



Cite this: *Phys. Chem. Chem. Phys.*, 2025, 27, 1554

# Dissipative particle dynamics parametrisation using infinite dilution activity coefficients: the impact of bonding†

Rachel L. Hendrikse,<sup>a</sup> Carlos Amador<sup>b</sup> and Mark R. Wilson<sup>a</sup>

Dissipative particle dynamics (DPD) simulations have proven to be a valuable coarse-grained simulation technique for studying complex systems such as surfactant and polymer solutions. However, the best method to use in parametrising DPD systems is not universally agreed. One common approach is to map infinite dilution activity coefficients to the DPD simulation 'beads' that represent molecular fragments. However, we show that here that this approach can lead to serious errors when bonding beads together to create molecules. We show errors arise from the overlaps between bonded beads, which alters their solubility. In this article, we demonstrate how these bonding errors can be accounted for when defining DPD force fields using simple theoretical methods to account for the overlapping volumes, and we demonstrate the validity of our approach by calculating the partition coefficients for a series of solutes into two immiscible solvents.

Received 2nd October 2024,  
 Accepted 15th December 2024

DOI: 10.1039/d4cp03791j

[rsc.li/pccp](https://rsc.li/pccp)

## 1 Introduction

The dissipative particle dynamics (DPD) simulation method was initially introduced by Hoogerbrugge and Koelman,<sup>1</sup> and it has been continuously developed since then.<sup>2–5</sup> DPD is a mesoscopic simulation technique in which groups of atoms are represented by single beads, which is intended to reduce the computational demand of simulating single atoms. Mesoscale modelling provides the bridge between the atomistic methods and the macro-scale, retaining molecular details which are lost at a continuum scale, but making longer time and length scales more accessible than when using full atomistic modelling.

When first introduced, DPD was primarily used for qualitative study, with DPD parameters chosen to investigate general behaviours of systems such as surfactants<sup>6,7</sup> and polymer<sup>8–11</sup> solutions. These molecules are typically simulated using a simple bead-and-spring model, allowing one to bond various 'beads' together to study the behaviour of larger molecules. This has proven to be a largely successful approach for studying these types of systems, including the prediction of liquid crystal phase formation<sup>6,8,10</sup> and the nature of micellar solutions.<sup>7,10</sup>

A convenient method of relating atomistic simulations to mesoscopic simulation was presented by Groot and Warren<sup>2</sup>

who provided a mapping between DPD parameters and Flory–Huggins parameters. This mapping allows one to determine Flory–Huggins parameters using solubility parameters from molecular dynamics (or even experimental data) to determine how DPD beads should interact to model realistic systems. This approach has been widely used<sup>12–16</sup> in the literature as a means of simulating 'real molecules' using DPD, enabling quantitative predictions for various systems. This has made it possible to accurately calculate properties such as interfacial tension,<sup>12–14</sup> critical micelle concentrations<sup>14</sup> and other micelle properties<sup>14</sup> for surfactant systems, as well as polymer properties such as radius of gyration<sup>15,17</sup> and the end-to-end distance.<sup>18</sup>

While the mapping procedure provided by Groot and Warren<sup>2</sup> has been widely used, there are limitations to its applicability for different systems. For example, the original expressions they provided for the mapping were determined on the basis of simulations which relied on the phase separation of beads, meaning this approach is difficult to use for determining the interaction between two soluble bead types. They also performed simulations over a limited range of DPD interaction parameters, choosing to focus on ranges where linear relationships can be fitted to their calculated values. Therefore, strictly, their expressions should only be used in the range in which they were intended. Although, they are often used outside of this range. Other methods that have since been used to determine realistic DPD parameters include mapping to infinite dilution activity coefficients,<sup>19–23</sup> as an alternative to Flory–Huggins parameters. This has the benefit of being generally applicable to a greater range of system types, including bead types that are

<sup>a</sup> Department of Chemistry, Durham University, Durham, DH1 3LE, UK.

E-mail: [rachel.hendrikse@durham.ac.uk](mailto:rachel.hendrikse@durham.ac.uk)

<sup>b</sup> Procter and Gamble, Newcastle Innovation Centre, Whitley Road, Newcastle upon Tyne, NE12 9BZ, UK

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d4cp03791j>



soluble in one another. Another approach is by mapping interactions to octanol–water partition coefficients,<sup>24–26</sup> which has been shown to be successful for studying the behaviour of various surfactant systems.<sup>24–28</sup> Alternatively, others have calculated DPD parameters by matching to experimental liquid densities of mixtures.<sup>29,30</sup>

It is of particular interest in this work that, no matter which parametrisation approach is used, most existing DPD studies neglect the impact that bonding has on the resulting solubility of molecules. However, this has been shown to be important,<sup>31</sup> not least because the length of the bond can greatly influence the degree of overlap between the bonded beads. Bonding beads together has been shown to influence their solubility because of this overlapping bead volume. This oversight can lead to an incorrect determination of DPD parameters for bonded molecules, which is crucial for modelling the correct behaviour of systems. This is particularly true for systems involving shorter bond lengths, which typically result from a more finely coarse-grained system. Hence, dealing with this issue is essential for using DPD in quantitative studies of soft matter systems.

In this work, we show that DPD parameters can be accurately assigned based on their known infinite dilution activity coefficients, by using Widom insertion together with a simple approach to take into account the overlap of bonded beads. We validate our approach by simulating the partitioning of various solute molecules into two immiscible solvents, allowing us to determine partition coefficients as a means to check the predicted chemical potential of bonded beads and hence to check their infinite dilution activity coefficients. We suggest that this simple approach can make a major improvement to the parametrisation of DPD models.

## 2 Dissipative particle dynamics

### 2.1 Overview

In dissipative particle dynamics (DPD) atoms and molecules are modelled by simulating ‘DPD’ beads, where a single bead can represent a number of atoms. The total force acting on a single unbonded DPD ‘bead’ is the sum of the conservative  $\mathbf{F}_{ij}^C$ , dissipative  $\mathbf{F}_{ij}^D$ , and random force  $\mathbf{F}_{ij}^R$ . All of these forces act within a defined cut-off range  $R_C$ , beyond which the interaction force is zero. We choose to set the cut-off  $R_C = 1$ , in line with existing literature. Additional bonding  $\mathbf{F}_{ij}^B$  and angle forces  $\mathbf{F}_{ijk}^A$  contribute to longer molecules represented by a series of bonded beads.

The conservative force takes the form

$$F_{ij}^C = a_{ij} \left( 1 - \frac{|\mathbf{r}_{ij}|}{R_C} \right) \hat{\mathbf{r}}_{ij}, \quad (1)$$

where  $a_{ij}$  is the the magnitude of the interaction, the vector between beads  $i$  and  $j$  is calculated as  $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$  and the unit vector  $\hat{\mathbf{r}}_{ij} = \mathbf{r}_{ij}/|\mathbf{r}_{ij}|$ . The conservative force assigns a chemical identity to the beads *via* varying the value of the  $a_{ij}$  parameter, which can be defined using various approaches which will be

discussed in Section 2.2. The forces  $\mathbf{F}_{ij}^D$  and  $\mathbf{F}_{ij}^R$  are given by the following expressions.

$$\mathbf{F}_{ij}^D = -\gamma \omega^D(r_{ij}) (\hat{\mathbf{r}}_{ij} \cdot \mathbf{v}_{ij}) \hat{\mathbf{r}}_{ij}, \quad (2)$$

$$\mathbf{F}_{ij}^R = \sigma \omega^R(r_{ij}) \zeta_{ij} \hat{\mathbf{r}}_{ij} \Delta t^{-1/2}, \quad (3)$$

where  $\omega^D$  and  $\omega^R$  are weight functions that vanish when  $|r_{ij}| > R_C$ .  $\gamma$  is a friction coefficient and  $\sigma$  is the noise amplitude,  $\mathbf{v}_{ij} = \mathbf{v}_i - \mathbf{v}_j$  is the velocity between beads  $i$  and  $j$ ,  $\Delta t$  is the time step, and  $\zeta_{ij}(t)$  is a randomly fluctuating Gaussian variable with zero mean and unit variance.

In this work, we perform all of our simulations using the LAMMPS software, where the form of the weight function used is

$$\omega^D = \left( 1 - \frac{|r_{ij}|}{R_C} \right)^2. \quad (4)$$

To satisfy the fluctuation–dissipation theorem, the relationship between the weight functions must obey<sup>3</sup>

$$\omega^D = [\omega^R]^2 \quad (5)$$

and the relationship between the amplitudes is

$$\sigma^2 = 2\gamma k_B T, \quad (6)$$

where  $k_B$  is the Boltzmann constant and  $T$  is the temperature. In this work, we set  $k_B T = 1$  and set values for the constants  $\sigma = 3$  and  $\gamma = 4.5$ . For all simulations, we choose a time step of  $\Delta t = 0.01$ .

For bonded molecules, we also impose ‘bonding’ and ‘angle’ potentials taking the form of a harmonic spring,

$$U_{ij}^B = K_B (r_{ij} - l_0)^2, \quad (7)$$

$$U_{ijk}^A = K_A (\theta_{ijk} - \theta_0)^2, \quad (8)$$

where  $l_0$  is an equilibrium bond length,  $\theta_{ijk}$  is a bond angle between beads  $i, j$  and  $k$ , and  $\theta_0$  is an equilibrium angle, which we set as  $180^\circ$ . The constants  $K_B$  and  $K_A$  define the strength of the harmonic spring, and in this work, we choose  $K_B = 150k_B T/r_C^2$  and  $K_A = 5k_B T/\text{rad}^2$ .

