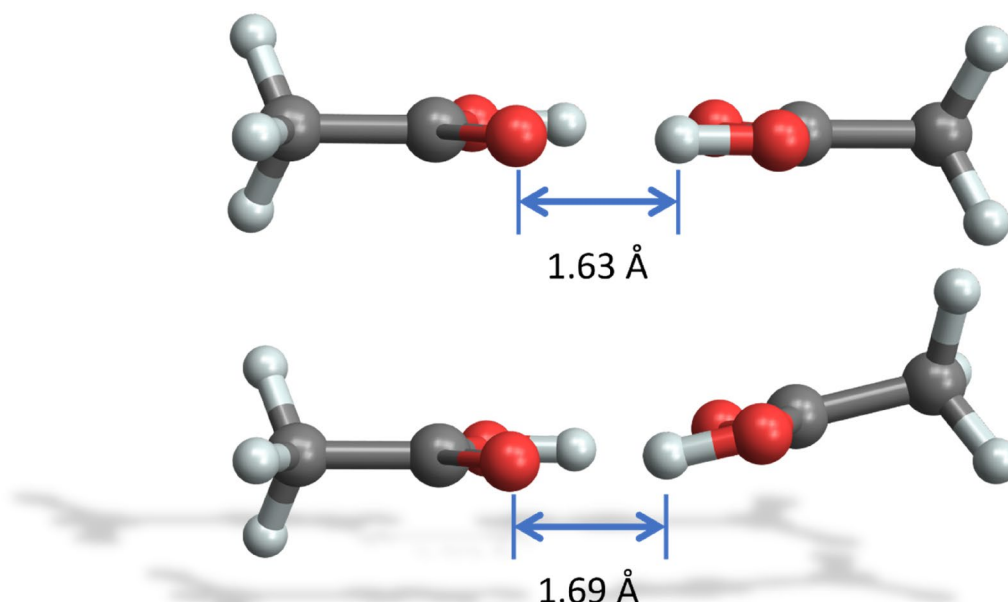


Fig. 5 Acetic acid dimer with the MMFF94 force field. Bottom: minimal sampled energy configuration C^* with $E_{ij}^{C^*}(4.91$ Å) = -15.9 kcal/mole. Top: optimized global minimum energy configuration C_{min} with $E_{ij}^{min} = E_{ij}^{C_{min}}(4.91$ Å) = -17.6 kcal/mole

Table 1 Force-field based intermolecular interaction energies in kcal/mole for the different dimers (averages are obtained at $T = 298K$)

Dimer	MMFF94			MM3			AMOEBA09			OPLS-AA (TIP3P)			OPLS-AA (TIP5P)		
	$\langle E_{ij} \rangle$	$E_{ij}^{C^*}$	E_{ij}^{min}	$\langle E_{ij} \rangle$	$E_{ij}^{C^*}$	E_{ij}^{min}	$\langle E_{ij} \rangle$	$E_{ij}^{C^*}$	E_{ij}^{min}	$\langle E_{ij} \rangle$	$E_{ij}^{C^*}$	E_{ij}^{min}	$\langle E_{ij} \rangle$	$E_{ij}^{C^*}$	E_{ij}^{min}
Et–Et	–0.4	–0.8	–0.8	–0.4	–0.8	–0.8	–0.7	–1.3	–1.4	–0.6	–1.2	–1.2			
EtOH–Et	–0.4	–1.0	–1.0	–0.4	–0.9	–0.9	–0.9	–2.0	–2.1	–0.7	–1.5	–1.5			
EtOH–EtOH	–4.3	–6.2	–6.5	–5.8	–7.3	–7.5	–4.0	–6.2	–6.6	–4.8	–6.8	–7.3			
H2O–Et	–0.2	–0.5	–0.5	–0.2	–0.5	–0.5	–0.5	–1.3	–1.3	–0.3	–0.9	–0.9	–0.3	–0.8	–0.8
H2O–EtOH	–4.9	–6.3	–6.7	–6.3	–7.5	–7.7	–4.8	–6.4	–6.5	–5.4	–6.7	–7.3	–5.5	–7.6	–10.9
H2O–H2O	–5.0	–6.3	–6.8	–5.9	–7.2	–7.4	–3.3	–4.9	–5.0	–5.1	–6.5	–6.9	–4.8	–6.5	–7.3
Me–Et	–0.3	–0.6	–0.6	–0.3	–0.6	–0.6	–0.4	–0.9	–0.9	–0.4	–0.8	–0.8			
Me–EtOH	–0.3	–0.7	–0.7	–0.3	–0.7	–0.7	–0.6	–1.7	–1.7	–0.5	–1.0	–1.0			
Me–H2O	–0.2	–0.3	–0.3	–0.2	–0.3	–0.3	–0.4	–1.2	–1.2	–0.2	–0.6	–0.6	–0.2	–0.5	–0.5
Me–Me	–0.2	–0.4	–0.4	–0.2	–0.4	–0.4	–0.4	–0.5	–0.5	–0.3	–0.5	–0.5			
Me2O–Et	–0.4	–1.0	–1.0	–0.4	–0.9	–1.0	–0.9	–1.9	–2.0	–0.8	–1.4	–1.4			
Me2O–EtOH	–3.8	–5.6	–5.9	–5.4	–6.6	–6.7	–3.3	–5.6	–5.9	–3.1	–5.2	–5.6			
Me2O–H2O	–4.5	–5.8	–6.1	–5.0	–6.2	–6.3	–4.3	–6.0	–6.2	–3.9	–5.2	–5.4	–3.2	–4.4	–4.5
Me2O–Me	–0.3	–0.7	–0.7	–0.3	–0.7	–0.7	–0.6	–1.5	–1.6	–0.5	–1.0	–1.0			
Me2O–Me2O	–1.2	–2.4	–2.4	–0.7	–1.5	–1.6	–1.6	–3.1	–3.2	–1.5	–2.4	–2.4			
MeOH–Et	–0.3	–0.8	–0.8	–0.3	–0.7	–0.7	–0.7	–1.7	–1.8	–0.6	–1.4	–1.4			
MeOH–EtOH	–4.7	–6.2	–6.4	–5.9	–7.3	–7.5	–4.6	–6.4	–6.7	–5.1	–6.7	–7.2			
MeOH–H2O	–4.6	–6.0	–6.5	–6.0	–7.6	–7.6	–3.8	–5.6	–5.8	–4.9	–6.4	–6.9	–6.4	–8.0	–10.8
MeOH–Me	–0.3	–0.6	–0.6	–0.3	–0.5	–0.5	–0.5	–1.4	–1.4	–0.4	–0.9	–0.9			
MeOH–Me2O	–4.2	–5.7	–5.9	–5.4	–6.6	–6.7	–4.0	–5.8	–6.2	–3.7	–5.2	–5.5			
MeOH–MeOH	–4.3	–5.9	–6.1	–5.8	–7.1	–7.2	–3.7	–5.6	–5.8	–4.5	–6.2	–6.6			
HAc–HAc	–15.4	–15.9	–17.6	–16.9	–17.4	–18.1	–15.4	–15.8	–18.0	–11.4	–12.2	–13.6			

of 1.63 Å and a distance of 2.62 Å between the donor and acceptor oxygen atoms (see Fig. 5).

