

This article was downloaded by:

On: 31 January 2011

Access details: *Access Details: Free Access*

Publisher *Taylor & Francis*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Molecular Simulation

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713644482>

DL_POLY_2 adaptations for solvation studies

P. -A. Cazade^a; P. Bordat^a; I. Baraille^a; R. Brown^a; W. Smith^b; I. T. Todorov^b

^a Institut pluridisciplinaire de recherche sur l'environnement et les matériaux, UMR 5254 du C.N.R.S. et de l'Université de Pau et des pays de l'Adour, Pau Cedex, France ^b Computational Science and Engineering Department, S.T.F.C. Daresbury Laboratory, Cheshire, UK

Online publication date: 28 January 2011

To cite this Article Cazade, P. -A. , Bordat, P. , Baraille, I. , Brown, R. , Smith, W. and Todorov, I. T.(2011) 'DL_POLY_2 adaptations for solvation studies', *Molecular Simulation*, 37: 1, 43 — 52

To link to this Article: DOI: 10.1080/08927022.2010.517531

URL: <http://dx.doi.org/10.1080/08927022.2010.517531>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

DL_POLY_2 adaptations for solvation studies

P.-A. Cazade^a, P. Bordat^a, I. Baraille^a, R. Brown^a, W. Smith^{b*} and I.T. Todorov^b

^a*Institut pluridisciplinaire de recherche sur l'environnement et les matériaux, UMR 5254 du C.N.R.S. et de l'Université de Pau et des pays de l'Adour, Avenue Pierre Angot, 64053 Pau Cedex, France;* ^b*Computational Science and Engineering Department, S.T.F.C. Daresbury Laboratory, Daresbury, Warrington, WA4 4AD Cheshire, UK*

(Received 14 July 2010; final version received 18 August 2010)

Adaptations of the molecular dynamics engine DL_POLY_2 were recently made to facilitate the investigation of questions arising mainly in the study of solvation dynamics. This paper presents the new facilities, with consideration of practical points for their efficient use and illustrative results obtained with them. Supplementary Information provides further details of their implementation to encourage future improvements and extensions.

The adaptations of DL_POLY_2 for solvation studies are:

- *Species by species and inter-species breakdown of the total energy* for systems containing several species, such as solutions or models of adsorption, with, as an illustration, diffusion of a particular water molecule in a microporous silica zeolite, silicalite. The new facilities allow correlation of the trajectory of the molecule with its interaction with the host matrix.
- *Solvation-induced spectral shifts*, illustrated by the absorption shift and Stokes shift of a cyanoaromatic dye in solution in methanol. We discuss the relative importance of different contributions to the total spectral shift.
- *Free energies by thermodynamic integration*, illustrated by calculation of the excess free energy of liquid dimethylsulphoxide. We examine the use of different weighting functions for the mixed Hamiltonian to avoid or to delay divergence in thermodynamic integration.

Keywords: molecular dynamics; solvation; spectroscopy; free energy

1. Introduction

DL_POLY is a general purpose, parallel molecular simulation package developed by Daresbury Laboratory in the UK to support academic research. Since its release in 1996, it has become widely used by research groups throughout the world and has been successfully exploited in numerous and varied studies of molecular systems which testify to its versatility [1,2]. The software and documentation is available free of cost under an academic licence from the website www.ccp5.ac.uk/DL_POLY/.

It is fair to say, however, that while DL_POLY provides a versatile and powerful molecular dynamics 'engine' capable of performing a wide range of simulations, it offers less in the provision of system analysis tools. The introduction of new features that extend the capabilities of the package in this regard is therefore of interest to many researchers. These features are already implemented, at least partially, in other molecular dynamics packages mainly devoted to biosimulations like Gromacs [3], Namd [4] or Charmm [5]. In this paper, we describe extensions of the DL_POLY_2 code to facilitate studies of solubility, which include the calculation of the energy decomposition

in terms of the different interacting species, solvent-induced spectroscopic shifts and the free energy of solvation.

2. Methods

The modifications described in this paper have been implemented in the DL_POLY_2 program and released as version 2.20. DL_POLY_2 is a modular code written primarily in Fortran 77 with components of Fortran 90. Extensive modifications of the existing code were required, most of which were confined to the subroutines that calculate energy terms and forces. An additional module was also created, named *solvation_module*, which hosts the solvation-specific arrays (and some variables) and gathers additional utilities (see below). This section gives an overview of the modifications. See the online Supplementary Information (SI) for more details of their implementation.

The energy/force subroutines affected were those which calculate the intramolecular terms: (bonds, valence angles, dihedrals and inversions); non-bonding pair

*Corresponding author. Email: w.smith@dl.ac.uk

interactions (van der Waals); coulombic (direct sum and Ewald methods); and three- and four-body non-bonding angular terms. Some subroutines were not modified due to the difficulty of isolating species-specific energy terms. These include the smoothed particle-mesh Ewald method [6] and the density-dependent interactions based on Tersoff's method [7] and the Finnis–Sinclair [8] and embedded atom model [9] for metallic structures.

Classical molecular dynamics simulations assume a force field or empirical potential energy function of the atomic (nuclear) coordinates, parameterised to represent a particular electronic energy surface, usually the electronic ground state. Calculation of solvent-induced spectral shifts required the introduction of a new feature: the 'ghost' molecule. The ghost molecule here represents part of the system, typically a chromophore in an excited state, for which different force field parameters apply. The ghost is located at the exact same position as the ground state chromophore, but it is not fully coupled to the other parts of the system. It functions as a probe of the interactions in the other state, e.g. those with the solvent environment surrounding the chromophore in its excited state. Formally, the ghost does not possess any dynamics of its own and exerts no force. It is simply mapped on to the location of the ground state structure, from where its interaction with the solvent is computed at regular time intervals. The ground state chromophore may be part of a larger molecule and can be defined, with the ghost molecule, as a contiguous part of it. For spectroscopic purposes, the remainder of the molecule is then treated on the same basis as the solvent.

The most obvious use of these facilities is associated with a change of the chromophore–solvent interactions in the excited state. Note, however, that it is quite permissible to change intramolecular terms such as bond stretching constants and reference bond lengths.

An option has also been provided in the program, which at intervals switches the identities of the ground state chromophore and the ghost molecule. As a result of this switch, the simulation is able to follow the relaxation of the solvent in response to the instantaneous creation of the excited state, responsible for the experimentally measurable time-dependent Stokes shift of dyes in solution and this data is available in the output files. Similarly, the relaxation after an instantaneous quench between the excited and ground states could be followed, with some adaptation by the user, to model vibrational cooling.