It is worth noting that, by default, bonded beads do not directly interact with each other in LAMMPS (except *via* the bonded interaction). This differs from other DPD software available. One of the significant drawbacks of having bonded beads interacting is that the resulting bond length can be significantly larger than what is set using the equilibrium bond length  $l_0$  parameter. While bonded interactions can be switched on in LAMMPS, for the bulk of this work we choose to leave bonded beads as non-interacting for this reason. However, as discussed below, the resulting bond length is also slightly different from  $l_0$ . We also note that it would be equally possible to leave bonded beads as interacting with each other, but this requires adjusting the setting of  $l_0$  accordingly to generate the correct effective distance between bonded beads. Given that this correction becomes dependent on a variety of factors, including the interaction  $a_{ij}$  parameter between the bonded beads themselves, this can be somewhat difficult in practice.



## 2.2 Determining $a_{ij}$ values (cross interactions)

In the existing literature, the most common approach used to determine DPD interaction parameters  $a_{ij}$  is from their Flory–Huggins  $\chi$  parameters, as first suggested by Groot and Warren.<sup>2</sup> In Flory–Huggins theory the total free energy of mixing is given as

$$\Delta G = k_B T \left( \frac{\phi_1}{n_1} \ln \phi_1 + \frac{\phi_2}{n_2} \ln \phi_2 + \chi \phi_1 \phi_2 \right) \quad (9)$$

where  $\phi_1$  and  $\phi_2$  are the volume fractions of the components 1 and 2,  $n_1$  and  $n_2$  are the number of segments per molecule. The  $\chi$  parameter accounts for the interaction between 1 and 2, and larger values indicate poor mixing. Groot and Warren also derived an expression relating  $\chi$  to the excess repulsion parameter  $\Delta a$  (where  $a_{ij} = a_{ii} + \Delta a$ ), where

$$\chi = 0.286 \Delta a. \quad (10)$$

for bead density  $\rho = 3R_C^{-3}$ . However, although this relation was derived for unbonded beads, many authors have applied it to bonded molecules. Additionally, little attention is often paid to the impact that the choice of equilibrium bond length has on the solubility behaviour.

Flory–Huggins parameters can be obtained from experimental measurement, but other developments have looked at alternative ways to calculate  $\chi$  when it is not experimentally available. This includes using Hildebrand solubility parameters  $\delta$ ,

$$\chi_{ij} = \frac{v_b}{k_B T} (\delta_i - \delta_j)^2 \quad (11)$$

where  $v_b$  is the average bead volume. The solubility parameter can be determined *via* a variety of approaches, including experimental methods,<sup>16,32</sup> molecular dynamics simulations,<sup>13,32</sup> or *via* COSMO-RS calculations.<sup>19,22</sup>

While the original Flory–Huggins parameterisation approach has proven to be extremely popular across literature, alternative approaches are available. One example is that of Travis *et al.*<sup>33</sup> where the conservative DPD interaction is based on regular solution theory (RST), which has the benefit of removing the restriction of equal repulsive interactions between like beads. Another is to determine  $a_{ij}$  using infinite dilution activity coefficients (IDAC),<sup>19,22</sup> will be the method of choice in our work.

The chemical potential  $\mu_A^*$  of a substance A in substance B is calculated as

$$\mu_A^* = \mu_A^\ominus + k_B T \ln(\gamma_A x_A) \quad (12)$$

where  $\mu_A^\ominus$  is the chemical potential of pure component A,  $\gamma_A$  is the activity coefficient,  $x_A$  is the mole fraction of A in B, and the units of the chemical potential  $\mu$  are energy per particle. The chemical potential can be calculated as a sum of two contributions

$$\mu = \mu_{\text{ex}} + \mu_{\text{id}} \quad (13)$$

where  $\mu_{\text{id}}$  is the chemical potential of an ideal gas, calculated as  $\mu_{\text{id}} = k_B T \ln \rho \Lambda^3$  where  $\Lambda$  is the thermal de Broglie wavelength and  $\rho$  is the number density.  $\mu_{\text{ex}}$  is the excess chemical potential, calculated using the Widom insertion method.

In the Widom insertion method, at each step, a ‘ghost’ particle is inserted at a random position in the simulation box.

The energy difference  $\Delta U$  between the system with and without this ghost particle is calculated, and the change in potential energy  $\Delta U$  comes from the potential interactions of the ghost particle with the actual particles in the box. The excess chemical potential can then be derived from the probability of a particle insertion being favourable, which depends on the average Boltzmann-weighted energy cost of insertion. The excess chemical potential  $\mu_{\text{ex}}$  is therefore calculated as

$$\mu_{\text{ex}} = -k_B T \ln \left\langle \exp \left( -\frac{\Delta U}{k_B T} \right) \right\rangle \quad (14)$$

where the angled brackets represent an average over many insertions. The process of inserting particles at random positions ensures the calculation captures a statistically relevant distribution of energy changes. This expression weights more favourable configurations more heavily, reflecting how particles naturally distribute themselves within a system. Therefore, one can use this method to determine the  $a_{ij}$  values which reproduce the correct IDAC values, which could be obtained *via* either experiment<sup>19,34</sup> or COSMO<sup>22,31</sup> approaches.

This approach has significant benefits over the method of relating Flory–Huggins parameters to their  $\Delta a$  value. Namely, the Groot and Warren<sup>2</sup> relation was derived by performing a sweep in the range  $8 \leq \Delta a \leq 25$ . Crucially, this relation was derived only for  $\Delta a$  values leading to phase separation, and therefore might not be reliable to determine interactions between components that are soluble within one another. The relation between  $\Delta a$  and  $\chi$  also becomes non-linear at large values of  $\Delta a$ ,<sup>35</sup> which is not accounted for in eqn (10).

It was shown previously,<sup>36</sup> that the Flory–Huggins  $\chi$ -parameter which is calculated from simulated phase behaviour agrees very well with the free energy difference between a monomer surrounded by solvent particles, and a solvent particle surrounded by solvent particles (*i.e.* essentially an approximate IDAC), for particular cases. Another interesting study<sup>35</sup> used the fact that, theoretically, one can relate the  $\chi$  parameter in terms of excess chemical potentials<sup>36</sup>

$$\chi_{ij} = \frac{1}{k_B T} \left( \frac{\Delta \mu_{\text{ex}}^{ii} - \Delta \mu_{\text{ex}}^{ij}}{V_i} + \frac{\Delta \mu_{\text{ex}}^{jj} - \Delta \mu_{\text{ex}}^{ij}}{V_j} \right) \quad (15)$$

where  $\mu_{\text{ex}}^{ij}$  is the excess chemical potential of species  $i$  in solvent  $j$ , and  $V_i$  is the molecular volume of  $i$ . Widom insertion can be used to calculate  $\mu_{\text{ex}}^{ij}$  as a function of  $a_{ij}$ , thereby providing another means to relate  $\chi$  to  $a_{ij}$ . Liyana-Arachchi *et al.*<sup>35</sup> compared the  $a_{ij}$  parameters which result from using eqn (10) *vs.* eqn (15), finding that the latter is significantly more accurate over a wider range of parameters/densities.

However, the approach of mapping to IDACs is reliant on the ability to correctly calculate the excess chemical potential  $\mu_{\text{ex}}^M$  of a molecule, which is the main consideration of the current work. Furthermore, it is worth noting that the method relies on the average  $\langle \exp(-\Delta U/k_B T) \rangle$ , which can be computationally sensitive. If  $\Delta U$  varies widely (such as for molecules of bonded beads), the exponent can lead to a broad distribution, which can make accurate averaging difficult due to large variances.



Therefore, a high number of insertions is required to converge to a reliable  $\mu_{\text{ex}}$ .

### 2.3 Effects of bonding

It is often assumed in the literature that one can determine the interactions of the beads individually (*i.e.* each bead represents a fragment of the molecule) and that bonding them together is a good representation of the molecule as a whole. From the point of view of using the Widom insertion technique to map to IDACs, this is equivalent to suggesting that the sum of the excess chemical potentials  $\mu_{\text{ex}}^{\text{S}}$  of beads which are bonded into a chain is equivalent to the total excess chemical potential for the whole molecule  $\mu_{\text{ex}}^{\text{M}}$ , *i.e.*

$$\mu_{\text{ex}}^{\text{M}} = \sum_{i=1}^N \mu_{\text{ex},i}^{\text{S}} \quad (16)$$

where  $N$  is the number of bonded beads in the chain.

While there are authors who have tried to quantify the effect that bonding has on the system properties and behaviours,<sup>37,38</sup> to the best of our knowledge, one of the only studies in existing literature which attempts to take into account the impact of bonding on solubility is that of Saathoff.<sup>31</sup> Saathoff<sup>31</sup> argues that, due to the overlapping nature of DPD beads, eqn (16) does not hold and the resulting chemical potential of bonded molecules is lower than expected. Therefore, corrections must be applied to  $\mu_{\text{ex}}$  to produce the desired value of  $\ln \gamma$ . Saathoff<sup>31</sup> achieved this by performing a series of simulations in which  $\mu_{\text{ex}}$  was calculated for bonded pairs of beads in a solvent, while the length of the bond was systematically varied. The resultant data was fitted to determine expressions to apply corrections to  $\mu_{\text{ex}}$ . However, here we argue that the corrections required can be accounted for by simply calculating the expected volume of the beads, which are overlapping (as a function of bond length), rather than depending on fitted expressions.

## 3 Calculating the chemical potential

In this section, we first consider how one might determine the excess chemical potential *via* Widom insertion for single, unbonded beads. Then we discuss our theoretical approach for how one might calculate the chemical potential of a series of bonded beads, taking into account their overlapping volumes as a result of bonding.