Mutual dimer calculations

Table 1 presents the mutual dimer calculations: it comprises $\langle E_{ij} \rangle$ and $E_{ij}^{C^*}$ for the detected minima (compare Fig. 2) as well as the global force-field-based energy minimum E_{ij}^{min} for force fields MMFF94, AMOEBA09, MM3 [34] and OPLS-AA [35, 36] with the water models TIP3P [37] and TIP5P [38]. As expected, the nonpolar pure alkyl (methane and ethane) dimers exhibit only small interaction energies, the acetic acid dimer with two hydrogen bonds shows the largest interaction, and the polar dimers are in between. There are clear differences between the force fields, with the OPLS-AA (TIP5P) interactions for the alcohol–water dimers being the strongest. On average, MM3 differs most significantly from the other force fields.

For the HAc–HAc and the Me–Me dimer the results for force fields OPLS-AA, AMOEBA09 and MMFF94 were compared with corresponding DFT single point calculations (denoted DFT sp). In addition, the spatial configurations with E_{ij}^{min} were used as start geometries

for DFT geometry optimizations (denoted DFT opt), see Table 2. The DFT calculations indicate that the automated pipeline leads to acceptable near minimum energies and corresponding spatial configurations—with individual exceptions: in contrast to the MMFF94 and OPLS-AA force fields, AMOEBA09 results in an eclipsed minimum energy configuration for the Me–Me dimer instead of a staggered one (see Fig. 6, this eclipsed configuration is maintained by the DFT geometry optimization), but this finding has no significant influence on the subsequent investigations due to its low energetic effect (see Table 2). The difference in interaction energies between the DFT opt and the OPLS-AA force field result for the HAc–HAc dimer is most significant.

Coordination numbers Z_{ij}

Table 3 contains the mutual coordination numbers Z_{ij} (at $T = 298K$), where a single molecule i is surrounded by 400 molecules j , using the MMFF94, OPLS-AA (with TIP3P and TIP5P water models), and AMOEBA09 force fields. For comparison, static packing [40] results are included which were taken from previous research [41] and obtained with the commercial Blends software

Table 2 Force-field based interaction energies $E_{ij}^{C^*}$ and E_{ij}^{min} with corresponding DFT interaction energies and configuration measures

Dimer		$E_{ij}^{C^*}$ [kcal/mole]		Distance [Å]		
		DFT sp	Force field	DFT = force field = fixed		
OPLS-AA	Me–Me	–0.6	–0.5	3.57 ^b		
	HAc–HAc	–14.5	–12.2	2.69 ^a		
AMOEBA09	Me–Me	–0.5	–0.5	3.74 ^{b,c}		
	HAc–HAc	–15.8	–15.8	2.76 ^a		
MMFF94	Me–Me	–0.6	–0.4	3.74 ^b		
	HAc–HAc	–11.5	–15.9	2.63 ^a		

Dimer		E_{ij}^{min} [kcal/mole]		Distance [Å]		
		DFT sp	Force field	DFT = force field = fixed		
OPLS-AA	Me–Me	–0.6	–0.5	3.57 ^b		
	HAc–HAc	–15.7	–13.6	2.68 ^a		
AMOEBA09	Me–Me	–0.5	–0.5	3.74 ^{b,c}		
	HAc–HAc	–17.6	–18.0	2.73 ^a		
MMFF94	Me–Me	–0.6	–0.4	3.73 ^b		
	HAc–HAc	–15.1	–17.6	2.62 ^a		

Dimer		DFT opt	Force field	DFT	Force field	Exp
OPLS-AA	Me–Me	–0.6	–0.5	3.57 ^b	3.57 ^b	3.85 ^b
	HAc–HAc	–19.8	–13.6	2.68 ^a	2.68 ^a	2.68 ^a
AMOEBA09	Me–Me	–0.5	–0.5	3.82 ^{b,c}	3.74 ^{b,c}	3.85 ^b
	HAc–HAc	–19.8	–18.0	2.68 ^a	2.73 ^a	2.68 ^a
MMFF94	Me–Me	–0.6	–0.4	3.57 ^b	3.73 ^b	3.85 ^b
	HAc–HAc	–19.8	–17.6	2.68 ^a	2.62 ^a	2.68 ^a

a: Hydrogen bond length (distance oxygen to oxygen), b: Distance between carbon atoms, c: Eclipsed minimum energy configuration

The Me–Me dimer distance denotes the distance between the carbon atoms (b), while the HAc–HAc dimer distance refers to the specific hydrogen bond length (labelled a). Experimental values (denoted Exp.) are taken from [29] for acetic acid and [33, 39] for methane. For the Me–Me dimer the AMOEBA09 force field yields an eclipsed minimum energy configuration instead of a staggered one (labelled c)

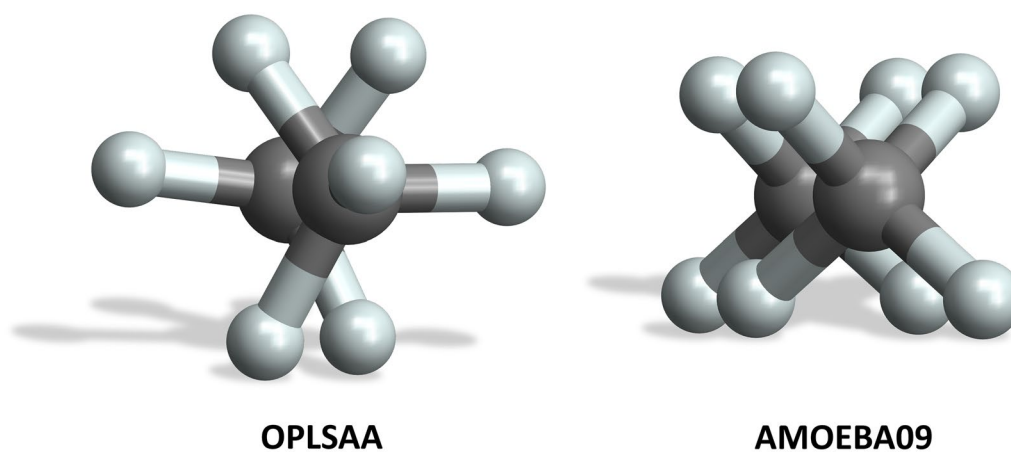


Fig. 6 Minimum energy $E_{ij}^{C^*}$ configuration for the Me–Me dimer: Left: OPLS-AA force field with staggered configuration. Right: AMOEBA09 force field with eclipsed configuration

Table 3 Coordination numbers Z_{ij} (at $T = 298\text{K}$)