Modifications to calculate the free energy of solvation are partly based on those made for the solvent-induced shift calculations. Thus, the Hamiltonian used to propagate the system in thermodynamic integration is a weighted average of those of the initial and final states, but one needs to configurationally average the full energy difference between the states. This is readily accomplished by adapting the 'ghost' molecule scheme.

3. Illustrations

3.1 Energy decomposition

The new version of DL_POLY distinguishes the contributions of different molecular species to the total energy. This is by means of a proper bookkeeping of the energy contributions and is of use in several fields. While we concentrate here on the configurational or potential energy, decomposition of the kinetic energy is provided too. In practice, it is the total temperatures that are reported. Decomposition of the kinetic energy helps to check thermal equilibrium between all components of a simulation coupled to the same temperature bath. Adaptation to distinguish rotational and translational temperatures should be straightforward. Other possibilities that come to mind include the study of the energy loss of particles bombarding a surface, as in sputtering [10] or radiation damage [11], simulations in the presence of external fields, which may lead to significant heating or cooling of the molecular system, and thermodynamics of solvation [12].

The new facility should be useful for the study of a variety of inhomogeneous systems, such as molecules at interfaces [13]. Here, we illustrate decomposition of the potential energy, by returning to a recent study of the diffusion of water in the orthorhombic phase of the pure silica zeolite silicalite [14]. The reader is referred to this paper for details of the methods and force fields (SPC/E for water and a modification of the BKS force field for silicalite).

The microporous zeolite contains two channel systems that intersect tangentially – straight channels along the crystal *b*-direction and zig-zag or sinusoidal channels in the *ac*-plane, lying along *a* (see Figure 1). Direct diffusion along the *c* crystal direction is thus impossible. Rather, diffusion along *c* occurs via alternate half zig-zag steps in the *ac*-plane, interspaced by (at least) half unit cell steps along the *b*-direction. Simulated annealing revealed a specific adsorption site of water in the channel intersections, adsorption energy *ca.* -40 kJ/mol, and unspecific adsorption in the channels, with adsorption energy *ca.* -21.5 kJ/mol.

Interestingly, plots of the mean squared displacement of water molecules in the zeolite showed that the onset of regular diffusion along the *c*-direction lags about 100 ps behind the onsets in the *a* and *b* directions. This suggests that molecules switching between the straight and the zig-zag channels must overcome a potential energy barrier. The obvious candidate is climbing out of the specific adsorption site. The energy decomposition facility in DL_POLY allows us to visualise this barrier. Figure 1(a) and (b) shows two orthogonal projections of the zeolite framework and part of the trajectory of a particular water molecule, in a simulation at 300 K.

The energy decomposition facility in DL_POLY was used to prepare Figure 1(c), showing the sum of the water-

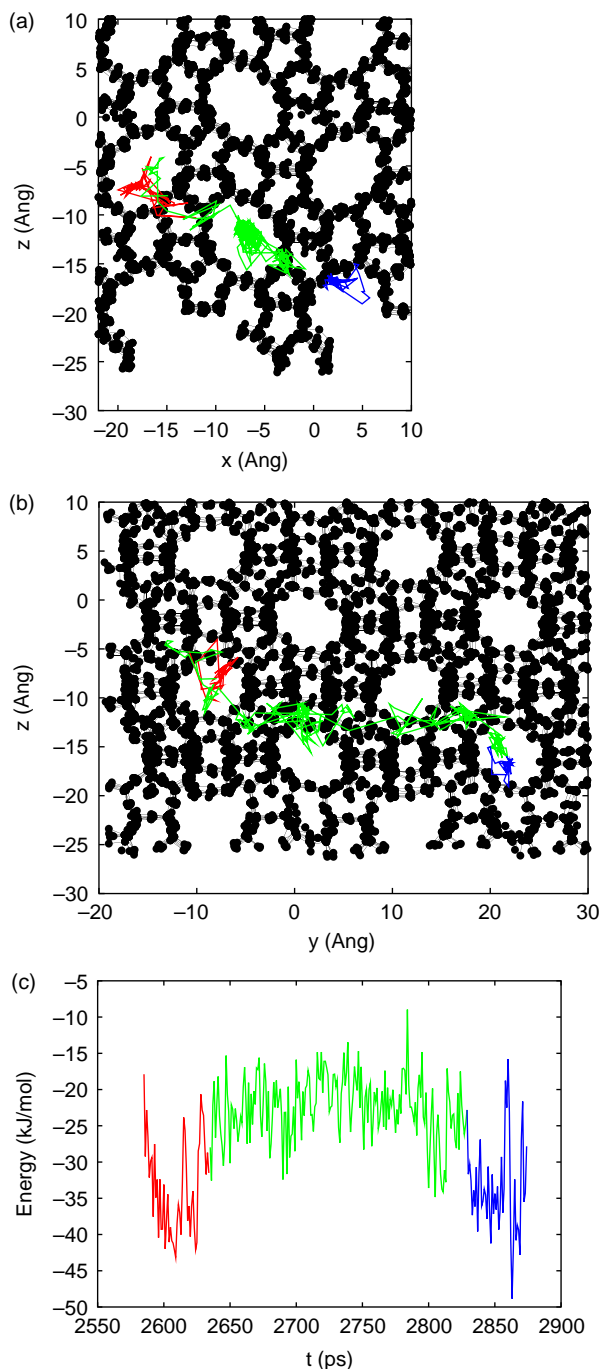


Figure 1. Illustration of the use of the energy decomposition facility. A single water molecule (coloured line shows centre of mass trajectory) diffuses 1.5 unit cells down the b crystal direction in silicalite (black), in a simulation at room temperature (see Ref. [14] for details). (a, b) Orthogonal projections on the ac and the bc planes. (c) Total water–host interaction energy vs. time. The targeted water molecule is at first trapped in a specific adsorption site (red) in the cavity formed by the intersection of the channels along a and b . A thermal fluctuation promotes it to the channel along b (green), where it diffuses until it is again trapped (blue) in another cavity (colour online). Water trapped in the cavities is about 20 kJ/mol more stable than during diffusion along the channel.

framework electrostatic and van der Waals energy terms. The trajectory is colour coded as a function of time as the molecule diffuses along the b -direction (also y in the orthorhombic unit cell). We highlight periods spent in the adsorption site at the beginning and the end, separated by diffusion down the channel. Overall, the molecule also moves half a unit cell along the c -direction (see above). Along the trajectory, we observe a ca. 20 kJ/mol jump up and down in the potential energy, consistent with the adsorption energies found in simulated annealing and with the activation of diffusion along c .