### 3.1 Single beads

To find a general relationship between  $a_{ij}$  and the excess chemical potential  $\mu_{\text{ex}}$  for single beads, we perform a parameter sweep in the range  $10 \leq a_{ij} \leq 100$ , with an interval of  $\Delta a_{ij} = 5$ . Following the standard approach,<sup>2</sup> we set all self-interactions to  $a_{ii} = 25$ . These simulations are conducted in the constant-NVT ensemble, in a cubic box with periodic boundaries and edge length  $20R_{\text{C}}$ . The beads are initialised with a random initial configuration using  $n = 24\,000$  simulation beads (box density  $\rho = 3R_{\text{C}}^{-3}$ ).

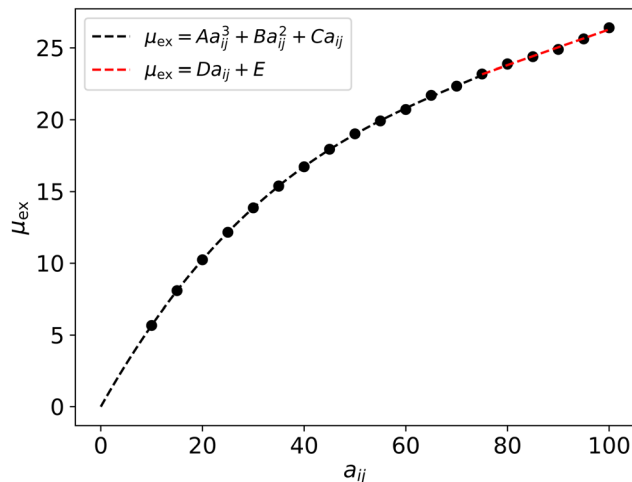


Fig. 1 Calculated relationship between the excess chemical potential  $\mu_{\text{ex}}$  and interaction parameter  $a_{ij}$  for the Widom insertion of single beads, for a solvent which interacts with  $a_{ij}$ .

From performing Widom insertions of single beads into a bath of non-bonded beads we determine how  $\mu_{\text{ex}}$  depends on the choice of cross-interaction parameter  $a_{ij}$ , which is shown in Fig. 1. We find that at higher values of  $a_{ij}$  the relation between  $a_{ij}$  and  $\mu_{\text{ex}}$  can be described linearly, while at lower values it cannot.

Therefore for  $a_{ij} < 70$ , we fit an expression of the form

$$\mu_{\text{ex}} = Aa_{ij}^3 + Ba_{ij}^2 + Ca_{ij} \quad (17)$$

which imposes that  $\mu_{\text{ex}} = 0$  when  $a_{ij} = 0$ . We calculate the coefficients to be  $A = 2.8248 \times 10^{-5}$ ,  $B = -6.391 \times 10^{-3}$  and  $C = 6.286 \times 10^{-1}$ . When  $a_{ij} \geq 70$  we fit the linear expression  $\mu_{\text{ex}} = Da_{ij} + E$  which has gradient  $D = 0.1244$  and intercept value  $E = 13.84$ .

It is worth noting that the fit shown in Fig. 1 is strictly speaking only valid for when  $a_{ii} = 25$ , and will vary slightly for different values of the solvent interaction. While intuitively one might expect eqn (17) to hold for all  $a_{ii}$  (due to the bulk density being set as  $\rho = 3R_{\text{C}}^{-3}$  in all cases), a change in  $\mu_{\text{ex}}$  results from slight differences in the radial distribution function at different  $a_{ii}$ . Examples of other solvent  $a_{ii}$  interactions can be found in the ESI.† However, in the partitioning calculations presented in this article, we only use  $a_{ii} = 25$  for single beads. We will show in Section 4.3 that sometimes it is necessary to use a different self-interaction value for the solvent when the beads are bonded, however, we will also show that eqn (17) still holds for the cases studied in this article.

### 3.2 Bonded beads

Suppose you have two overlapping spheres each with radius  $R$  which are separated by  $d$ , the volume of the overlapping portion is

$$V = \frac{\pi}{12}(4R + d)(2R - d)^2. \quad (18)$$

Now suppose you have a series of  $N$  bonded beads with cut-off radius  $R_{\text{C}}$  and bond length  $l_0$ . If  $l_0 \geq R_{\text{C}}$ , the volume of the



overlapping beads (for the whole molecule)

$$V_2 = \frac{\pi(N-1)}{12}(4R_C + l_0)(2R_C - l_0)^2 \quad (19)$$

while the non-overlapping volume can be calculated as

$$V_1 = \frac{N4\pi R_C^3}{3} - 2V_2. \quad (20)$$

The derivation of these expressions for volume is presented in more detail in the ESI.† For beads with interaction value  $a_{ij}$ , the interaction value inside the overlapping portions is effectively  $2a_{ij}$ . Using eqn (17) one can determine  $\mu_{\text{ex}}(a_{ij})$  and  $\mu_{\text{ex}}(2a_{ij})$ . From this, we hypothesise that the value of excess chemical potential for the molecule  $u_{\text{ex}}^{\text{M}}$  can be determined as

$$u_{\text{ex}}^{\text{M}} = \frac{3}{4\pi R_C^3}(V_1\mu_{\text{ex}}(a_{ij}) + V_2\mu_{\text{ex}}(2a_{ij})) \quad (21)$$

This expression relates to the assumption that we expect the excess chemical potential  $\mu_{\text{ex}}$  to be directly proportional to the volume of the inserted bead or fragment. We justify this assumption with additional theory provided in the ESI.†

If  $l_0 < R_C$ , then for molecules of length  $N \geq 3$ , there is the possibility of three overlapping spheres, as illustrated in Fig. 2.

The volume of the region where three beads overlap can be calculated as

$$V_3 = \frac{2\pi(N-2)}{3}(2R_C + l_0)(R_C - l_0)^2 \quad (22)$$

while the expression for the regions consisting of two overlapping beads,  $V_2$ , becomes

$$V_2 = \frac{\pi(N-1)}{12}(4R_C + l_0)(2R_C - l_0)^2 - 2V_3 \quad (23)$$

and the volume for the non-overlapping regions,  $V_1$ , becomes

$$V_1 = \frac{N4\pi R_C^3}{3} - 2V_2 - 3V_3. \quad (24)$$

Similarly to before, this allows us to calculate an excess chemical potential for the molecule as

$$u_{\text{ex}}^{\text{M}} = \frac{3}{4\pi R_C^3}(V_1\mu_{\text{ex}}(a_{ij}) + V_2\mu_{\text{ex}}(2a_{ij}) + V_3\mu_{\text{ex}}(3a_{ij})) \quad (25)$$

Once again further details of the derivation of eqn (22)–(24) can be found in the ESI.†

One simplification we make in our approach is that we use the bond length setting in eqn (7) to determine the distance between beads, and hence calculate the overlap. In practice,

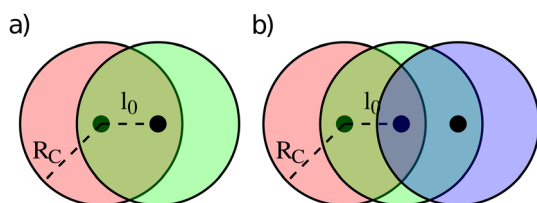


Fig. 2 Overlapping beads when  $l_0 < R_C$ , for cases where  $N = 2$  (a) and  $N = 3$  (b).

over the simulation period the bond length will fluctuate. In dissipative particle dynamics, there is normally a relatively broad distribution of bond lengths, depending on the strength chosen for the bonding. This means that the volume of overlap fluctuates also. Given that the average volume of overlap is not exactly the same as that determined from the average bond length (since volume scales with  $l_0^3$ ), our approach is a slight approximation. However, we show in this article that simply using  $l_0$  for the volume calculation is still an effective method for correcting the chemical potential of the beads.

One further simplification that we have made is to derive eqn (25) under the assumption that conformational effects can be neglected for relatively small molecules. However, it is worth noting that as the molecule grows in size, conformational effects may become more important. This is likely to be true for longer polymers where parts of the solute molecule have the capacity to fold in on themselves and overlap.

## 4 Partition coefficients

In this section, we test our expressions for the excess chemical potential by applying them in the prediction of partition coefficients and comparing with simulated values. This approach stems from the fact that we can relate the excess chemical potential (predicted *via* Widom insertion) with the resulting partition coefficient of a solute into two immiscible liquids, a relation which will be presented in this section.

We simulate a box containing two immiscible phases and insert a small amount of solute into the system. This allows us to vary the interaction parameters between the solute and the two solvent phases to extensively test eqn (21) and (25). For solute partitioning into two solvents  $\alpha$  and  $\beta$ , the partition coefficient is defined as

$$K = \frac{c^\alpha}{c^\beta} \quad (26)$$

where  $c^\alpha$  and  $c^\beta$  are the concentrations in the two phases. Therefore, using eqn (26), we can determine  $K$  from our simulation experiments.