Dimer	Z_{ij}					Z_{ji}				
	MMFF94	OPLS-AA (TIP3P)	OPLS-AA (TIP5P)	AMOEBA 09	STATIC PACKING	MMFF94	OPLS-AA (TIP3P)	OPLS-AA (TIP5P)	AMOEBA 09	STATIC PACKING
Et–Et	12.0	11.8	11.8	12.0	5.6	12.0	11.8	11.8	12.0	5.6
EtOH–Et	12.9	12.6	12.6	12.6	5.8	10.4	10.3	10.3	10.2	5.3
EtOH–EtOH	11.4	11.1	11.1	11.0	5.6	11.4	11.1	11.1	11.0	5.6
H2O–Et	8.7	8.2	8.2	8.4	4.5	16.3	16.4	17.3	17.2	6.9
H2O–EtOH	7.2	7.0	7.0	6.8	4.3	17.9	17.8	18.4	18.7	7.3
H2O–H2O	11.0	10.9	10.8	10.8	5.5	11.0	10.9	10.8	10.8	5.5
Me–Et	10.0	9.8	9.8	9.7	4.9	14.9	14.8	14.8	14.6	6.3
Me–EtOH	8.8	8.6	8.6	8.3	4.7	16.1	15.8	15.8	15.6	6.6
Me–H2O	13.3	13.2	14.2	13.7	6.1	10.6	10.2	10.2	10.1	5.0
Me–Me	12.3	12.1	12.1	11.9	5.6	12.3	12.1	12.1	11.9	5.6
Me2O–Et	13.0	12.8	12.8	12.7	5.7	11.0	10.7	10.7	10.7	5.3
Me2O–EtOH	11.5	11.1	11.1	11.0	5.5	11.5	11.4	11.4	11.4	5.6
Me2O–H2O	18.0	17.8	18.7	18.5	7.3	7.5	7.2	7.3	7.5	4.3
Me2O–Me	16.3	15.9	15.9	15.7	6.6	9.3	8.9	8.9	8.8	4.7
Me2O–Me2O	12.2	11.6	11.6	11.7	5.6	12.2	11.6	11.6	11.7	5.6
MeOH–Et	11.0	10.8	10.8	10.8	5.2	12.9	12.7	12.7	12.3	5.9
MeOH–EtOH	9.4	9.3	9.3	9.3	5.0	13.3	13.4	13.4	13.0	6.2
MeOH–H2O	14.8	14.7	15.1	15.2	6.5	8.2	8.1	8.0	8.4	4.7
MeOH–Me	13.6	13.4	13.4	13.2	6.0	10.5	10.5	10.5	9.9	5.2
MeOH–Me2O	10.0	9.7	9.7	9.8	5.0	13.8	13.6	13.6	13.4	6.1
MeOH–MeOH	11.5	11.2	11.2	11.3	5.6	11.5	11.2	11.2	11.3	5.6
HAc–HAc	9.7	10.2	10.2	10.1	5.6	9.7	10.2	10.2	10.1	5.6

The static packing results from previous research are added for comparison only

of Materials Studio [42] using the Condensed-phase Optimized Molecular Potentials for Atomistic Simulation Studies (COMPASS) force field [43].

Figures 7 and 8 show the similarity of trends in coordination number assignment. While the MD-derived coordination numbers are similar for the different force fields used, the results for the COMPASS static packing approach are significantly reduced by about 50%. The spread of the values derived from MD (Fig. 7) is significantly higher than that of the static packing approach. Interestingly, linear mapping of the coordination numbers to the interval [0,1] yields an approximate overlap of the results (Fig. 8).

Repulsion parameters a_{ij}

For the different force fields, particle sets for mesoscopic DPD simulations with the isotropic mutual repulsions a_{ij} for methane (Me), ethane (Et), methanol (MeOH), ethanol (EtOH), dimethyl ether (Me2O), and water (H2O) were constructed. The off-diagonal a_{ij} values of a particle set were linearly scaled with MFsim [30] so that the maximum absolute deviation between the smallest a_{ij} value and the diagonal values $a_{ii} = 24.83$ (for a thermodynamic

temperature of 298 K) is 20. For the OPLS-AA force field the water models TIP3P and TIP5P were considered, denoted OPLS-AA (TIP3P) and OPLS-AA (TIP5P). An additional OPLS-AA (TIP5P) particle set, denoted OPLS-AA (TIP5P $Z_{ij} = 1$), was created which is solely based on the minimal averaged intermolecular energy $\langle E_{ij} \rangle$ with a fixed coordination number $Z_{ij} = 1$ for all dimers. With force field MM3, interaction energies can be calculated only, so a combined particle set was created using MM3 for interaction energy calculation and the MMFF94 force field for MD-based coordination number estimation, denoted MM3/MMFF94. The particle set from [41, 44], based on the COMPASS force field, is used for comparison.

Figure 9a–f display the different repulsion parameters a_{ij} . The red dashed line indicates the diagonal value of 24.83. A crucial difference between the different particle sets is the water–methanol (H2O–MeOH) repulsion: for the COMPASS, OPLS-AA (TIP5P) and OPLS-AA (TIP5P, $Z_{ij} = 1$) force fields, this repulsion is the smallest one (and thus the base value for off-diagonal repulsion parameter scaling), whereas for force fields OPLS-AA (TIP3P), AMOEBA09, MMFF94, and MM3/MMFF94

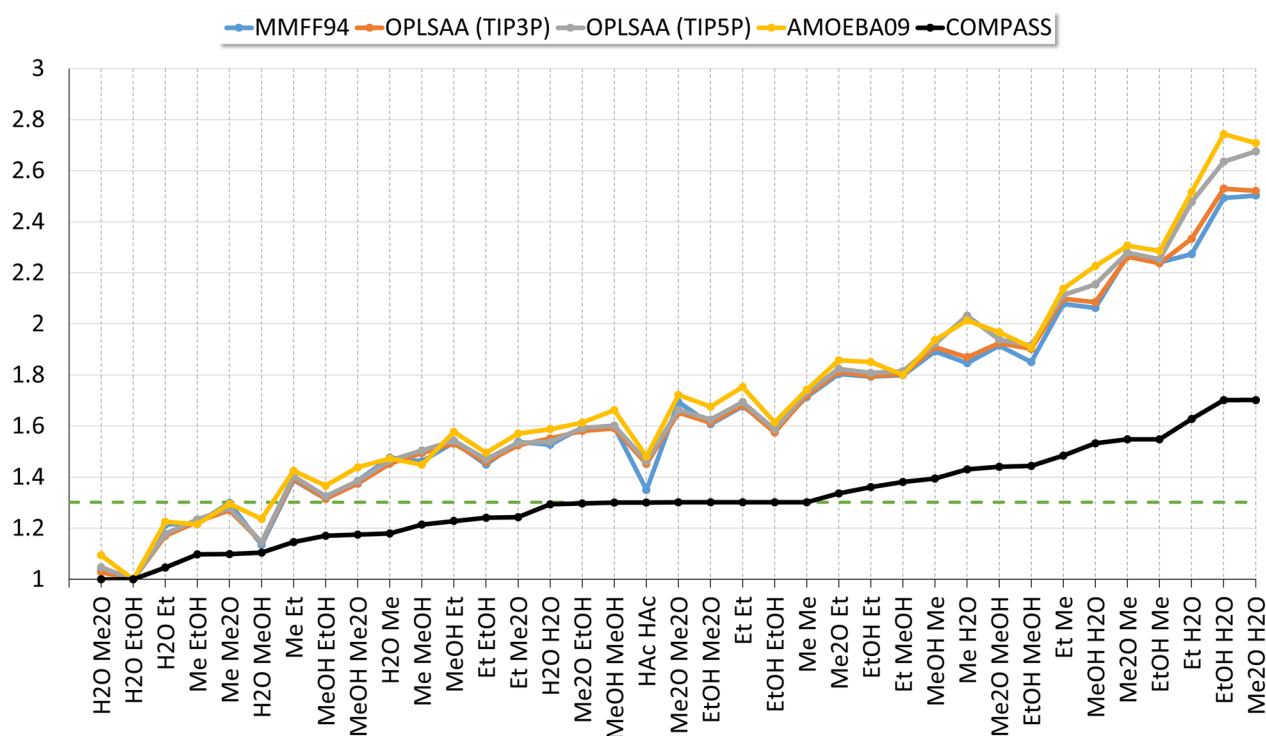


Fig. 7 Coordination numbers Z_{ij} , scaled to 1 for each force field. The black line indicates the COMPASS static packing coordination numbers for the homo dimers