3.2 Solvation and Stokes shifts of optical spectra

3.2.1 Classical simulation of solvation shifts

The second feature introduced in the new solvation module of DL_POLY is the ability to propagate a simulation with one force field, and to periodically calculate what the configurational energy would be if part or all of the system were represented by different intra- or intermolecular force fields. The secondary force field is in no way connected with the dynamics. It merely serves to probe what the interactions *would be*. Use of this feature causes the energy decomposition to come into play at regular intervals, successively once each time with the force field used for the dynamics and with the probed force field, as described in more detail in the SI.

This facility was introduced with the aim of simulating a system switching between electronic energy surfaces. While the transition process itself obviously is an electronic process, beyond the reach of classical molecular dynamics, it often may reasonably be approximated by a sudden jump between energy surfaces (Franck–Condon approximation [15]), before and after which the nuclear motions may be described by classical mechanics.

The facility is quite general. Both intra- and intermolecular interactions may be altered in the second or ‘ghost’ force field. A possible use that comes to mind is the study of the elastic relaxation around ‘molecular magnet’ transition metal complexes [16]. Here, we illustrate the facility by simulating the solvent-induced spectral shifts of electronic absorption and fluorescence transitions in a liquid solution.

Figure 2 illustrates schematically a molecule undergoing an optical absorption–emission cycle from a configuration close to the ground state minimum. Within the present description with empirical force fields, the absorption energy of the system at time t is

$$h\nu_{\text{abs}}(t) = E^e(Q(t), R(t)) - E^g(Q(t), R(t)), \quad (1)$$

where the ground and excited state energies, E^g and E^e , respectively, are described by different force fields involving in principle both intramolecular (Q) and intermolecular (R) coordinates. For example, dependence

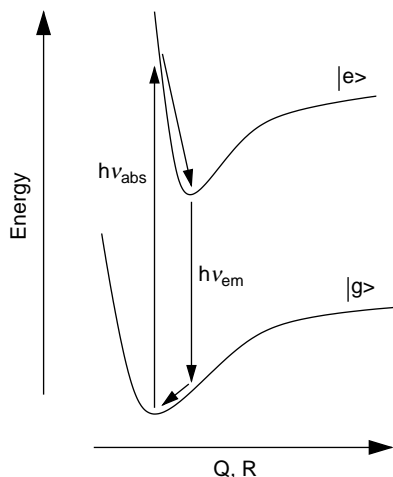


Figure 2. Schematic representation of an absorption–emission process, in which a photon (energy $h\nu_{\text{abs}}$) is absorbed with the nuclei in a configuration close to the minimum of the ground state energy surface ($|g\rangle$), carrying the system to the excited state energy surface ($|e\rangle$). The nuclei subsequently relax towards the minimum of the excited state energy surface. Emission (photon energy $h\nu_{\text{em}}$) carries the system back to the ground state. The surfaces depend on both intra- (Q) and intermolecular coordinates (R).

on the intramolecular coordinates could arise from a shift of the equilibrium geometry of the molecule and of its vibrational frequencies, reflecting the fact that in the quantum description, the electrons are less tightly bound to the nuclei in the excited state. Intermolecular interactions generally are sufficiently small compared to intramolecular interactions to be treated as an additive perturbation, so that the ground state energy can be written as a sum of intra- and intermolecular contributions:

$$E^g(Q, R) = E_{\text{intra}}^g(Q) + E_{\text{inter}}^g(R), \quad (2)$$

and similarly for the excited state e . The solvation *shift* of the optical absorption spectrum follows from sampling $E_{\text{inter}}^e(R) - E_{\text{inter}}^g(R)$ at regular intervals and binning the results. If the actual shape of the absorption spectrum is required, sampling $E_{\text{intra}}^e(Q) - E_{\text{intra}}^g(Q)$ is not satisfactory since it will not generate the progression of intramolecular vibronic bands of polyatomic molecules. A better plan is to obtain the vibronic progression either from an experiment in a weakly interacting solvent (e.g. hexane) or from a quantum calculation of vibronic frequencies and Franck–Condon overlap factors. The full vibronic (line) spectrum of the molecule *in vacuo* may subsequently be convoluted with the distribution of the solvation shifts.

Two main kinds of contribution to the solvation shift are electrostatic interactions and dispersion forces. The effects of both are explored here. They may be modelled, respectively, by introducing different partial charges or different van der Waals parameters in the excited state

force field. Variations of the charges may be deduced from quantum chemistry, followed by population analysis or electrostatic potential fitting. The variations are not systematic, but depend on the polar nature of the molecule and on the electronic states involved. For example, intramolecular charge transfer in common laser dyes such as coumarins usually increases in the $\pi - \pi^*$ excited state, but decreases in the case of $n - \pi^*$ transitions. On the other hand, some systematic variation of van der Waals parameters between the ground and excited states can be expected, since molecules in their excited states usually are less tightly bound and more polarisable. Thus, one would expect to have to increase both the van der Waals radii and the depths of the potential wells (greater polarisability in the excited state implies larger coefficients for the $-1/r^6$ terms).

Note in passing that the solvent shift of the emission spectrum is in general *not* obtained just by reversing the roles of the ground and excited states. Absorption of a photon produces a new electronic state that in general will not be in equilibrium with the solvent cage corresponding to the ground state. Spontaneous emission is a stochastic process with a Poisson distribution of waiting times, with a characteristic mean lifetime, say τ_F , of the order of a few nanoseconds in the case of strongly emitting fluorescent dyes. Only if the solvent relaxation is notably faster than the emission process will the emission correspond to that from a dye at equilibrium in a relaxed solvent cage. In general then, the emission solvation shift is time dependent, leading to a time-dependent ‘Stokes shift’ between the absorption and the emission spectrum. In fact, measurement and simulation of the time-dependent shift has long been used [17] to probe solvent relaxation on timescales of the order of 100 fs–10 ns, a subject of topical interest being solvation in ionic liquids [18].

Because fluorescence emission is not an equilibrium process, provision is made in the solvation module of DL_POLY to simulate the whole cycle in Figure 2. The user can thus switch at regular intervals between using the ground and excited state force fields, respectively, for propagation of the system, while probing the other force field. By repeating this process and binning the solvation shift at different lags after the switch, one can build up the time-dependent emission spectrum. Beware that in practice this will be expensive. Commonly, at least a thousand runs are needed to average fluctuations of the relaxation of the solvent cage. The length of the runs will naturally depend on the viscosity, but a few tens of picoseconds would not be unusual for common laboratory solvents at room temperature. Adding to this the relaxation of the ground state back to its equilibrium solvent cage to complete the cycle in Figure 2, one should thus expect to simulate a total of 10–100 ns of dynamics.