At equilibrium, the chemical potentials of the solute in the two solvents are equal. By making use of this fact, as well as decomposing the chemical potential into its excess and ideal parts (eqn (13)), this allows us to express the partition coefficient  $K$  at infinite dilution in terms of the difference between the excess chemical potentials of the two phases<sup>39</sup>

$$K = \frac{c^\alpha}{c^\beta} = \exp\left(\frac{\mu_{\text{ex}}^\beta - \mu_{\text{ex}}^\alpha}{k_B T}\right). \quad (27)$$

In this section, we compare the partition coefficients of single bead solutes (Section 4.1), with those for bonded molecules (Section 4.2). Following this, we also investigate the effects of bonding the solvent molecules on  $K$  (Section 4.3). In all cases, we perform simulations consisting of two-phase separating solvents (type A and type B beads), defining the partition coefficient as  $K = c^A/c^B$ . To ensure their phase separation we set their interaction to be  $a_{AB} = 100$  which results



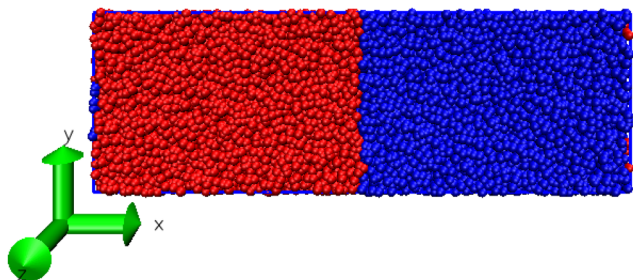


Fig. 3 Phase separation of beads of type A (red) and type B (blue) when their DPD interaction values are set as:  $a_{AB} = 100$ ,  $a_{AA} = 25$  and  $a_{BB} = 25$ .

in a negligible solubility of type A beads in type B and *vice versa*. We set all self-interactions to be  $a_{ii} = 25$ , unless specified otherwise. All simulations are conducted in boxes of size  $60 \times 20 \times 20$  with a bead density of  $\rho = 3R_C^{-3}$ . We refer to our solute beads as type C. The phase separation produced by types A and B (in the absence of solute beads) is illustrated in Fig. 3.

#### 4.1 Non-bonded solute

In order to first test the applicability of eqn (27), we perform a set of simulations in which DPD beads are unbonded, single beads. We examine the effects of varying the strength of the interaction with the solute beads (type C) and examine the partitioning for varying choices of  $a_{AC}$  and  $a_{BC}$ . We find that the value of  $K$  calculated is weakly dependent on the solute concentration, as shown in Fig. 4 (which is to be expected as we have assumed infinite dilution). Therefore, we find the value of  $K$  at infinite dilution by extrapolating to a concentration of 0, in order to directly compare with that predicted using eqn (27). The results of this study are shown in Table 1, showing good agreement between the theoretically predicted value and the simulated values. This also provides confidence in the calculated relationship between  $a_{ij}$  and the chemical potential which is shown in Fig. 1.

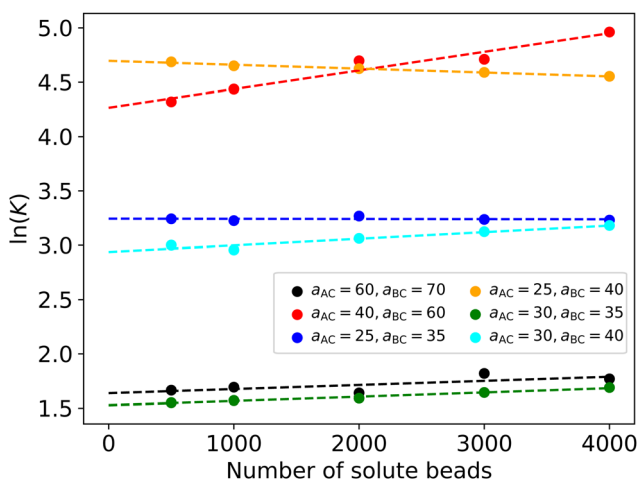


Fig. 4 The partitioning ( $K$ ) of unbonded solute beads (C type) in unbonded solvent beads (A and B type). There is a weak dependence of  $\ln K$  on the concentration of the solute.

Table 1 Calculation of the natural logarithm of partition coefficient  $K$  where  $K = c^A/c^B$ . For the simulated values, we also report the estimated uncertainty in our values in brackets as the standard deviation of the intercept

$a_{AC}$	$a_{BC}$	$\ln K$ (calculated using eqn (27))	$\ln K$ (simulated)
25	35	3.22	3.244 (0.016)
25	40	4.56	4.697 (0.007)
30	35	1.71	1.528 (0.008)
30	40	2.86	2.938 (0.026)
40	60	4.08	4.267 (0.058)
60	70	1.56	1.638 (0.052)

#### 4.2 Bonded solute

We next investigate the effects of bonding the solute molecules together, while keeping the solvent as single beads. We investigate for solute molecules with various lengths ( $N = 2, 3, 4$ ). The bond length  $l_0$  is also expected to play a role in solubility, as this significantly influences the degree of overlap between bonded beads. Therefore we test three bond lengths  $l_0 = 0.4R_C$ ,  $l_0 = 0.6R_C$  and  $l_0 = 1.0R_C$ . Similarly to in Section 4.1, we simulate various solute concentrations and extrapolate to zero solute concentration to determine a value for  $\ln K$ .

If we assume that the chemical potential for the molecules can be calculated using eqn (16) (*i.e.* the overlap is not taken into account), Fig. 5a shows the relationship between the simulated, and predicted values for  $\ln K$ . We observe that the deviation between the predicted and simulated values is largest when the bond length is shorter, and also increases as the value of  $K$  increases. This is hypothesised to be due to the increasing volume of overlap between the bonded beads at shorter bond lengths.

However, if we calculate the predicted value of  $\ln K$  using our volume-correction equations (eqn (21) and (25)) the relationship between the simulated and predicted values greatly improves, as shown in Fig. 5b. To quantify the degree of improvement, we calculate mean squared error values for the two different bond lengths before and after applying the correction, which are shown in Table 2. We note the significant improvement for the shorter bond lengths of  $l_0 = 0.4R_C$  and  $l_0 = 0.6R_C$ , while there is minimal change when  $l_0 = 1R_C$ . Note that all values shown in Fig. 5a and b can be found in the ESI.†

#### 4.3 Bonding the solvent

In this section, we investigate the effect of bonding the solvent. We perform simulations in which we determine  $\ln K$  for a single bead solute (type C), which partitions into two solvents where one of those solvents consists of bonded beads (type A) and the other solvent remains unbonded (type B). Similarly to before, we investigate the effects of the number of bonded beads in the molecule, as well as the bond length.

Fig. 6a shows the predicted  $\ln K$  vs. the simulated values for a variety of choices for  $a_{AB}$  and  $a_{AC}$ , where all self-interactions are kept as  $a_{ii} = 25$ . We see a small difference between the simulated and the predicted values. We hypothesise that, due to the bonding of the solvent, there is a difference in  $K$  which results from the change in density that bonding produces. The act of bonding molecules together decreases the pressure



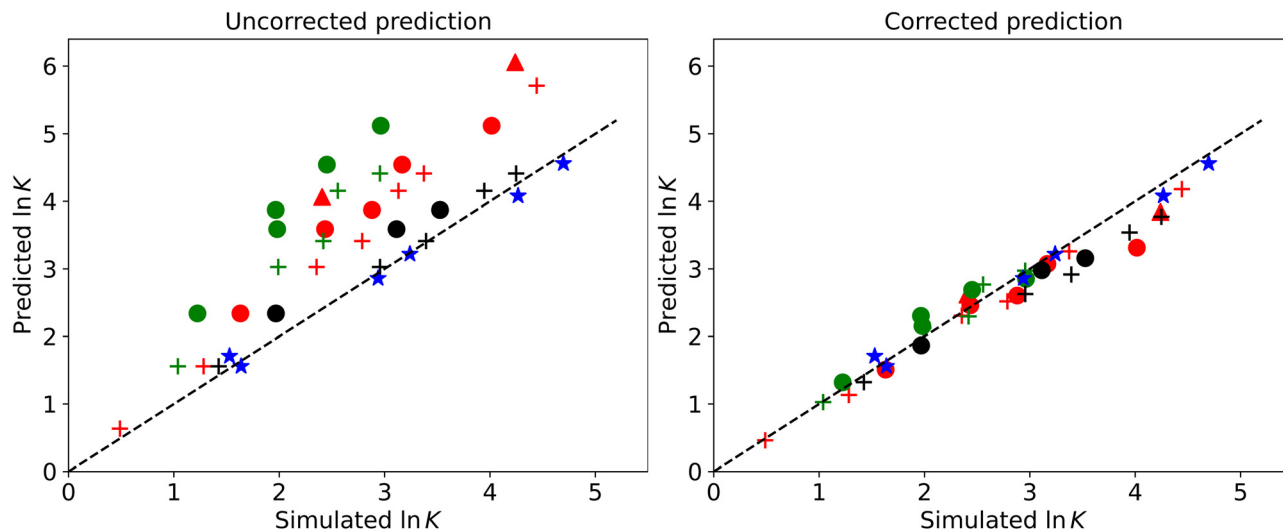


Fig. 5 Relationship between the predicted and simulated values of  $\ln K$ . The predicted values are calculated assuming that the excess chemical potential of bonded molecules can be calculated using (a) eqn (16) (*i.e.* uncorrected) and (b) eqn (21) (*i.e.* corrected). The symbol represents the number of bonded beads in the molecule: 2 (+), 3 (O) and 4 ( $\Delta$ ), while the colour represents the bond length  $l_0$ :  $0.4R_C$  (green),  $0.6R_C$  (red) and  $1.0R_C$  (black). The unbonded single-bead solute is represented by the blue (\*) points. Note that error bars are smaller than the symbol size, but values can be found in the ESI.†

Table 2 Mean square error (MSE) of  $\ln K$  between the predicted and simulated values when using the uncorrected form of the chemical potential vs. the corrected form. Note that bond length is in units of  $R_C$

$K$ prediction	Bond length	MSE
Uncorrected	0.4	2.35
Corrected	0.4	0.030
Uncorrected	0.6	1.19
Corrected	0.6	0.070
Uncorrected	1.0	0.072
Corrected	1.0	0.112

of the system, therefore increasing the equilibrium density, resulting in larger excess chemical potentials than those

determined in Section 3.1 (eqn (17)). This pressure decrease results from the fact that bonding reduces the number of interacting particles in the system.