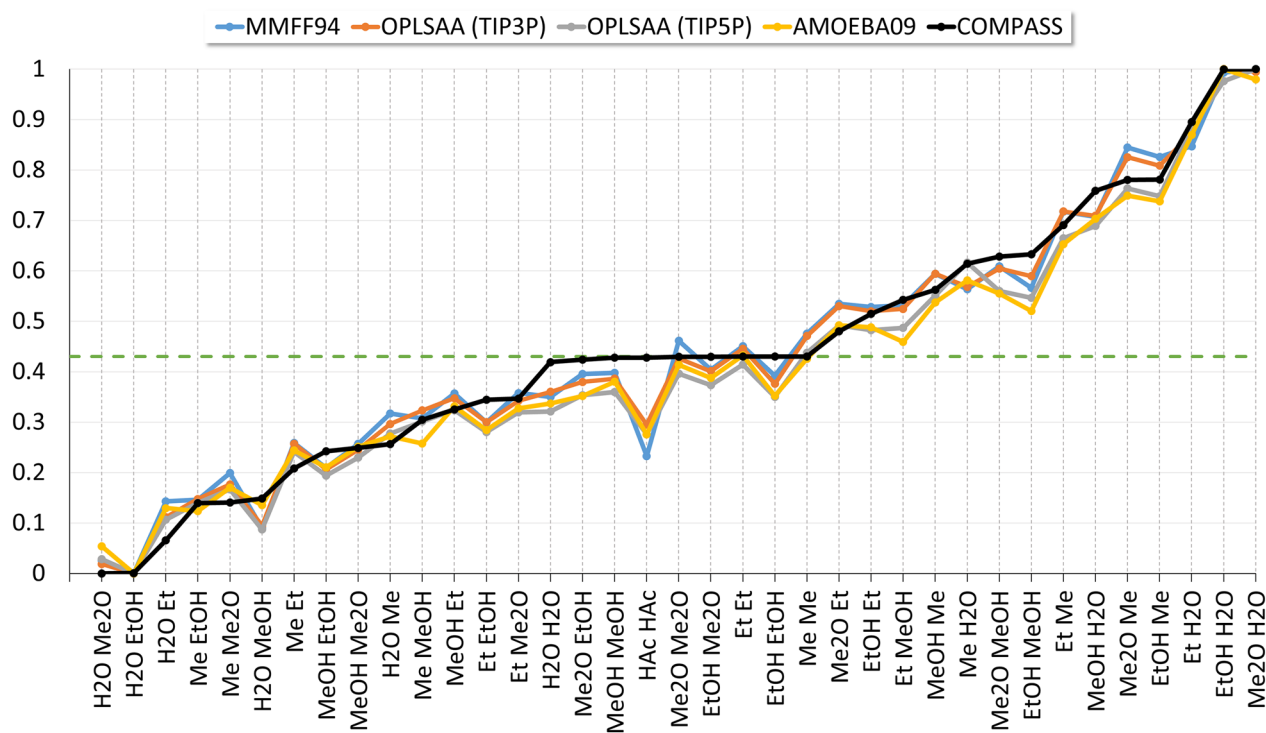


Fig. 8 Coordination numbers Z_{ij} , linearly mapped to interval [0,1] for each force field. The black line indicates the COMPASS static packing coordination numbers for the homo dimers

the water–dimethyl ether (H₂O–Me₂O) repulsion is minimal. Interestingly, the different water models TIP3P and TIP5P of the OPLS-AA force field exhibit this crucial difference, emphasizing their relevance. Note, that

the differences of the non-water repulsions for OPLS-AA (TIP3P) and OPLS-AA (TIP5P) are caused by the different scaling due to these different base values for off-diagonal repulsion parameter scaling. The COMPASS

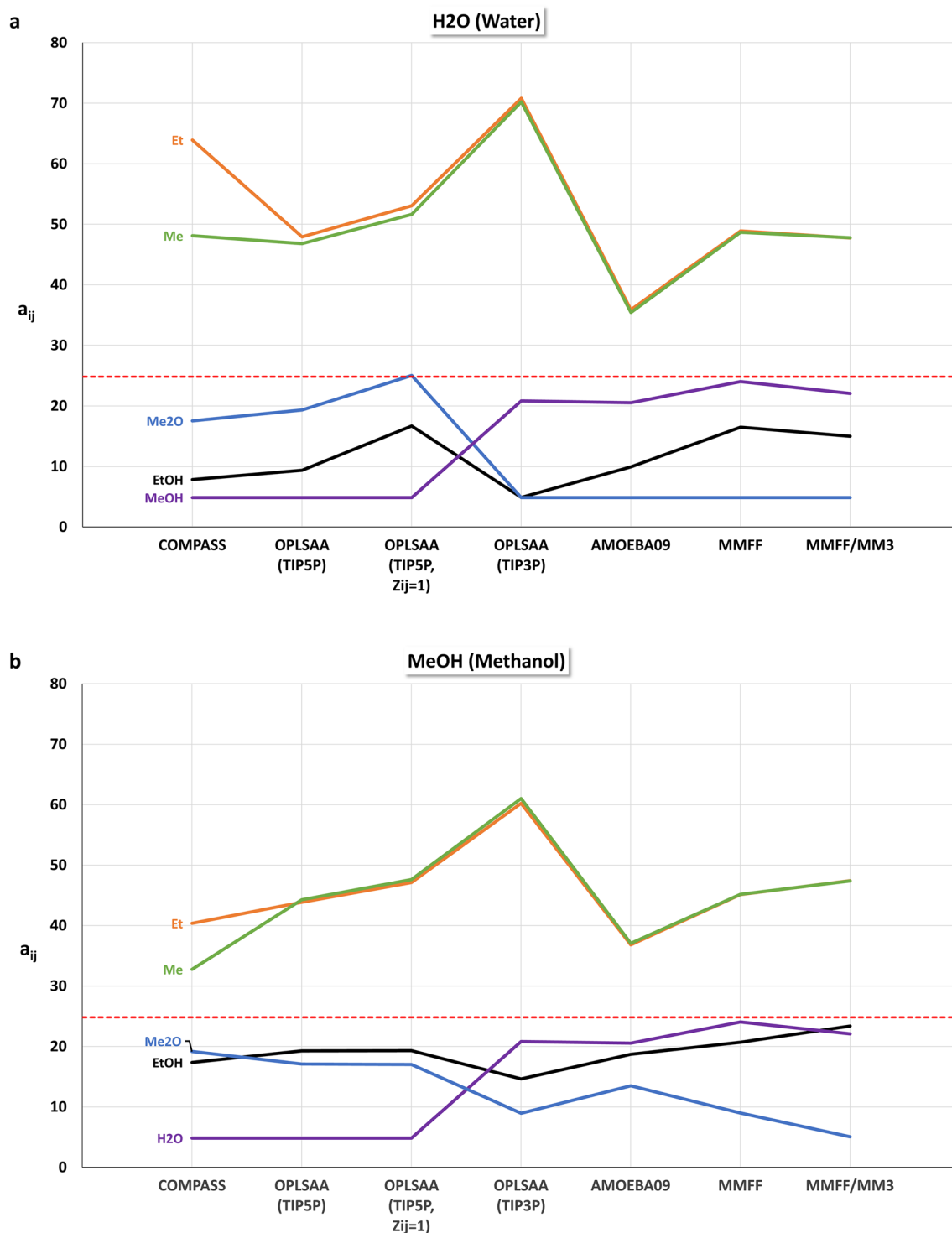


Fig. 9 Scaled repulsion parameters a_{ij} . The red dashed line indicates the diagonal value of 24.83 for homo dimers