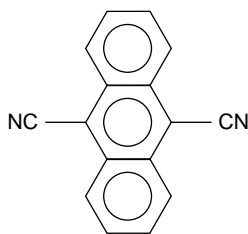


Figure 3. 9,10-Dicyanoanthracene.

3.2.2 Example: 9,10-dicyanoanthracene in methanol

As an example, we now consider the solvation shifts of 9,10-dicyanoanthracene (DCA, Figure 3) in methanol. DCA is of interest as a widely used photosensitiser in homogeneous and inhomogeneous photo-oxidation [20]. Table 1 shows measured absorption and emission shifts for the first excited singlet state of DCA, relative to the transition in vacuum. Full details of the simulations are available in Ref. [19], our purpose here being to illustrate the use of the solvation module in DL_POLY. Briefly, simulations were performed at room temperature and 1 atm pressure, in the Berendsen *NPT* ensemble [21], on a single DCA molecule in a box of 2197 methanol molecules (OPLS all atom model [22]). The model of DCA (equilibrium geometry, harmonic bond, angle and dihedral force constants, point charges) was developed with reference to B3LYP/6-311G(d,p) DFT calculations. Partial atomic charges from an electrostatic potential fit were small (the largest being $\pm 0.45e$ on the cyano group). The molecule is centro-symmetric, so it has no permanent dipole in either electronic state and charge jumps between the ground and excited states were very small, ca. $0.02e$. Solute-solvent van der Waals forces were represented by Lennard-Jones potentials. The contribution to the configurational energy of interactions between the solvent (*S*) and the dye (*D*) in state *s* (*s* = *g*, *e*) is thus a sum over pairs of atoms *i*, *j* separated by distances *r_{ij}*:

$$V_{\text{solute-solvent}}^s = \sum_{i \in D, j \in S} \frac{q_i^s q_j}{4\pi\epsilon_0 r_{ij}} + 4\pi\epsilon_{ij}^s \left(\left(\frac{\sigma_{ij}^s}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}^s}{r_{ij}} \right)^6 \right).$$

We took initial parameters for the ground state of DCA from fits to 6 – exp interaction models in the literature, which were subsequently reoptimised to improve the predicted structures and enthalpies of sublimation of two different cyanoanthracene crystals [19]. Finally, cross-terms with the OPLS model of methanol were determined using the Lorentz–Berthelot mixing rules.

Taking into account the remarks made in Section 3.2.1, the excited state was modelled by increasing the depths of the Lennard-Jones potential wells (ϵ 's) on N and C atoms

Table 1. Absorption solvation shift and Stokes shift of DCA in methanol.

	Absorption shift (cm ⁻¹)	Stokes shift (cm ⁻¹)
Exp. ^a	– 1360	– 700
Calc.	– 1300	– 480

^aRef. [19].

(ring and cyano groups) or their radii (σ 's). Indeed, as expected, analysis of the molecular orbitals shows that *H*'s do not participate in the transition. The solvation facilities in DL_POLY can be used to identify the contributions of different components in the total solvation shifts. We discuss below the influence on the total solvation shift (summing all contributions) of varying separately or together the force field parameters controlling electrostatics or the van der Waals forces. All the data shown may be extracted with the new facilities and toolkit provided with DL_POLY.

Figure 4 shows the influence on the solvation shifts of varying either the solute point charges *q* or the depths ϵ of the Lennard-Jones potential wells. From symmetry, the lowest order permanent multipole of DCA is a quadrupole. The contribution of electrostatic interactions (shifted potential approximation) to the total solute-solvent

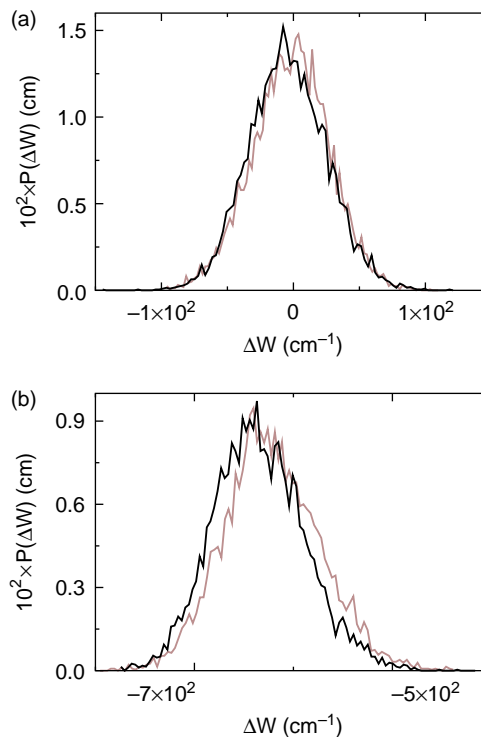


Figure 4. Distributions of the absorption (black) and the emission (grey) solvation spectral shifts of DCA in methanol as either charges only (a) or ϵ 's only (b) are changed in the excited state, averaged over 1500 configurations from runs of 10 ns.

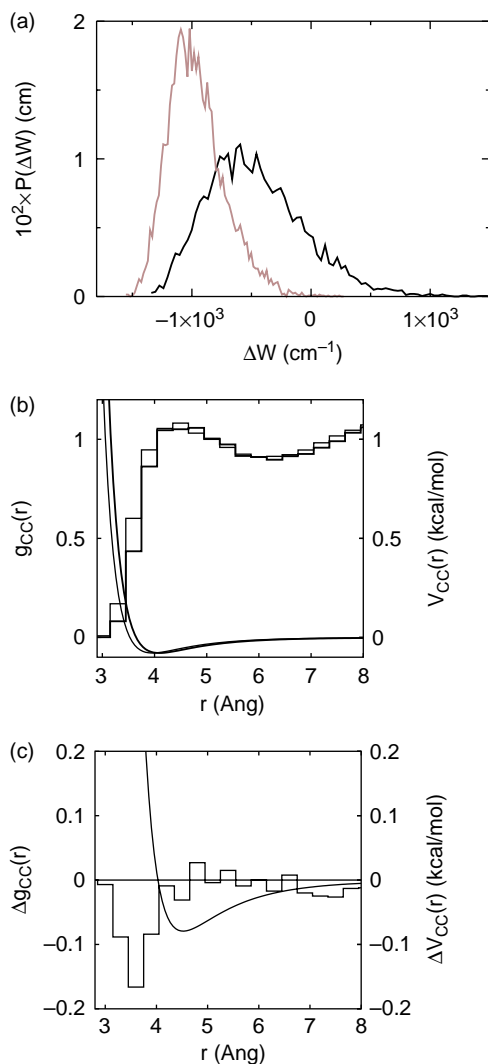


Figure 5. (a) Distributions of the absorption (black) and the emission (grey) solvation spectral shifts of DCA in methanol when the σ 's only are changed in the excited state, averaged over 1500 configurations from runs of 10 ns; (b) first peak of the C-C pair correlation function between DCA and methanol, (thin, ground state, $g_{CC}^0(r)$; thick, excited state, $g_{CC}^*(r)$), superposed on the corresponding Lennard-Jones interactions $V_{CC}^0(r)$, $V_{CC}^*(r)$; (c) differences $\Delta g_{CC}(r) = g_{CC}^*(r) - g_{CC}^0(r)$ and $\Delta V_{CC}(r) = V_{CC}^*(r) - V_{CC}^0(r)$.