We aim to compensate for this by altering the self-interaction value for the bonded solvent. The relationship between the number of bonded beads and the pressure is shown in Fig. 7. We calculate this pressure by performing simulations in a cubic domain of length  $L = 20R_C$  with a bead density of  $3R_C^{-3}$ . For a given bonded solvent, we choose to set the self-interaction  $a_{AA}$  to reproduce the same pressure as that of an unbonded solvent (for which we find a pressure of  $P = 23.7$ ). This is to ensure that in all cases the pure solvent produces a bead density of  $\approx 3R_C^{-3}$ , when simulated in coexistence with the unbonded solvent. We find

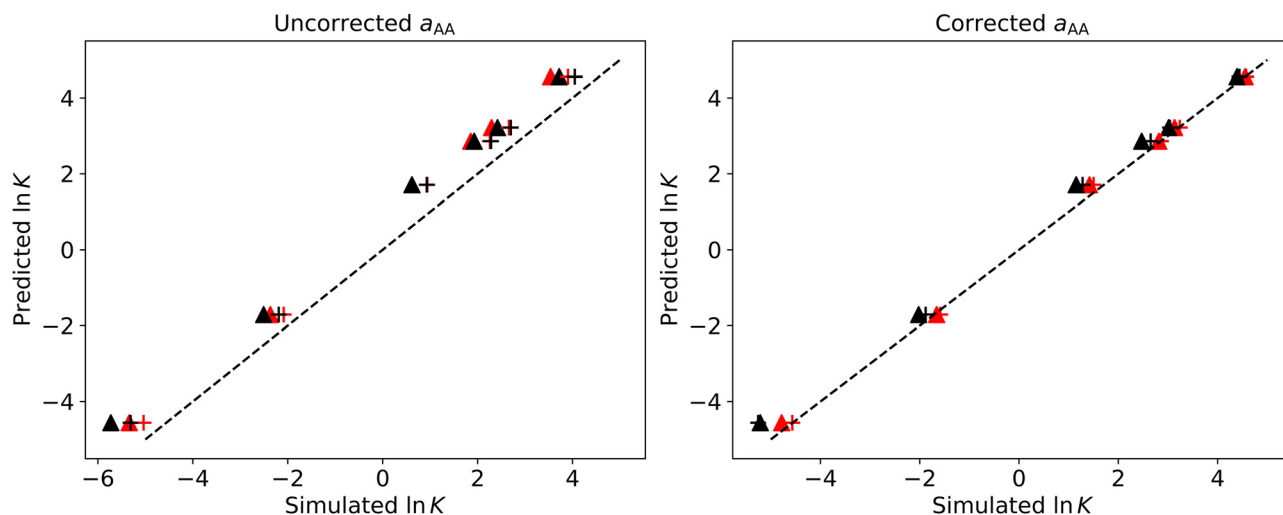


Fig. 6 Comparison between the predicted and simulated  $\ln K$  values for varying  $a_{AB}$  and  $a_{AC}$  choices. Fig (a) shows the partitioning when all self-interactions are set to  $a_{ij} = 25.0$ , while (b) shows the partitioning behaviour when the self-interaction of the bonded beads ( $a_{AA}$ ) is corrected based on the pressure. Note that error bars are smaller than the symbol size, but values can be found in the ESI.†



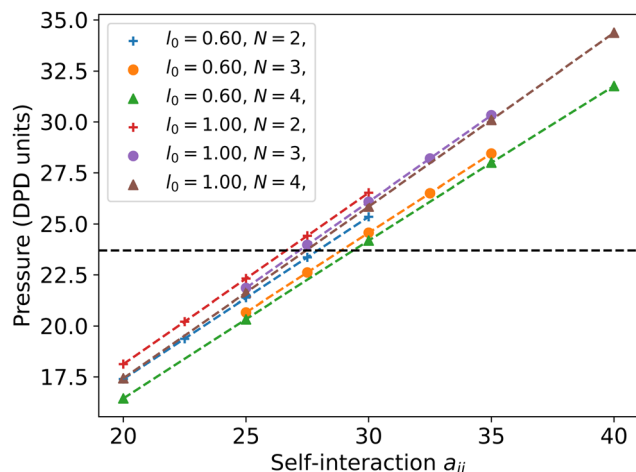


Fig. 7 Bulk pressure in simulations with number density 3, where the bond length  $l_0$  and the number of bonded beads  $N$  is varied. Also shown (black) is the bulk pressure of unbonded beads at the same number density.

that this approach works well for cases where a bond length of  $l_0 = 0.6R_C$  and  $l_0 = 1R_C$  are used, however, is not a perfect approach for correcting the density when  $l_0 = 0.4R_C$ . In this case, we suspect that other factors need to be taken into account to perfectly correct the density upon bonding. Therefore in this section, we focus on bond lengths  $l_0 = 0.6R_C$  and  $l_0 = 1R_C$ . The corrected  $a_{AA}$  values used for the simulations in this section are presented in Table 3. Fig. 8 provides an example of how the solvent density changes due to this updated self-interaction value.

We note here that eqn (17) was calculated by performing Widom insertions into a solvent of unbonded beads with  $a_{ii} = 25$ . We expect the relation between  $\mu_{ex}$  and  $a_{ij}$  to alter slightly depending on the solvent parameters *i.e.* if  $a_{ii} \neq 25$ , or if the solvent is bonded (or not). However, we find a negligible difference between the fit presented in eqn (17), and the fit for when the solvent is bonded and has the  $a_{ij}$  values we use in this article. Further information supporting this is presented in the ESL.†

Fig. 6b shows the new predicted vs simulated relationship for  $\ln K$ . We see that the updated calculated values more closely match the simulated results once the self-interaction for the bonded solvent is updated. Similarly to previously, we calculate the mean squared error (MSE) for the uncorrected and corrected cases, which is shown in Table 4, where we observe that the MSE increases with more bonds in the solvent and that the correction greatly decreases the MSE between the simulated and predicted values of  $\ln K$ . Note that all values shown in Fig. 6 can be found in the ESL.†

Table 3 Solvent self-interaction values  $a_{AA}$  which reproduce the same pressure as non-bonded beads ( $P = 23.7$ ). Bond length is given in units of  $R_C$

$N$	Bond length	$a_{AA}$
2	0.6	27.9
3	0.6	28.9
4	0.6	29.4
2	1.0	26.6
3	1.0	27.2
4	1.0	27.5

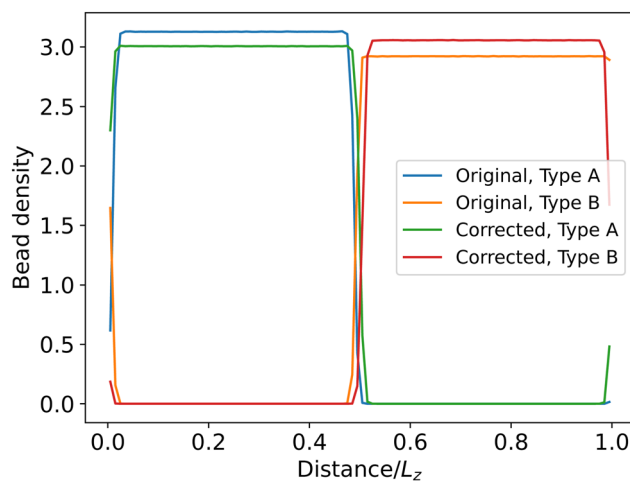


Fig. 8 Example of the improvement of the bead density when the self-interaction for bonded beads is altered. Example for when four A-type molecules are bonded together with equilibrium bond length  $l_0 = 0.6R_C$ . Type-B beads are left unbonded and  $a_{AB} = 100$ . In the original case  $a_{AA} = a_{BB} = 25$ , while in the corrected case  $a_{AA} = 29.4$  and  $a_{BB} = 25.0$ .

It may seem counter-intuitive that the effects of bonding the solute molecules appear more significant than the effects of bonding the solvent molecules. However, this conclusion is expected from a theoretical point of view. In the case of the solvent, the effect of bonding beads does not significantly alter the degree of overlap, as in a bulk solvent of single beads there is expected to be a degree of overlap anyway. This contrasts with what is discussed in Section 4.2, for solutes, where the correction applied in calculating  $\mu_{ex}^M$  accounts for a difference between beads which have no overlap at all, and those which overlap because they are bonded.

This point can be illustrated by considering the effect that bonding has on the radial distribution function of the fluid. We simulate a selection of pure systems in cubic simulation boxes, for which the radial distribution function  $g(r)$  is shown in Fig. 9. The simulations are conducted in boxes with edge length  $L = 20R_C$  and density  $\rho = 3R_C^{-3}$ . Therefore, the only differences between the three cases shown is in the choice of bonding, and the repulsion parameter  $a_{ii}$ . As previously, we choose to set this so that the pressure is the same in all simulations (the values used in Fig. 9 can be found in Table 3).

For a system consisting of single beads, if our beads were non-interacting, the average distance  $d$  between beads could be

Table 4 Comparison of the mean squared error (MSE) for the simulated value of  $\ln K$  (compared with the calculated value using eqn (25)) when the self-interaction of the solvent is corrected to account for the density change of the solvent upon bonding

$N$ (solvent)	Bond length	MSE (uncorrected)	MSE (corrected)
2	0.6	0.344	0.010
2	1.0	0.378	0.131
4	0.6	0.857	0.024
4	1.0	0.900	0.176



calculated as

$$d = \left(\frac{1}{\rho}\right)^{\frac{1}{3}} \quad (28)$$

All of our simulations are conducted with a bead density of  $\rho = 3R_C^{-3}$ , which would lead to an average distance  $d = 0.7R_C$ . However, as our beads are interacting (primarily *via* the repulsive conservative force) the average distance between unbonded beads is larger, as shown in Fig. 9.