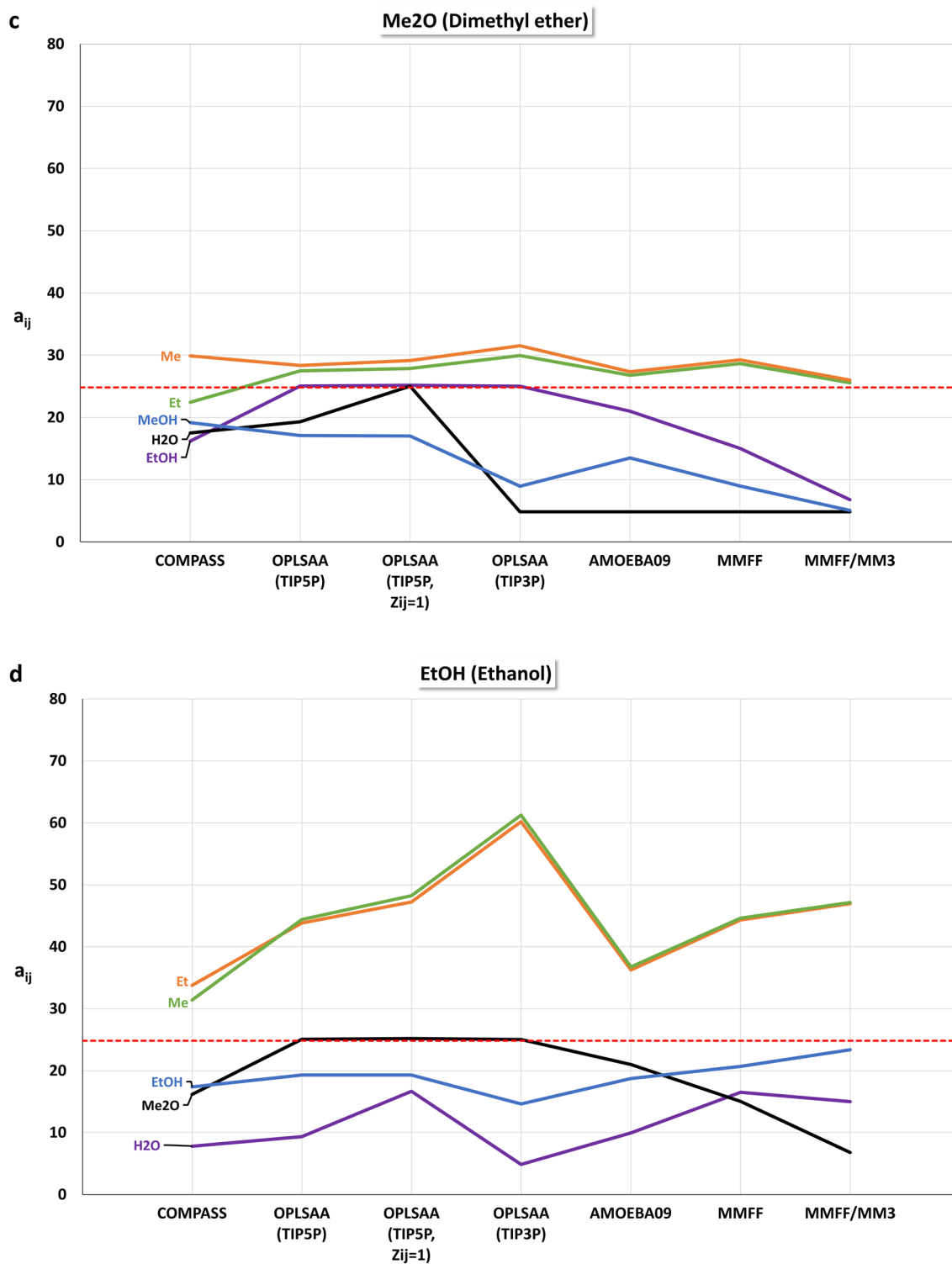


Fig. 9 continued

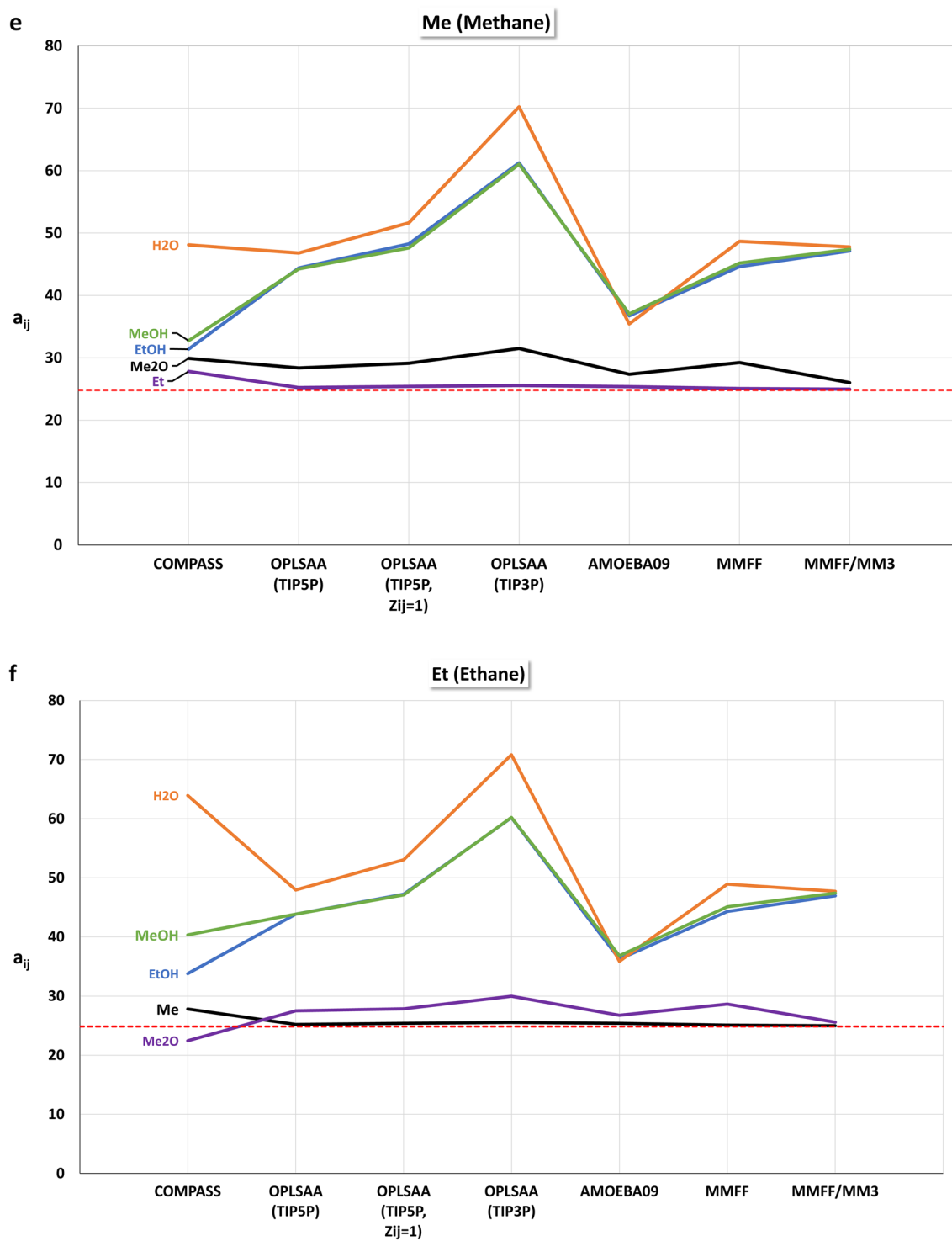


Fig. 9 continued

force field exhibits an extraordinary difference for the water–methane (H₂O–Me) and water–ethane (H₂O–Et) repulsions, which is otherwise not visible. A significant

difference between the OPLS-AA (TIP5P) and OPLS-AA (TIP5P, $Z_{ij} = 1$) force field is the water–dimethyl ether (H₂O–Me₂O) repulsion. These obvious differences

between the force fields lead to different levels of usefulness for DPD simulation approaches, where even a single difference in dimer interaction can become a decisive factor.