interaction is correspondingly small, i.e. solvation is essentially non-polar. It is therefore not surprising that, even in the polar solvent methanol, the electrostatic contribution to the solvation shifts is very small. Fluctuations of the solvent cage lead to a broadening of about 100 cm⁻¹ FWHM. Next, the solute-solvent Lennard-Jones potential well depths are increased in Figure 4(b) by 7%, this value being a typical order of magnitude of the ratio of solvation absorption spectral shifts to enthalpies of sublimation in non-polar systems [23,24]. Increasing only the depth of the Lennard-Jones wells induces a solvation red shift of ≈ 620 cm⁻¹ in both the absorption and

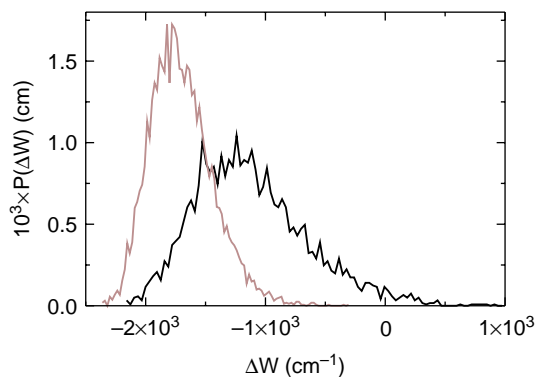


Figure 6. Spectral shift distribution for the absorption (black) and the emission (grey) of DCA in methanol as charges σ and ϵ change.

the emission. The absorption and emission shifts differ little because the small increase in the depth of the potential wells does not influence the *structure* of the solvent cage, i.e. relaxation in the excited state is negligible in this case. Note that in both the above cases, the distributions are symmetrical about their mean values.

Increasing just the Lennard-Jones radii for solute-solvent interactions, here by 5%, has a stronger effect, see Figure 5(a). The mean absorption shift, 600 cm⁻¹, is larger than previously. The mean emission shift is about 1020 cm⁻¹. The distribution of the solvation shifts is sharper and less skewed in the excited state. Figure 5(b) shows that the solvation shell expands slightly in response to the creation of the excited state, as shown by the CC pair correlation functions between methanol and DCA in its ground and excited states, $g_{CC}^0(r)$ and $g_{CC}^*(r)$. This is highlighted as a hole in $\Delta g_{CC}(r) = g_{CC}^*(r) - g_{CC}^0(r)$ between 3 and 4 Å in Figure 5(c). Observe in Figure 5(b) that some equilibrium C-C distances are mapped into the repulsive wall of the excited state potential function, explaining the tail of positive solvation shifts observed in the simulation. On the other hand, equilibrated pair distances in the excited state tend to be mapped back to the attractive, but (compared to the wall) flatter part of the ground state interaction potential, beyond its minimum. Hence, the distribution of equilibrium emission shifts is sharper, less skewed and on average greater than that of absorption (see also Figure 5(c)).

Figure 6 shows the total solvation shifts when all three changes are combined. The average solvation shifts and the Stokes shift between absorption and emission are in fair agreement with the experimental values (Table 1).

3.3 Free energy differences

3.3.1 Theory

The third new capability added to DL_POLY is thermodynamic integration to obtain free energies of

solvation (a feature which can also be exploited in studies not related to solubility). Atoms of the system are classified as an unchanged part, the solvent, S , an initial solute or state of a solute, A , or a final solute, B , which just might be A , but with a different force field. One wishes to estimate the free energy difference between states described by initial and final Hamiltonians H_0 and H_1 defined by

$$H_0 = K_S + K_A + V_{AA} + V_{AS} + V_{SS}, \quad (3)$$

$$H_1 = K_S + K_B + V_{BB} + V_{BS} + V_{SS}, \quad (4)$$

where K and V refer, respectively, to kinetic and potential energy contributions. Thermodynamic integration is a well-established technique for calculating free energies (see [25,26]). It is based on the concept of a ‘mixed’ Hamiltonian which combines two systems together in one:

$$H_\lambda = (1 - \lambda)H_0 + \lambda H_1, \quad (5)$$

where $\lambda(0 \leq \lambda \leq 1)$ is the so-called ‘mixing parameter’. For such a Hamiltonian, it is easy to show that in the canonical (NVT) ensemble,

$$\frac{\partial A}{\partial \lambda} = \langle H_1 - H_0 \rangle_\lambda, \quad (6)$$

in which A represents the Helmholtz free energy and the term $\langle H_1 - H_0 \rangle_\lambda$ is the average energy difference between *full* Hamiltonians H_1 and H_0 in a system governed by the *mixed* Hamiltonian (5) with a fixed value of λ . It follows from Equation (6) that

$$\Delta_0^1 A = \int_0^1 \langle H_1 - H_0 \rangle_\lambda d\lambda, \quad (7)$$

where $\Delta_0^1 A$ is the free energy difference between the independent systems described by Hamiltonians H_0 and H_1 . In practice, integral (7) is calculated numerically from a series of averages, each obtained from a molecular dynamics simulation using a distinct value of λ .