The process of bonding leads to the appearance of peaks at approximately integer values of the equilibrium bond lengths. However, bonding does not significantly alter the average overlap volume in the systems. Upon bonding, there is effectively a narrowing of the distribution of bead distances. The system loses some of its closest beads (reflected by only a significant decrease in  $g(r)$  for  $r < 0.5R_C$ ) but also some that would otherwise have been slightly further apart ( $g(r)$  decrease in the range  $0.7R_C < r < 0.9R_C$ ).

Finally, we also test systems with both bonded solute and bonded solvent, to test combining the bonding effects discussed thus far. Exact details of this study can be found in the ESI,<sup>†</sup> however, we find that we can combine the corrections discussed thus far to make good predictions for the expected chemical potential, and hence partitioning.

#### 4.4 Solute–solute interactions

Until now, the examples studied considered solute molecules consisting of a single type of bead. However, a solute molecule could potentially be made up of various bead types. In this section, we briefly discuss the bonding overlap effects which would need to be considered if one wishes to define a full set of interactions for more complex systems.

If we consider an example system where the solute molecule is made up of three beads and a fourth solvent bead (Fig. 10), there are 6 cross interaction  $a_{ij}$  values to determine. Three of

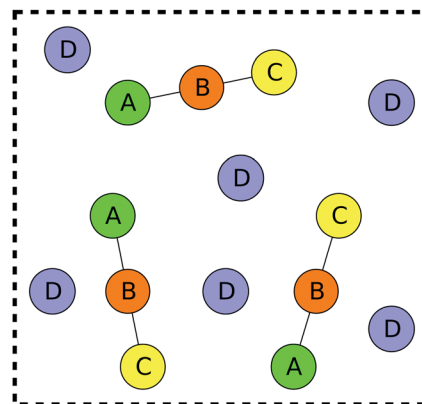


Fig. 10 Simple example system of a solute consisting of three bead types (A)–(C) in a solvent (D).

those interactions are related to how the solute interacts with the solvent ( $a_{AD}$ ,  $a_{BD}$  and  $a_{CD}$ ). In this case, the values are best determined by aiming to reproduce the correct activity coefficients of each of the solute beads at infinite dilution in the solvent (assuming that the system to be simulated consists of significantly more solvent molecules than solute). There are, however, three other cross-interaction values that are between the solute molecules themselves ( $a_{AB}$ ,  $a_{AC}$  and  $a_{BC}$ ). In this case, one could choose to determine the cross-interaction in two different ways depending on which bead is chosen for the Widom insertion. For example,  $a_{AB}$  could be determined by matching to the IDAC of bead A in bead B, or *vice versa*.

Upon bonding of the solute beads, bead overlaps lead to an effective change in the IDAC coefficient, as discussed previously. However, unlike in Section 4.2, the effective interaction (as a result of bonding) has changed for both beads involved in the interaction. Therefore, we now consider the effect that bonding has on the solute–solute interactions.

Fig. 11 summarises the bonding effects addressed in Section 4.2 (interactions between a bonded solute and the solvent) and Section 4.3 (interactions between a solute and a bonded solvent). In this figure, we also illustrate the effects addressed in this section between the bonded solute molecules.

If, for example, we consider the interaction between bead A (left-hand molecule) and bead B (right-hand molecule) in Fig. 11c, this could be determined *via* matching to the IDAC of bead A in a bath of bead B. In this case, then the effect of bonding on the first molecule can partially be taken into account in the same way as described in Section 4.3, by altering the interaction to account for its overlap with its bonded neighbour. However, since bead B in the second molecule is also overlapping with its neighbours there is an extra impact of bonding.

However, the effects of bonding on the second molecule cannot necessarily be accounted for in the same manner as the first. As the interaction was determined *via* the insertion of a single bead into a bath of beads, there was already a degree of overlap of the B-type beads in the bath. Now that the molecule has been bonded to its neighbours, the degree of overlap has changed very little, the impact being the type of bead it is

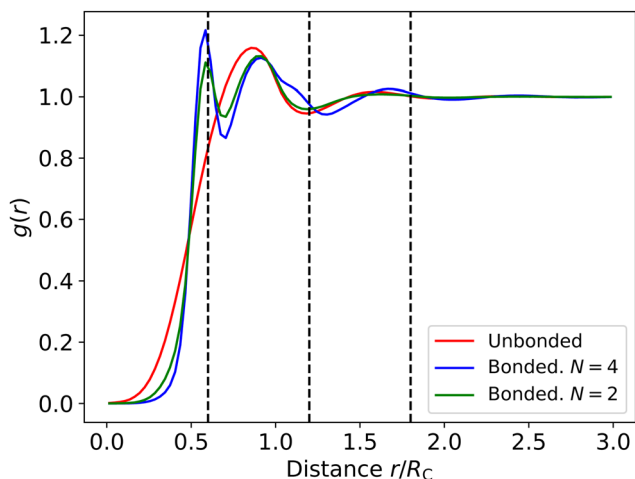


Fig. 9 Radial distribution function  $g(r)$  for a variety of simple test cases. In the cases of bonded simulations, we use an equilibrium bond length of  $l_0 = 0.6R_C$  for molecules of two different lengths  $N$ . Also illustrated vertically are integer values of the equilibrium bond length.



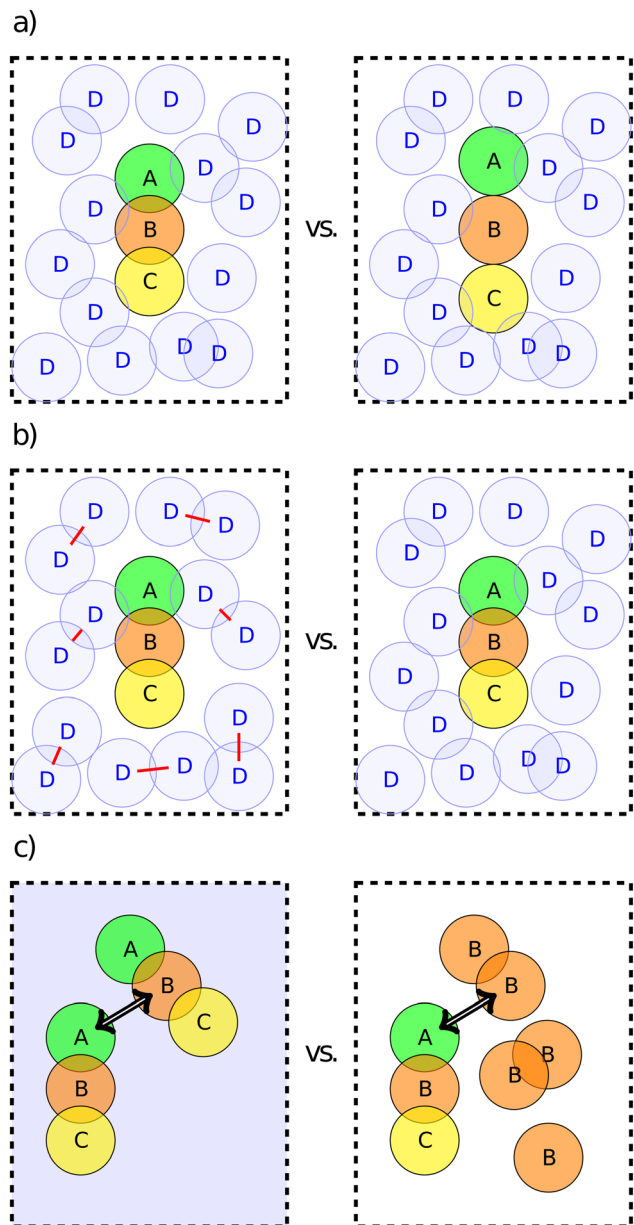


Fig. 11 Effect of bonding (a) the solute on its interaction with the solvent (Section 4.2); (b) the solvent on its interaction with the solute (Section 4.3); (c) the solute on its interaction with other solute molecules (Section 4.4).

overlapping with. This is similar to the effects of bonding the solvent discussed in the previous section.

For example, if we refer again to our example case, we can determine the effect the overlap of the second solute molecule will have on the excess chemical potential  $\mu_{\text{ex}}$ . If we assume that the average volume of overlap  $V_{\text{OL}}$  between a pair of beads is unchanged before and after bonding, then

$$\Delta\mu_{\text{ex}} \approx \frac{1}{2} \frac{V_{\text{OL}}}{V_{\text{Bead}}} \left( \frac{\mu_{\text{ex}}(a_{\text{AB}} + a_{\text{AA}}) + \mu_{\text{ex}}(a_{\text{AC}} + a_{\text{AB}})}{\mu_{\text{ex}}(2a_{\text{AB}})} \right) \quad (29)$$

where  $V_{\text{Bead}}$  is the volume of a single bead. Of course, this is a simplification, and the choice of bond length will influence if there is any change in the volume of overlap. However, we find

that  $\Delta\mu_{\text{ex}}$  is generally going to take a relatively small value. For example, if we take some values illustrative of a realistic case, in which the bond length is  $l_0 = 0.6R_{\text{C}}$ ,  $a_{\text{AA}} = 25$ ,  $a_{\text{AC}} = 25$ , and  $a_{\text{AB}} = 50$ , we find that the Widom insertion approach would find a value of  $\mu_{\text{ex}} = 19.0$  for the insertion of bead A in bead B. Eqn (29) predicts that the bonding of bead B would change this by  $\Delta\mu_{\text{ex}} \approx 0.9 \frac{V_{\text{OL}}}{V_{\text{Bead}}}$ , which is relatively small.