DPD simulations

The created particle sets were used for DPD simulations of mixtures of water with the non-ionic polyoxyethylene alkyl ether surfactant $C_{10}E_4$, where “ C_{10} ” indicates

the number of 10 carbon atoms in the alkyl chain of the lyophobic part, and “ E_4 ” represents the number of 4 lyophilic ethylene oxide units [44]. A stable lamellar L_α phase is formed by a $C_{10}E_4$ /water mixture around 298 K with a $C_{10}E_4$ mass fraction of 0.75. The performance of the particle sets for the different force fields may be evaluated by monitoring the emergent formation of $C_{10}E_4$ bilayers from initial random mixing in the simulation box [44]. The DPD simulations are carried out with the settings

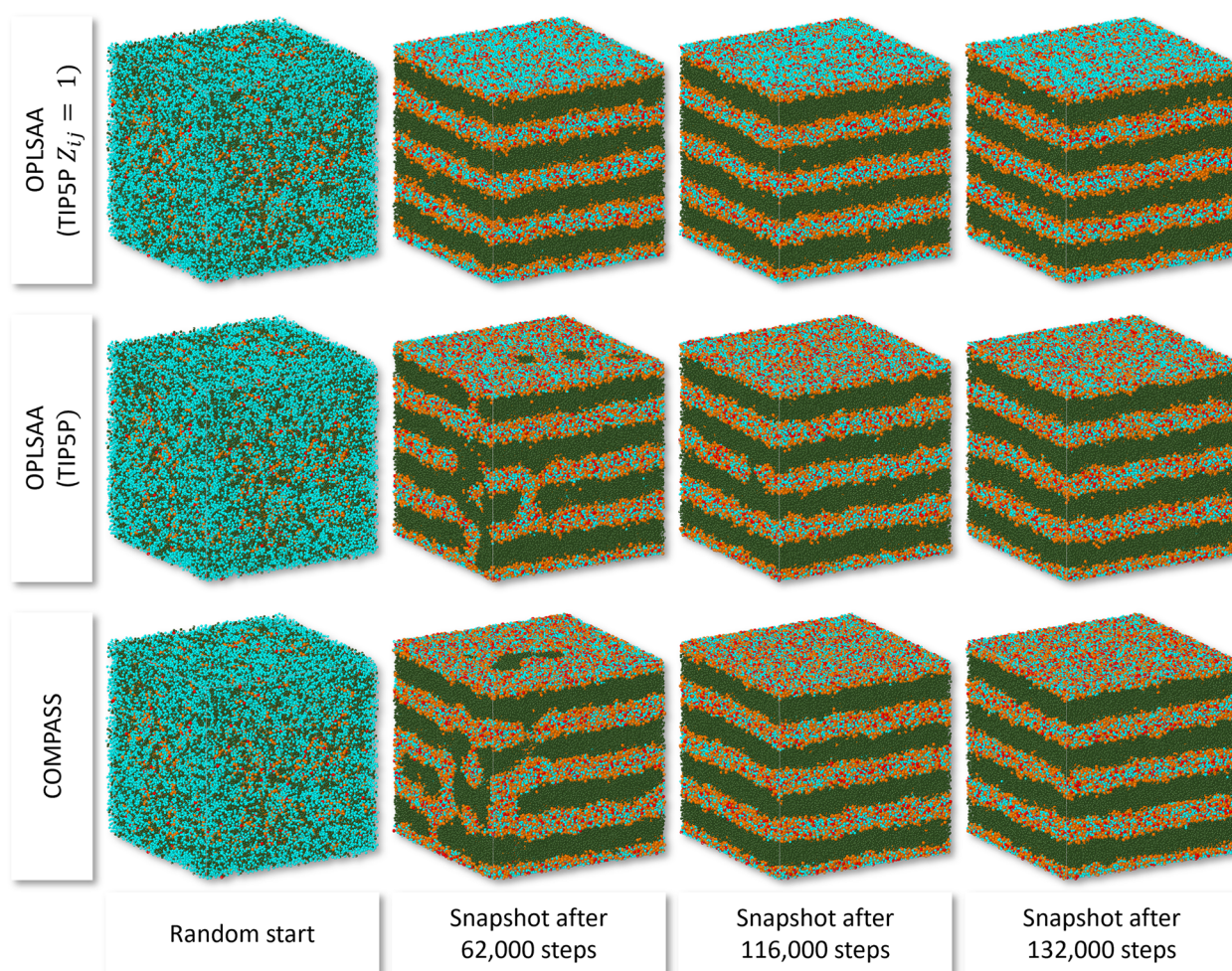


Fig. 10 Stacked bilayer superstructure formation from random mixing of a $C_{10}E_4$ /water mixture for particle sets OPLS-AA (TIP5P $Z_{ij} = 1$) (top row), OPLS-AA (TIP5P) (middle row), and COMPASS (bottom row) with front view of the simulation box with vertical z-axis. Particle colours: Me: Olive, Me2O: Orange, MeOH: Red, H2O: Cyan

Table 4 Bilayer convergence for different particle sets (with an integration step size of 0.04)

Force field	COMPASS	MMFF94*	MMFF94/MM3	AMOEA09*	OPLS-AA (TIP3P)	OPLS-AA (TIP5P)	OPLS-AA (TIP5P, $Z_{ij} = 1$)
Convergence [in 1000 simulation steps]	116	522	–	860	846	132	62

* Bilayer superstructure not parallel to xy-plane

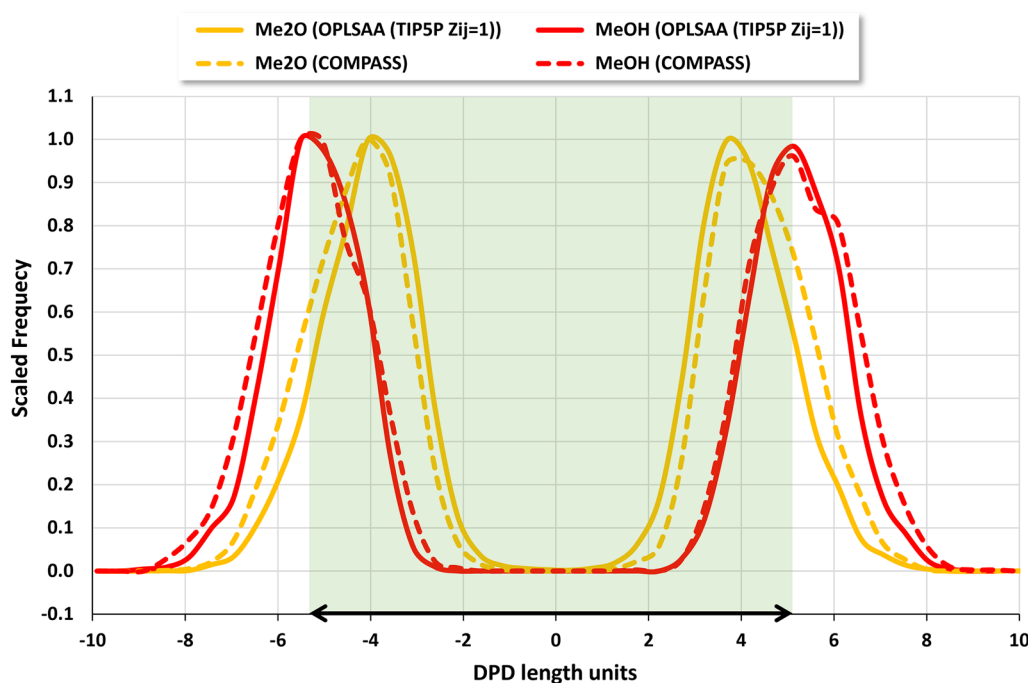


Fig. 11 Me2O and MeOH particle distribution snapshots along the z-axis perpendicular to a single emerged $C_{10}E_4$ bilayer for the OPLS-AA (TIP5P $Z_{ij} = 1$) (solid lines) and COMPASS (dashed lines) particle sets. The highlighted area corresponds to the bilayer width, indicated by the double arrow

outlined in [44], using the SPICES 9Me–4Me2O–MeOH fragmentation for $C_{10}E_4$ [45, 46], an integration step size of 0.04, and a deactivated periodic boundary in z-direction (to induce bilayer formation in the xy-plane).