One has, of course, to be able to simulate the system described by the mixed Hamiltonian. Mixing the full Hamiltonians, necessary for estimation of Gibbs free energy differences, leads to problems arising from an effective rescaling of the atomic masses in the kinetic contribution, and to instabilities in the integration of the equations of motion. Though DL_POLY does permit scaling of kinetic energies if required, it is fortunately not necessary for calculations of the free energy of solvation in the canonical ensemble. For this purpose, one may redefine $H(\lambda)$ as

$$H(\lambda) = K + (1 - \lambda)V_0 + \lambda V_1, \quad (8)$$

where K is the kinetic energy common to both states 0 and 1, and V_0 and V_1 are the potential energies:

$$V_0 = V_{AA} + V_{AS} + V_{SS}, \quad (9)$$

$$V_1 = V_{BB} + V_{BS} + V_{SS}, \quad (10)$$

$$K = K_A + K_B + K_S. \quad (11)$$

These equations describe transformation of a solute A and non-interacting degrees of freedom corresponding to B into solute B with non-interacting degrees of freedom of A . The mixed Hamiltonian is thus:

$$H(\lambda) = K_A + K_B + K_S + (1 - \lambda)(V_{AA} + V_{AS}) + \lambda(V_{BB} + V_{BS}) + V_{SS}. \quad (12)$$

Thus, unlike the mixing rule (5), even when atoms are added or deleted on going from A to B , there is no kinetic energy correction term, since, even at $\lambda = 0$ (or 1), A and B are both simultaneously present, one physically and the other as degrees of freedom.

The estimate of the Helmholtz free energy difference is thus:

$$\Delta_0^1 A = \int_0^1 \langle V_1 - V_0 \rangle_\lambda d\lambda. \quad (13)$$

The reader will notice that for $\lambda \neq 0, 1$, solutes A and B interact indirectly via their interactions with a common solvent S , notably their contributions to the pressure. It is therefore important to ensure that the numerical integration of (13) is pursued sufficiently close to the limits.

It will furthermore be noticed that the $\lambda = 0, 1$ states above are not directly relevant to measurable free energy differences. Typically one wishes to compute the free energy difference between dissolving A or B in S , in which case account has to be taken of the free energy difference involved in returning (resp. taking) A (B) to or from their respective reference states, as illustrated in Figure 7. When the Hamiltonians are mixed, the value $\Delta_0^1 A$ computed above should be corrected by adding $\langle H_{AA} \rangle - \langle H_{BB} \rangle$. Similarly, the correction in potential mixing is $\langle V_{AA} \rangle - \langle V_{BB} \rangle$.

3.3.2 Mixing rules

It is well known that calculating the required averages in Equation (6) or (13) is problematic when λ approaches the values of 0 or 1. Near these extremes, it is the full difference of the Hamiltonians which is sampled, but in configurations whose evolution is governed by the *weighted* Hamiltonian H_λ . Typically, this leads to sampling bad van der Waals contacts, so that the averages

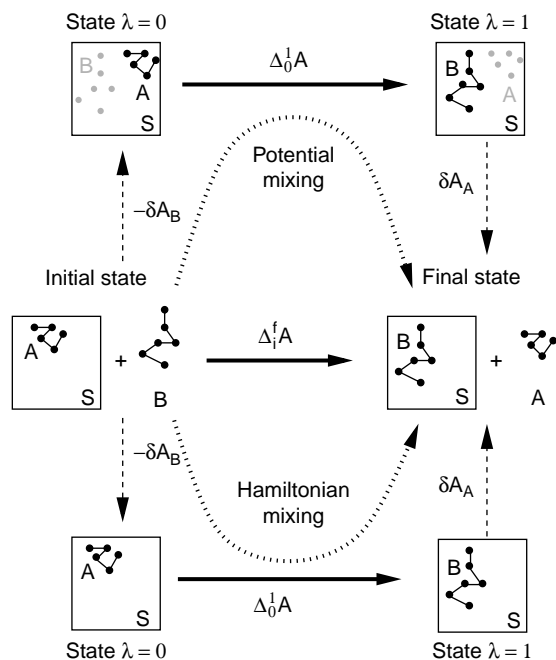


Figure 7. Thermodynamic cycles for the solvation free energy difference of two species A and B in a solvent S. Initial state: A in solution and B in some reference state; final state: B in solution and A withdrawn, in its reference state. Upper cycle: at the end points of integration in potential mixing, B (resp. A) is present as degrees of freedom. The corrections $-\delta A_B$ and δA_A correct for suppressing/reinstating internal interactions in B or A. Lower cycle: in Hamiltonian mixing, the corrections must allow for annihilation (creation *ex nihilo*) of B (A).

are subject to large fluctuations. A useful technique to help with this difficulty is to replace the simple mixing parameter λ by a function $f(\lambda)$ so that Equation (7) for example becomes

$$\Delta_0^1 A = \int_0^1 \left\langle (H_1 - H_0) \left(\frac{\partial f(\lambda)}{\partial \lambda} \right) \right\rangle_\lambda d\lambda. \quad (14)$$

The function $f(\lambda)$ is chosen so that its derivative with respect to λ tends to zero at $\lambda = 0$ and $\lambda = 1$, while $f(\lambda)$ itself tends to 0 and 1, respectively, at these locations. Appropriate forms of $f(\lambda)$ thus dampen the large fluctuations associated with the end points of the thermodynamic integration and obviate the need for simulations at the end points where the averages are formally zero. This strategy shifts the numerical difficulty from integrating divergences at the end points to accurate sampling of a fast-varying function in a narrow internal subinterval as done by Beutler et al. [27]. In DL_POLY, we have included the following forms for $f(\lambda)$:

- (1) Standard linear mixing:

$$f(\lambda) = \lambda,$$

- (2) Non-linear mixing:

$$f(\lambda) = 1 - (1 - \lambda)^k,$$

where k is a positive integer;

- (3) Trigonometric mixing:

$$f(\lambda) = \frac{1}{2} \{1 + \sin[\pi(\lambda - 1/2)]\},$$

- (4) Error function mixing:

$$f(\lambda) = \frac{\alpha}{\sqrt{\pi}} \int_0^\lambda \exp[-\alpha^2(x - 1/2)^2] dx,$$

where α^2 is a real parameter of order 10–11; and

- (5) Polynomial mixing:

$$f(\lambda) = 1 - (1 - \lambda)^k \sum_{i=0}^{k-1} \frac{(k-1+i)!}{(k-1)!i!} \lambda^i,$$

where k is a positive integer.

Figure 8 compares these weighting functions. Forms (3)–(5) have desirable properties with regard to both end points of the thermodynamic integration, while form (2) is suitable for systems with problems at one end point, as when atoms are created or deleted only. Not all forms of weighting are equally proficient at damping the fluctuations at the extremes of the range of λ , as will be seen in examples presented below.