However, we also note that for the cross-interaction values that are between the solute molecules, it may be best to determine the  $a_{ij}$  value for each pair as an average of the two IDAC values from one bead in another. For example, one will determine a different value calculating  $a_{\text{AB}}$  depending on if one matches to the IDAC of bead A in bead B, or visa versa. Therefore, a potential choice is to determine both and take the average value *i.e.*  $a_{\text{AB}} = \frac{1}{2}(a_{\text{A} \rightarrow \text{B}} + a_{\text{B} \rightarrow \text{A}})$ . This is likely to be an effective approach if the two approaches yield similar interaction values (that is  $a_{\text{A} \rightarrow \text{B}}$  and  $a_{\text{B} \rightarrow \text{A}}$  are fairly similar).

## 5 Altering intramolecular interactions

As briefly mentioned in Section 2.1, by default LAMMPS does not allow beads which are directly bonded to interact. However, this behaviour can be switched on using the ‘special bonds’ function, and here we investigate the impact that switching these interactions on has on our results. This is motivated by the fact that many popular DPD codes, such as DL\_MESO,<sup>40</sup> choose to allow this interaction.

We repeat a small subset of our calculations from Section 4.2 (a bonded solute in an unbonded solvent), but allowing the bonded beads to also interact with each other *via* the usual DPD repulsive force. The first noticeable difference between the two cases is the average bond length. In our original simulations, although the equilibrium bond length is set as  $l_0 = 0.6R_{\text{C}}$ , the average resulting bond length in the simulation is  $0.58R_{\text{C}}$ . This decrease, compared to the  $l_0$ , is due to the pressure of the system. However, if desired, we could force the bond length to be closer to the value set in eqn (7) by increasing the strength of the bonding potential parameter  $K_{\text{B}}$ . In our repeated simulations with bonded interactions turned on, we find that the average bond length in the system increases to  $0.61R_{\text{C}}$ . This increase is due to the added interparticle repulsion between two bonded beads. Details of these calculations of apparent bond length can be found in the ESI.†

A comparison of the partitioning  $\ln K$  for the two cases is shown in Fig. 12, where the simulation data is summarised in the ESI.† Due to the slight change in bond length, one might expect to see a difference in the resulting partitioning due to the change in the degree of overlap. However, despite the slight change in average bond length, there is no significant change in partitioning within the error of the simulated value of  $\ln K$ . If we consider eqn (19) we can determine that the change in the overlap between two beads would only vary by  $\approx 4\%$  upon the bond length changing from  $0.58R_{\text{C}}$  to  $0.61R_{\text{C}}$ , which would not result in a significant change in the chemical potential, likely



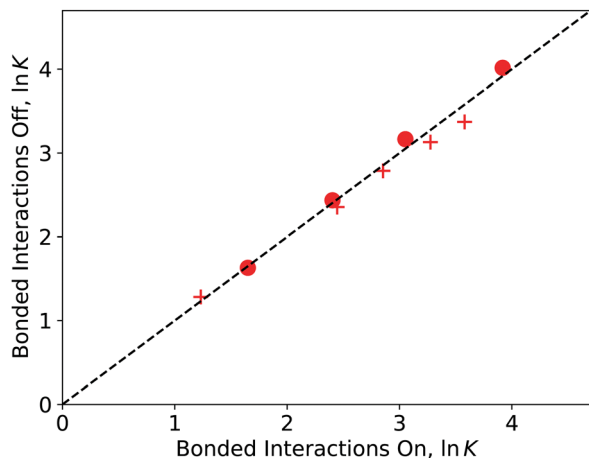


Fig. 12 A comparison of the partitioning  $\ln K$  which results when the bonded interactions are turned on and turned off. The symbol represents the number of beads in the bonded solvent: 2 (+) and 3 (O). The equilibrium bond length in all cases is set as  $l_0 = 0.6$ . Note that error bars are smaller than the symbol size, but values can be found in the ESI.†

indicating why we see very little variation between the two simulation cases in Fig. 12.

We also now consider the effect that including bonded interactions has on a bonded solvent, repeating the pressure calculation originally shown in Fig. 7. Due to the increase in the number of interactions in the fluid, turning the bonded interactions on alters the relationship between the pressure and the self-interaction value. The new  $a_{ii}$  value required to produce the same pressure as a non-bonded solvent is shown in Table 5 (full figure included in the ESI†). Comparing Table 3 and, we see that altering the nature of bonded bead interactions alters the pressure when the bond length is shorter, but not when it is longer. This is because the longer bond length of  $l_0 = 1R_C$  is the same as the cut-off,  $R_C$ , and so beads have negligible interaction with each other even when bonded interactions are switched on. However, the difference in pressure for shorter bond lengths can be accounted for when setting the self-interactions (as discussed above). Therefore, we conclude that the results presented in this paper are applicable to systems with bonded beads with excluded bead interactions turned on or off.

## 6 Application to real systems

In this section, we briefly demonstrate the applicability of our model to real systems by studying the partitioning of three

Table 5 Solvent self-interaction values  $a_{AA}$  which reproduce the same pressure as non-bonded beads ( $P = 23.7$ ), when the fluid has bonded beads are allowed to interact with each other

$N$	Bond length	$a_{AA}$
2	0.6	27.5
3	0.6	28.3
4	0.6	28.6
2	1.0	26.6
3	1.0	27.2
4	1.0	27.5

different molecules into a water/dodecane system. The coarse-graining of the solutes, along with the representation for dodecane, are shown in Fig. 13. The coarse-grained modelling for these three systems is very different, highlighting the versatility of the method for real molecules.

For each solute we calculate the interaction parameters  $a_{ij}$  using our equations to correct  $\mu_{\text{ex}}$  due to their overlap. We then determine the resulting partition coefficient  $K$ , and compare with experimental data. In each case, we also calculate another set of interaction parameters  $a_{ij}$  for if we were to ignore the overlap of the beads, to show the improvement that our method has over this approach.

### 6.1 Dodecane and water solvents

We use a single water bead to represent multiple water molecules, and dodecane is modelled by four bonded beads of the same type. In our simulations, we use a constant bead density of  $\rho = 3R_C^{-3}$ . Therefore we can find a value  $R_C$  (and a means for conversation into real units) by matching to the experimentally known density of dodecane at room temperature. This approach leads us to determine  $R_C = 6.56 \text{ \AA}$ . From this, we define the coarse-graining of water  $N_W$  by also matching to the density of water at room temperature, finding that 1 water bead represents  $N_W \approx 3.14$  water molecules.

To determine an appropriate bond length for the dodecane molecule, we use the fact the C–C–C angles are approximately  $109.5^\circ$ , and that the experimental C–C bond is  $1.543 \text{ \AA}$ .<sup>41</sup> The separation between three carbon atoms is therefore approximately  $3.77 \text{ \AA}$ , meaning an equilibrium bond length of  $0.6R_C$  is a good choice.

### 6.2 Solutes

As shown in Fig. 13, diethyl carbonate can be represented by two beads of the same type, due to its symmetry. This means that eqn (19), (20), and (21) can be directly applied to account for the overlap. For diethyl carbonate, we use an equilibrium

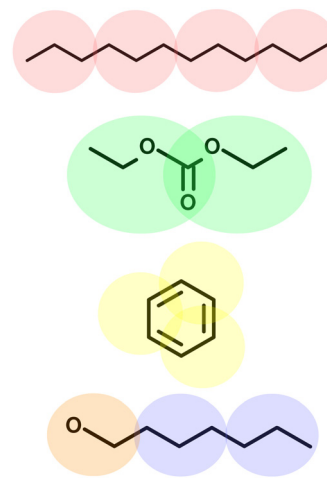


Fig. 13 Coarse-graining used for various real molecules simulated in this section. Note that the beads are for illustrative purposes only and do not indicate the shape or size of the beads nor the degree of the overlap.



bond length of  $0.6R_C$ , which converts to a realistic bond length for this molecule.

For simulating heptanol, we need to use two different types of beads. This means that we need to calculate the volume of the overlapping portions separately, rather than using eqn (19) and (20). Then we correct each bead's  $\mu_{\text{ex}}$  individually, due to the overlapping portions. However, practically this is simple and not much more complicated than our previous equations (which assume each bonded bead interacts with the solvent with the same  $a_{ij}$  parameter). Once again, we use an equilibrium bond length of  $0.6R_C$ , since this is representative of the real bond length.

Finally, we model the benzene as a triangle of bonded beads, where an angle constraint is imposed at every corner, and the equilibrium angle is set to  $60^\circ$ . Using that the average C–C bond in a benzene molecule is  $1.39 \text{ \AA}$ , we determine the separation between bonded beads should be  $2.41 \text{ \AA}$ , meaning an equilibrium bond length of  $0.4R_C$  is a good choice.

The choice to model benzene using three beads in a triangle formation has been used in other DPD models<sup>24</sup> and makes sense from a geometric point of view. However, this formation of beads has a very large degree of overlap and so we expect that in most previous author's models, the chemical potential of rings is not fully reflected. Using our method to calculate  $\mu_{\text{ex}}$  requires us to know the degree of overlap each bead has with one another. Our previous expressions for volume are based on a linear construction of the molecule, therefore we need another approach. The analytical expression for the volumes of overlap in a ring is quite complex compared to the cases used in the article so far. However, the volume of overlap can very easily be calculated using a Monte Carlo simulation. Therefore eqn (25) can still be applied to calculate the excess chemical potential, once these volumes have been found. Details of this calculation, including the volumes  $V_1$ ,  $V_2$  and  $V_3$  for different bond lengths, can be found in the ESI.†

The target values for excess chemical potential  $\mu$  are obtained from experimental data. The exact  $a_{ij}$  values used in each simulation case can be found in the ESI,† along with details of the experimental data used and additional information about these simulations.