For particle set OPLS-AA (TIP5P $Z_{ij} = 1$), a stacked bilayer superstructure emerges at simulation step 62,000, see Fig. 10 and Table 4, which was even below the COMPASS particle set from [44] with 116,000 steps, where the emerged bilayer structure corresponds well to the one reported in [44], see Fig. 11. The OPLS-AA (TIP5P) particle set required more than twice as many simulation steps (132,000) and the OPLS-AA (TIP3P) particle set more than tenfold as many (846,000 steps). The AMOEBA09 and MMFF94 particle sets show a bilayer superstructure formation, but the bilayers do not align parallel to the xy-plane (as induced by the periodic boundary conditions) but parallel to the yz-plane. Using the hybrid MM3/MMFF94 particle set, no bilayer was formed within 1,000,000 simulation steps. For the specified DPD simulation task, the OPLS-AA (TIP5P $Z_{ij} = 1$) particle set can be regarded as the most suitable choice, which is in good agreement with experimental findings.

Conclusions

The outlined automated comprehensive calculation of intermolecular interaction energies based on molecular force-fields shows satisfactory results for small

molecule interactions and can even be successfully used to estimate mesoscopic simulation parameters. Special care should always be taken, as individual force fields can lead to erroneous results. Therefore, despite all automation, manual checking of the results is still essential.

The new calculation pipeline can be easily extended to additional force fields (which may require their conversion into the Tinker format). Therefore, calculating differential pair interaction energies will advance with the progress in improving the underlying molecular force fields. Due to the modularized pipeline approach, alternative modeling packages can also be used for certain computational tasks if they can provide the required specific functions.

The robustness of the outlined computational pipeline can be seen as a crucial advance, as the estimation of differential pair interaction energies along alternative computational paths often led to ambiguous results, which in particular prevented the construction of consistent particle–particle repulsions for larger mesoscopic particle sets [47]. The construction of mesoscopic sets with dozens of particles (including hundreds or thousands of mutual particle repulsions) is now within practical reach.

Availability and requirements

Project name: Mesoscopic Interaction Parameter Estimation with Tinker for Java (MIPET4Java)

Project home page: <https://github.com/zielesny/MIPET>

Current version: v1.0.0.0

Operating system(s): Windows (×64), Linux (×64)

Programming language: Java

Other requirements: Java v17.0.4 or higher, Tinker

Molecular Modeling Package v8.10.2

Licence: GPL-3.0

Any restrictions to use by non-academics: While the backbone code of this project is not restricted to academic use, the Tinker Molecular Modeling Package is subject to corresponding restrictions, see [21] for license details.

Abbreviations

AMOEBA09	Atomic Multipole Optimized Energetics for Biomolecular Simulation force field 09
BSSE	Basis set superposition error
COMPASS	Condensed-phase Optimized Molecular Potentials for Atomistic Simulation Studies
CDK	Chemistry Development Kit
DFT	Density Functional Theory
DPD	Dissipative Particle Dynamics
LMOD	Low-mode conformational search
MD	Molecular dynamics
MM3	Molecular mechanics force field 3
MMFF94	Merck molecular force field 94
OCVM	Optimally Conditioned Variable Metric
OPLS-AA	All-atom optimized potentials for liquid simulations force field
RMS	Root mean square
SAPT	Symmetry-Adapted Perturbation Theory

Acknowledgements

The authors would like to thank Tim Clark for valuable initial discussions as well as his “door-opener” support, and especially Andreas Heßelmann for important methodological advice. Jay Ponder provided extensive help and advice on the Tinker package, for which we are very grateful. Karina van den Broek and Matthias Eppe kindly supported the initial concept, Veit Hucke thankfully contributed the rotation matrix code. Jan-Mathis Hein and Martin Urban generously helped with the molecular coding and the construction of start geometries. Jonas Schaub thankfully revised the GitHub repository and tested the installation procedure. Support from the Carl Zeiss Foundation and the Open Access Publication Fund of the Westphalian University of Applied Sciences is gratefully acknowledged. Lastly, the authors would like to thank the communities that created the open software libraries used in this work.

Author contributions

FB and MD designed, developed, and tested the calculation pipeline. HL helped with all force field and molecular dynamics related issues. CS and AZ conceived the study. All authors read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. This work was supported by the Carl-Zeiss-Foundation.

Availability of data and materials

The source code of the calculation pipeline and the description for downloading the Tinker Molecular Modeling Package are available on GitHub at <https://github.com/zielesny/MIPET>.

Declarations

Competing interests

AZ is co-founder of GNWI—Gesellschaft für naturwissenschaftliche Informatik mbH, Dortmund, Germany.

Author details

¹Institute for Bioinformatics and Chemoinformatics, Westphalian University of Applied Sciences, August-Schmidt-Ring 10, 45665 Recklinghausen, Germany. ²Erlangen National High Performance Computing Center (NHR@FAU), Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Martensstraße 1, 91058 Erlangen, Germany. ³Institute for Inorganic and Analytical Chemistry, Friedrich-Schiller University, Lessing Strasse 8, 07743 Jena, Germany.