In general thermodynamic integration, if the initial and final systems are very similar, for example substitution of a chemical group by another of a very similar size and interactions, the perturbation induced by the substitution is minor. The simple linear mixing rule may then give reasonable results. For more complex substitutions, it is undesirable to sample the full interactions $\langle V_1 - V_0 \rangle$ when particles are actually disappearing in the dynamics. This

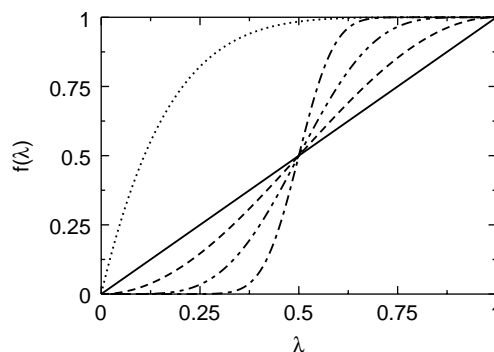


Figure 8. Weighting functions $f(\lambda)$ of the mixing parameter λ in thermodynamic integration: linear (straight line), non-linear (here $k = 6$, dotted line), trigonometric (dashed line), error function (dot-dashed line) and polynomial ($k = 6$, double-dot-dashed line).

situation can be avoided by using one of the mixing rules (3)–(5). On the other hand, the faster the approach of $f(\lambda)$ to 0 or 1 at the end points, the steeper will be its variation in the centre of the interval. Attention should then be paid to the spacing of the discrete λ values to achieve accurate results. These points are illustrated in the next section.

3.3.3 Illustration: excess free energy of liquid dimethylsulphoxide

We illustrate the calculation of free energy differences with the Helmholtz excess free energy of dimethylsulphoxide (DMSO), for an all atom model reparameterised by Bordat et al. [28]. Simulations were performed in the Berendsen *NVT* ensemble [21], at room temperature and atmospheric pressure. Formally, as λ ranges from 0 to 1, we insert one extra molecule in a box already containing 430 other molecules (representing the experimental density). In practice, the transformation is run in reverse, gradually removing the extra molecule. DMSO was represented as a rigid unit. A 2 fs timestep was used and electrostatic interactions were treated by the reaction field method with a dielectric constant $\epsilon = 46$.

We evaluate the excess free energy by removing one molecule from the box between $\lambda = 1$ and $\lambda = 0$. Referring to Figure 7, solute *A* is absent and solute *B* is a DMSO molecule, here a rigid unit. Since we used potential energy mixing, the correction to $\Delta_0^1 A$ is $\langle V_{AA} \rangle - \langle V_{BB} \rangle$. However, $\langle V_{AA} \rangle$ vanishes trivially and $\langle V_{BB} \rangle = 0$ for the rigid unit modelling of the DMSO molecule. Therefore, the experimentally relevant excess free energy $\Delta_i^f A$ is in this case simply $\Delta_0^1 A$.

Removing a molecule represents a rather violent perturbation for the rest of the solution which has to fill in the hole left by the disappearing molecule. We apply and comment on all the available mixing rules to illustrate the points made in the previous section.

Taking account of the direction in which the transformation is run in the simulations, the estimated excess free energy is

$$\Delta_{\text{exc}} F = -\Delta_1^A A = -\int_1^A \frac{dA}{d\lambda} d\lambda. \quad (15)$$

$\Delta_{\text{exc}} F$ is plotted in Figure 9 as a function of the lower bound Λ , for the non-linear mixing rule (2) with $k = 3$ to $k = 10$. Note the logarithmic spacing of the λ points used to try to control the divergence of the integrand close to $\lambda = 0$. Despite this precaution, the integral is unconverged (non-zero slopes of the graphs) for all values of k .

Figure 10 compares the linear mixing (rule (1)) with the trigonometric and the error function weightings. As expected, linear mixing leads to poor convergence. The trigonometric weighting function gives better results, though visibly still not completely converged. Error

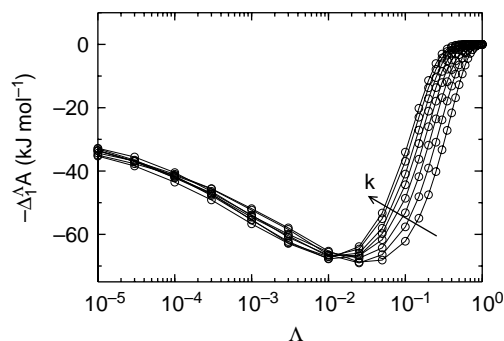


Figure 9. Convergence of the excess free energy of DMSO as a function of the lower bound of thermodynamic integration, $\Delta_1^A A$, Equation (15), for mixing rule number (2) for different exponents k in the [3–10] range. Convergence of the integral is unsatisfactory in all cases.

function weighting gives fully converged results, even for Λ as large as 0.1. On the other hand, the integral varies very fast for Λ close to 0.5, requiring more sampling points than the trigonometric function. The excess free energy estimated from the error function is -35.4 kJ/mol.

Finally, Figure 11 shows $\Delta_1^A A$ obtained with the polynomial weighting function. This mixing rule, displayed in Figure 8 for $k = 6$, has a similar behaviour to the error function weighting, leading to a converged value for the excess free energy. We suspect that a compromise between convergence at the end points and not-too fast decay of the polynomial function in the middle of the interval, can be more easily struck by tuning the k exponent than by tuning α in the error function weighting (in fact, α can be changed only by recompiling the code). Here, $k = 6$ yields good results. Higher values lead to a behaviour similar to the error function. $\Delta_1^A A$ depends slightly on k in Figure 11. This reflects room for improvement of the sampling points in the centre of the interval [0,1]. However, for $k = 3$ to $k = 8$, the excess free energy converges quickly towards a value of -30.3 kJ/mol. This result is close to the experimental value of -29.7 kJ/mol [29] and to a previous thermodynamic integration calculation [28], with the program YASP [30]. In

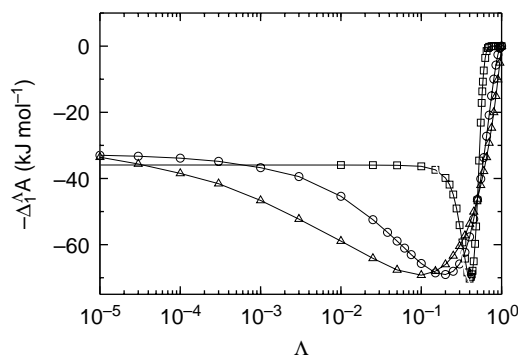


Figure 10. $\Delta_1^A A$ for the linear (triangles), trigonometric (circles) and error function (squares) mixing rules.