### 6.3 Partitioning results

Table 6 shows the results of our partitioning study for each molecule. We show the values of  $\ln K$  obtained when the overlap of beads is not taken into account (uncorrected) when calculating the  $a_{ij}$  interaction parameters for the molecule, compared to when we use our approach to account for the overlap (corrected) when determining suitable  $a_{ij}$  values.

We observe a significant improvement in our corrected parameters compared to the uncorrected ones, and we see this for all solute cases tested. This is particularly notable for our benzene representation. For all solutes, the use of uncorrected parameters leads the solute molecules to behave with significantly less hydrophobicity than they should. However, correcting the parameters for the overlap generates a much more realistic behaviour. Generating the correct behaviour, and in particular

**Table 6** The simulated values for  $\ln K$  in the cases where the volume of overlap is not accounted for (uncorrected) and it is accounted for using our method (corrected). Experimental  $\ln K$  are obtained from ref. 42, where  $K$  is defined as the ratio of the concentration in the dodecane to the water

Solute	Simulated $\ln K$ (uncorrected)	Simulated $\ln K$ (corrected)	Experimental $\ln K$
Diethyl carbonate	0.817	1.168	1.3
Heptanol	0.484	1.965	1.9
Benzene	2.542	4.475	5.1

the correct hydrophobicity, is crucial for many types of systems, but in particular for self-assembling systems such as surfactants.

## 7. Conclusions and discussion

In this article, we have shown that there are effects resulting from the bonding of beads, which are generally not considered in the parameterisation of DPD models. However, we show that these bonding effects can be compensated for by considering the degree of overlap between bonded beads. The degree of overlap depends on the chosen bond length, and shorter bond lengths lead to the need to consider the effects of bonding more carefully. Although, as highlighted in Section 5, the average bond length is slightly different to its harmonic potential setting (*i.e.* that set by eqn (7) using parameter  $l_0$ ), we show that this deviation is small enough that we can use  $l_0$  to estimate the volume of overlap of the beads, and hence correct the chemical potential.

We have also considered more complex cases, where molecules can be made up of a combination of fragment types. While we have only considered linearly bonded molecules in this work with a single bond length, the methodology presented could easily be extended to consider more complex molecules, including those containing rings or branching *etc.*

Finally, it is worth considering the general applicability of the results presented in this work. In recent years, DPD has become a very versatile tool for the simulation of soft matter systems. Accordingly, we expect the method presented here to find uses in the simulation of surfactants, where accurate parametrization is required to obtain predictions of interfacial tensions between fluids<sup>20,43</sup> and for predicting critical micelle concentrations;<sup>21,25</sup> for model membranes where partitioning of small molecules through the membrane is strongly dependent on the accurate parametrization of the interactions with lipid molecules;<sup>44</sup> for DPD studies of peptides and proteins<sup>45</sup> where folding depends critically on accurate monomer–monomer  $a_{ij}$  values; in addition to a wide variety of other problems.

## Author contributions

RH: conceptualization, data curation, formal analysis, investigation, methodology, software, validation, visualization, writing – original draft. CA: project administration, writing – review and editing. MW: funding acquisition, project administration, supervision, writing – review and editing.



## Data availability

The data supporting this article have been included as part of the ESI.†

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

The authors would like to acknowledge the support of Durham University for the use of its HPC facility, Hamilton. RH would like to acknowledge support through an EPSRC/P&G funded Prosperity Partnership grant EP/V056891/1.

## Notes and references

- P. J. Hoogerbrugge and J. M. V. A. Koelman, *Europhys. Lett.*, 1992, **19**, 155–160.
- R. D. Groot and P. B. Warren, *J. Chem. Phys.*, 1997, **107**, 4423–4435.
- P. Español and P. Warren, *Europhys. Lett.*, 1995, **30**, 191–196.
- R. D. Groot, *J. Chem. Phys.*, 2003, **118**, 11265–11277.
- C. Marsh, G. Backx and M. Ernst, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1997, **56**, 1676–1691.
- T. L. Rodgers, O. Mihailova and F. R. Siperstein, *J. Phys. Chem. B*, 2011, **115**, 10218–10227.
- Z. Li and E. E. Dormidontova, *Macromolecules*, 2010, **43**, 3521–3531.
- M. Lísal and J. K. Brennan, *Langmuir*, 2007, **23**, 4809–4818.
- A. G. Schlijper, P. J. Hoogerbrugge and C. W. Manke, *J. Rheol.*, 1995, **39**, 567–579.
- A. A. Gavrillov, Y. V. Kudryavtsev and A. V. Chertovich, *J. Chem. Phys.*, 2013, **139**, 224901.
- K. Zhang and C. W. Manke, *Mol. Simul.*, 2000, **25**, 157–166.
- S. Chen, N. Phan-Thien, X.-J. Fan and B. C. Khoo, *J. Non-Newtonian Fluid Mech.*, 2004, **118**, 65–81.
- S. Wang, S. Yang, R. Wang, R. Tian, X. Zhang, Q. Sun and L. Liu, *J. Pet. Sci. Eng.*, 2018, **169**, 81–95.
- E. Mayoral and A. G. Goicochea, *J. Chem. Phys.*, 2013, **138**, 094703.
- Y.-D. He, Y.-H. Tang and X.-L. Wang, *J. Membr. Sci.*, 2011, **368**, 78–85.
- R. Groot and K. Rabone, *Biophys. J.*, 2001, **81**, 725–736.
- S. Sengupta and A. Lyulin, *Polymers*, 2020, **12**, 907.
- T. Ni, G.-S. Huang, P. Gao, Y.-T. Xu and M.-Z. Yang, *Polym. J.*, 2011, **43**, 635–641.
- A. Vishnyakov, M.-T. Lee and A. V. Neimark, *J. Phys. Chem. Lett.*, 2013, **4**, 797–802.
- X. Wang, K. P. Santo and A. V. Neimark, *Langmuir*, 2020, **36**, 14686–14698.
- M.-T. Lee, R. Mao, A. Vishnyakov and A. V. Neimark, *J. Phys. Chem. B*, 2016, **120**, 4980–4991.
- R. L. Hendrikse, C. Amador and M. R. Wilson, *Soft Matter*, 2023, **19**, 3590–3604.
- R. L. Hendrikse, C. Amador and M. R. Wilson, *Soft Matter*, 2024, **20**, 6044–6058.
- R. L. Anderson, D. J. Bray, A. S. Ferrante, M. G. Noro, I. P. Stott and P. B. Warren, *J. Chem. Phys.*, 2017, **147**, 094503.
- R. L. Anderson, D. S. D. Gunn, T. Taddese, E. Lavagnini, P. B. Warren and D. J. Bray, *J. Phys. Chem. B*, 2023, **127**, 1674–1687.
- R. L. Anderson, D. J. Bray, A. Del Regno, M. A. Seaton, A. S. Ferrante and P. B. Warren, *J. Chem. Theory Comput.*, 2018, **14**, 2633–2643.
- M. Panoukidou, C. R. Wand, A. Del Regno, R. L. Anderson and P. Carbone, *J. Colloid Interface Sci.*, 2019, **557**, 34–44.
- R. L. Hendrikse, A. E. Bayly and P. K. Jimack, *J. Phys. Chem. B*, 2022, **126**, 8058–8071.
- D. J. Bray, R. L. Anderson, P. B. Warren and K. Lewtas, *J. Chem. Theory Comput.*, 2020, **16**, 7109–7122.
- D. J. Bray, R. L. Anderson, P. B. Warren and K. Lewtas, *J. Phys. Chem. B*, 2022, **126**, 5351–5361.
- J. Saathoff, *J. Chem. Phys.*, 2018, **148**, 154102.
- H. Rezaei and H. Modarress, *Chem. Phys. Lett.*, 2015, **620**, 114–122.
- K. P. Travis, M. Bankhead, K. Good and S. L. Owens, *J. Chem. Phys.*, 2007, **127**, 014109.
- E. Mayoral and E. Nahmad-Achar, *J. Chem. Phys.*, 2012, **137**, 194701.
- T. P. Liyana-Arachchi, S. N. Jamadagni, D. Eike, P. H. Koenig and J. I. Siepmann, *J. Chem. Phys.*, 2015, **142**, 044902.
- C. M. Wijmans, B. Smit and R. D. Groot, *J. Chem. Phys.*, 2001, **114**, 7644–7654.
- M. R.-B. A. Gama Goicochea and R. López-Rendón, *Mol. Phys.*, 2007, **105**, 2375–2381.
- M.-T. Lee, A. Vishnyakov and A. V. Neimark, *J. Phys. Chem. B*, 2013, **117**, 10304–10310.
- A. S. Paluch, S. Parameswaran, S. Liu, A. Kolavennu and D. L. Mobley, *J. Chem. Phys.*, 2015, **142**, 044508.
- S. M. Michael, A. Seaton, R. L. Anderson and W. Smith, *Mol. Simul.*, 2013, **39**, 796–821.
- F. H. Allen, O. Kennard, D. G. Watson, L. Brammer, A. G. Orpen and R. Taylor, *J. Chem. Soc., Perkin Trans. 2*, 1987, S1–S19.
- M. H. Abraham and W. E. Acree Jr., *New J. Chem.*, 2004, **28**, 1538–1543.
- H. Alasiri and W. G. Chapman, *J. Mol. Liq.*, 2017, **246**, 131–139.
- H.-M. Ding, W.-D. Tian and Y.-Q. Ma, *ACS Nano*, 2012, **6**, 1230–1238.
- A. Vishnyakov, D. S. Talaga and A. V. Neimark, *J. Phys. Chem. Lett.*, 2012, **3**, 3081–3087.