Received: 13 February 2024 Accepted: 22 July 2024

Published online: 08 August 2024

References

- Engel T, Gasteiger J (2018) Chemoinformatics: basic concepts and methods. Wiley-VCH, Weinheim
- Stone AJ (2013) The theory of intermolecular forces, 2nd edn. Oxford University Press, Oxford
- Heßelmann A, Korona T (2014) Intermolecular symmetry-adapted perturbation theory study of large organic complexes. *J Chem Phys* 141:094107. <https://doi.org/10.1063/1.4893990>
- Heßelmann A (2018) DFT-SAPT intermolecular interaction energies employing exact-exchange Kohn-Sham response methods. *J Chem Theory Comput* 14:1943–1959. <https://doi.org/10.1021/acs.jctc.7b01233>
- Hill TL (1986) An introduction to statistical thermodynamics. Dover Publications, New York
- Groot RD, Warren PB (1997) Dissipative particle dynamics: bridging the gap between atomistic and mesoscopic simulation. *J Chem Phys* 107:4423–4435. <https://doi.org/10.1063/1.474784>
- Rackers JA, Wang Z, Lu C et al (2018) Tinker 8: software tools for molecular design. *J Chem Theory Comput* 14:5273–5289. <https://doi.org/10.1021/acs.jctc.8b00529>
- Davidon WC (1975) Optimally conditioned optimization algorithms without line searches. *Math Program* 9:1–30. <https://doi.org/10.1007/BF01681328>
- Kolossvary I, Guida WC (1999) Low-mode conformational search elucidated: application to C39H80 and flexible docking of 9-deazaguanine inhibitors into PNP. *J Comput Chem* 20:1671–1684. [https://doi.org/10.1002/\(SICI\)1096-987X\(19991130\)20:15%3C1671::AID-JCC7%3E3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1096-987X(19991130)20:15%3C1671::AID-JCC7%3E3.0.CO;2-Y)
- Gonzalez A (2010) Measurement of areas on a sphere using fibonacci and latitude-longitude lattices. *Math Geosci* 42:49–64. <https://doi.org/10.1007/s11004-009-9257-x>
- Halgren TA (1996) Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J Comput Chem* 17:490–519. [https://doi.org/10.1002/\(SICI\)1096-987X\(199604\)17:5:6%3C490::AID-JCC1%3E3.0.CO;2-P](https://doi.org/10.1002/(SICI)1096-987X(199604)17:5:6%3C490::AID-JCC1%3E3.0.CO;2-P)
- Riddick JA, Bunger WB, Sakano T, Weissberger A (1986) Organic solvents: physical properties and methods of purification, 4th edn. Wiley, New York
- PubChem Water. <https://pubchem.ncbi.nlm.nih.gov/compound/962>. Accessed 5 Feb 2024
- Zhao YH, Abraham MH, Zissimos AM (2003) Fast calculation of van der Waals volume as a sum of atomic and bond contributions and its application to drug compounds. *J Org Chem* 68:7368–7373. <https://doi.org/10.1021/jo034808o>
- Chemistry Development Kit. <https://cdk.github.io/>. Accessed 5 Feb 2024
- Willighagen EL, Mayfield JW, Alvarsson J et al (2017) The Chemistry Development Kit (CDK) v2.0: atom typing, depiction, molecular formulas, and substructure searching. *J Cheminform* 9:33. <https://doi.org/10.1186/s13321-017-0220-4>
- Andersen HC (1980) Molecular dynamics simulations at constant pressure and/or temperature. *J Chem Phys* 72:2384–2393. <https://doi.org/10.1063/1.439486>

18. Allen MP, Tildesley DJ (2017) Computer simulation of liquids, 2nd edn. Oxford University Press, Oxford
19. MIPET. <https://github.com/zielesny/MIPET>. Accessed 5 Feb 2024
20. MIPET-v1.0.0.0. <https://github.com/zielesny/MIPET/releases/tag/MIPET>. Accessed 5 Feb 2024
21. Tinker Molecular Modeling Package. <https://dasher.wustl.edu/tinker/>. Accessed 5 Feb 2024
22. MIPET/README.md at main zielesny/MIPET. <https://github.com/zielesny/MIPET/blob/main/README.md>. Accessed 5 Feb 2024
23. Wolfram Mathematica: Modern Technical Computing. <https://www.wolfram.com/mathematica/>. Accessed 5 Feb 2024
24. MIPET/Visualizaton Mathematica notebooks at main zielesny/MIPET. <https://github.com/zielesny/MIPET/tree/main/Visualizaton%20Mathematica%20notebooks>. Accessed 5 Feb 2024
25. (2022) AMD Ryzen™ Threadripper™ PRO 5995WX Prozessor. <https://www.amd.com/de/products/cpu/amd-ryzen-threadripper-pro-5995wx>. Accessed 5 Feb 2024
26. Frisch MJ, Trucks GW, Schlegel HB, et al (2016) Gaussian 16 Rev. C.01. <https://gaussian.com/gaussian16/>. Accessed 5 Feb 2024
27. Dennington R, Keith TA, Millam JM (2019) GaussView Version 6. <https://gaussian.com/gaussview6/>. Accessed 5 Feb 2024
28. Chai J-D, Head-Gordon M (2008) Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys Chem Chem Phys* 10:6615–6620. <https://doi.org/10.1039/B810189B>
29. MIPET/Gaussian job files at main zielesny/MIPET. <https://github.com/zielesny/MIPET/tree/main/Gaussian%20job%20files>. Accessed 5 Feb 2024
30. van den Broek K, Daniel M, Epple M et al (2020) MFSim—an open Java all-in-one rich-client simulation environment for mesoscopic simulation. *J Cheminformatics* 12:29. <https://doi.org/10.1186/s13321-020-00432-9>
31. van den Broek K, Kuhn H, Zielesny A (2018) Jdpd: an open java simulation kernel for molecular fragment dissipative particle dynamics. *J Cheminformatics* 10:25. <https://doi.org/10.1186/s13321-018-0278-7>
32. MIPET/C10E4-water bilayer formation study at main zielesny/MIPET. <https://github.com/zielesny/MIPET/tree/main/C10E4-water%20bilayer%20formation%20study>. Accessed 5 Feb 2024
33. Derissen JL (1971) A reinvestigation of the molecular structure of acetic acid monomer and dimer by gas electron diffraction. *J Mol Struct* 7:67–80. [https://doi.org/10.1016/0022-2860\(71\)90008-1](https://doi.org/10.1016/0022-2860(71)90008-1)
34. Allinger NL, Yuh YH, Lii JH (1989) Molecular mechanics. The MM3 force field for hydrocarbons. 1. *J Am Chem Soc* 111:8551–8566. <https://doi.org/10.1021/ja00205a001>
35. Jorgensen WL, Tirado-Rives J (1988) The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J Am Chem Soc* 110:1657–1666. <https://doi.org/10.1021/ja00214a001>
36. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 118:11225–11236. <https://doi.org/10.1021/ja9621760>
37. Jorgensen WL, Chandrasekhar J, Madura JD et al (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935. <https://doi.org/10.1063/1.445869>
38. Mahoney MW, Jorgensen WL (2000) A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J Chem Phys* 112:8910–8922. <https://doi.org/10.1063/1.481505>
39. Matthews GP, Smith EB (1976) An intermolecular pair potential energy function for methane. *Mol Phys* 32:1719–1729. <https://doi.org/10.1080/00268977600103031>
40. Fan CF, Olafson BD, Blanco M, Hsu SL (1992) Application of molecular simulation to derive phase diagrams of binary mixtures. *Macromolecules* 25:3667–3676. <https://doi.org/10.1021/ma00040a010>
41. Truszkowski A, Epple M, Fiethen A et al (2013) Molecular fragment dynamics study on the water–air interface behavior of non-ionic polyoxyethylene alkyl ether surfactants. *J Colloid Interface Sci* 410:140–145. <https://doi.org/10.1016/j.jcis.2013.07.069>
42. BIOVIA Materials Studio - BIOVIA - Dassault systèmes®. <https://www.3ds.com/products-services/biovia/products/molecular-modeling-simulation/biovia-materials-studio/>. Accessed 5 Feb 2024
43. Sun H (1998) COMPASS: an ab initio force-field optimized for condensed-phase applications overview with details on alkane and benzene compounds. *J Phys Chem B* 102:7338–7364. <https://doi.org/10.1021/jp980939v>
44. Bänsch F, Steinbeck C, Zielesny A (2023) Notes on molecular fragmentation and parameter settings for a dissipative particle dynamics study of a C10E4/water mixture with lamellar bilayer formation. *J Cheminformatics* 15:23. <https://doi.org/10.1186/s13321-023-00697-w>
45. van den Broek K, Daniel M, Epple M et al (2018) SPICES: a particle-based molecular structure line notation and support library for mesoscopic simulation. *J Cheminformatics* 10:35. <https://doi.org/10.1186/s13321-018-0294-7>
46. Zielesny A (2021) SPICES—A particle-based Molecular Structure Line Notation and Support Library for Mesoscopic Simulation. <https://github.com/zielesny/SPICES>. Accessed 5 Feb 2024
47. Truszkowski A (2015) Simulation von Peptiden, Proteinen und Biomembranen mit Molekularer Fragmentdynamik (MFD). Dissertation, University of Duisburg-Essen, Germany

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.