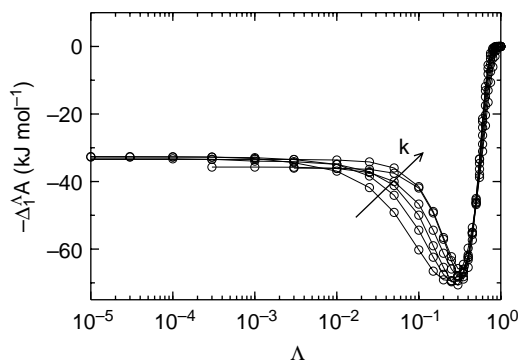


Figure 11. Δ_1^A for mixing rule (5) for k in the [3–8] range.

conclusion, thermodynamic integration can provide accurate results, provided attention is paid both to the mixing rule, determined by the similarities or differences characterising the end points, and to the sampling points, determined in turn by the weighting function.

4. Conclusions

It is to be hoped that the new facilities for solvation studies in DL_POLY will be relevant to the work of a wide range of users. Further developments should be possible, such as adapting the force-field switching facility from its present application in spectral shift calculations to the calculation of vibrational relaxation. Another possible use is to switch between the full molecular force field of a system and one either with attenuated interactions or a coarse-grained model, which might prove useful in some problems of exploring the configurational space.

Acknowledgements

This work was supported by the Franco-British Alliance programme.

References

- [1] W. Smith, C.W. Yong, and P.M. Rodger, *DL_POLY: Applications to molecular simulation*, Mol. Simul. 28 (2002), pp. 385–471.
- [2] W. Smith, *Special issue: DL_POLY: Applications to molecular simulation II*, Mol. Simul. 32 (2006), pp. 933–1121.
- [3] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, *GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation*, J. Chem. Theory Comput. 4 (2008), pp. 435–447.
- [4] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, and K. Schulten, *Scalable molecular dynamics with NAMD*, J. Comput. Chem. 26 (2005), pp. 1781–1802.
- [5] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M.J. Karplus, *CHARMM: A program for macromolecular energy, minimization, and dynamics calculations*, J. Comput. Chem. 4 (1983), pp. 187–217.
- [6] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, and L.G. Pedersen, *A smooth particle mesh Ewald method*, J. Chem. Phys. 103 (1995), pp. 8577–8593.
- [7] J. Tersoff, *Modelling solid state chemistry: Interaction potentials for multicomponent systems*, Phys. Rev. B 39 (1989), pp. 5566–5568.
- [8] M.W. Finnis and J.E. Sinclair, *A simple empirical N body potential for transition metals*, Philos. Mag. A 50 (1984), pp. 45–55.
- [9] M.S. Daw and M.I. Baskes, *Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals*, Phys. Rev. B 29 (1984), pp. 6443–6453.
- [10] V. Georgieva, I.T. Todorov, and A. Bogaerts, *Molecular dynamics simulation of oxide thin film growth: Importance of the inter-atomic interaction potential*, Chem. Phys. Lett. 485 (2010), pp. 315–319.
- [11] I.T. Todorov, J.A. Purton, N.L. Allan, and M.T. Dove, *Simulation of radiation damage in gadolinium pyrochlores*, J. Phys. Condens. Matter 18 (2006), pp. 2217–2234.
- [12] M.E. Lee and N.F.A. van der Vegt, *Molecular thermodynamics of methane solvation in tert-butanol–water mixtures*, J. Chem. Theory Comput. 3 (2007), pp. 194–200.
- [13] J. Chowdhary and B.M. Ladanyi, *Hydrogen bond dynamics at the water/hydrocarbon interface*, J. Phys. Chem. B 113 (2009), pp. 4045–4053.
- [14] P. Bordat, P.A. Cazade, I. Baraille, and R. Brown, *Host and adsorbate dynamics in silicates with flexible frameworks: Empirical force field simulation of water in silicalite*, J. Chem. Phys. 132 (2010), pp. 94501/1–94501/8.
- [15] P.W. Atkins, *Molecular Quantum Mechanics*, Clarendon Press, Oxford, 1990.
- [16] P. Guionneau, M. Marchivie, G. Bravic, J.F. Létard, and D. Chasseau, *Structural aspects of spin crossover. Example of the [Fe II Ln (NCS) 2] complexes*, Top. Curr. Chem. 234 (2004), pp. 97–128.
- [17] E.A. Carter and J.T. Hynes, *Solvation dynamics for an ion pair in a polar solvent: Time-dependent fluorescence and photochemical charge transfer*, J. Chem. Phys. 94 (1991), pp. 5961–5979.
- [18] H. Jin, X. Li, and M. Maroncelli, *Heterogeneous solute dynamics in room temperature ionic liquids*, J. Phys. Chem. B 111 (2007), pp. 13473–13478.
- [19] P.A. Cazade, *Interactions chromophore-solvent dans des matériaux hybrides*, Université de Pau et des pays de l'Adour, 64053 Pau Cedex, France (2008).
- [20] C. Cantau, S. Larribau, T. Pigot, M. Simon, M. Maurette, and S. Lacombe, *Oxidation of nauseous sulphur compounds by photocatalysis or photosensitization*, Catal. Today 122 (2007), pp. 27–38.
- [21] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, and J.R. Haak, *Molecular dynamics with coupling to an external bath*, J. Chem. Phys. 81 (1984), pp. 3684–3690.
- [22] W.L. Jorgensen, D.S. Maxwell, and J. Tirado-Rives, *Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids*, J. Am. Chem. Soc. 118 (1996), pp. 11225–11236.
- [23] R.J. Babbitt and M.R. Topp, *Jet spectroscopy of anthracene–alkane complexes*, Chem. Phys. Lett. 135 (1987), pp. 182–188.
- [24] E.U. Wallenborn, R. Brown, and U.P. Wild, *Analysis of the optical spectra of aromatic-alkane clusters*, J. Chem. Phys. 107 (1997), pp. 8338–8348.
- [25] M.P. Allen and D.J. Tildesley, *Computer Simulation of Liquids*, Clarendon Press, Oxford, 1989.
- [26] D. Frenkel and B. Smit, *Understanding Molecular Simulation*, Elsevier/Academic Press, New York, 2002.
- [27] T.C. Beutler, A.E. Mark, R.C. van Schaik, P.R. Gerber, and W.F. van Gunsteren, *Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations*, Chem. Phys. Lett. 222 (1994), pp. 529–539.
- [28] P. Bordat, J. Sacristan, D. Reith, S. Girard, A. Glättli, and F. Müller-Plathe, *An improved dimethyl sulfoxide force field for molecular dynamics simulations*, Chem. Phys. Lett. 374 (2003), pp. 201–205.
- [29] J.A. Riddick, W.B. Bunger, and T.K. Sakand, *Organic Solvents Physical Properties and Methods of Purification*, New York, Wiley, 1986.
- [30] F. Müller-Plathe, *YASP: A molecular simulation package*, Comput. Phys. Commun. 78 (1993), pp. 77–94.