



ACADÉMIE
DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus

Chimie

Mario Araujo-Rocha, Alejandro Diaz-Marquez and Guillaume Stirnemann

On the validity of some equilibrium models for thermodiffusion


Volume 27, Special Issue S5 (2024), p. 95-106

Online since: 31 May 2024

Part of Special Issue: French Network on Solvation (GDR 2035 SolvATE)

Guest editor: Francesca Ingresso (Université de Lorraine–CNRS, LPCT UMR 7019, Nancy, France)

<https://doi.org/10.5802/crchim.283>

 This article is licensed under the
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.
<http://creativecommons.org/licenses/by/4.0/>



The Comptes Rendus. Chimie are a member of the
Mersenne Center for open scientific publishing
www.centre-mersenne.org — e-ISSN : 1878-1543



Research article

French Network on Solvation (GDR 2035 SolvATE)

On the validity of some equilibrium models for thermodiffusion

Mario Araujo-Rocha^{Ⓜ, a}, Alejandro Diaz-Marquez^{Ⓜ, a} and Guillaume Stirnemann^{Ⓜ, *, a, b}

^a CNRS Laboratoire de Biochimie Théorique, Institut de Biologie Physico-Chimique, Université Paris Cité, 13 rue Pierre et Marie Curie, 75005, Paris, France

^b PASTEUR, Département de chimie, École normale supérieure, PSL University, Sorbonne Université, CNRS, 75005 Paris, France

E-mails: araujorochoa@ibpc.fr (M. Araujo-Rocha), diazmarquez@ibpc.fr (A. Diaz-Marquez), guillaume.stirнемann@ens.psl.eu (G. Stirnemann)

Abstract. When applied to binary solutions, thermal gradients lead to the generation of concentration gradients and thus to inhomogeneous systems. While being known for more than 150 years, the molecular origins for this phenomenon are still debated, and there is no consensus on the underlying physical models or theories that could explain the amplitude of the concentration gradient in response to a given temperature gradients. Notably, there have been some attempts to relate this non-equilibrium, steady-state manifestation, to equilibrium properties of these solutions, for example, to the temperature dependence of the self-diffusion coefficient or to the solvation free energies of each of their components. Here, we use molecular dynamics simulations on dilute solutions containing molecular-size solutes, both in a thermophoretic setting as well as under equilibrium conditions, to test the validity of such models. We show that these approaches are inadequate and lead to completely uncorrelated estimates as compared to those based on the out-of-equilibrium measurements. Crucially, they fail to explain the strong mass dependence (to which thermodynamics or single-particle diffusion are insensitive) observed in the simulations and measured in the experiments. However, our results suggest an interesting correlation between the amplitude of the short-time molecular motion and that of the concentration gradient that would deserve future investigations.

Keywords. Thermophoresis, Solvation free energy, Translational diffusion.

Funding. European Research Council under the European Union's Eighth Framework Program (H2020/2014-2020)/ERC Grant Agreement No. 757111, "Initiative d'Excellence" program from the French State (Grant "DYNAMO", ANR-11-LABX-0011-01).

Manuscript received 13 October 2023, revised 20 December 2023, accepted 10 January 2024.

1. Introduction

When subjected to an inhomogeneous spatial temperature distribution, a liquid mixture's local composition usually becomes temperature- (and thus space-) dependent. This phenomenon is known as thermodiffusion, thermophoresis (especially when

considering large particles in a solvent), or Soret effect [1–3]. In order to explain such experimental observations, that can be quantified with a wide range of techniques [4], a typical phenomenological approach consists of considering a particle current that is proportional to the concentration and to the temperature gradient (through a factor known as "thermal diffusion coefficient") and which adds up with the regular Fickian diffusion current, proportional

*Corresponding author

(through the regular diffusion coefficient) to the concentration gradient [2]. At steady state, the absence of net fluxes results in an exponentially distributed concentration of particles as a function of temperature, which is consistent with the experimental measurements. The temperature dependence of the concentration is called the Soret coefficient, and it corresponds to the ratio of thermal and regular diffusion coefficients. This phenomenon is observed for many different systems, across many lengthscales, and the present contribution will exclusively discuss and focus on molecular-size diffusing objects.

These considerations remain purely phenomenological. A vast body of experimental and computational work has attempted to go beyond this picture and characterise the molecular determinants of thermodiffusion (see, e.g., [2]). While being an “old” problem, this remains a partly open question and a timely topic, with very recent work testing phenomenological thermodynamic models using molecular dynamics simulations of Lennard-Jones particles [5–7].

It is now widely accepted that the Soret coefficient can be decomposed as a sum of a purely chemical component (which depends on the interatomic interactions between the components of the mixture), a mass-dependent component, sensitive to the difference in masses between the mixture particles, and a component sensitive to the difference in the moments of inertia [2,8,9]. This decomposition was extensively studied through molecular simulations of hard-spheres and Lennard-Jones binary fluids [7]. Very recent work completed this picture by highlighting the effect of the mass dipole [10] (i.e., the first moment of the mass distribution in the molecules). If we take the example of aqueous mixtures, the so-called chemical component has been related to the hydrophobic/hydrophilic nature of the solutes, in particular, to the strength of their hydrogen-bond interactions with water [3,11,12].

Although interesting, making such connections does not really provide an underlying molecular picture of the phenomenon. At the turn of the 20th century, Einstein’s work paved the way for a molecular understanding of single particle diffusion [13]. The diffusion coefficient was thus connected to the mean square displacement of molecules, as well as to the atomistic details of the diffusing particle (e.g., its radius) and of the diffusing medium (e.g., its viscosity).

As opposed to regular diffusion, a major challenge is that thermophoresis is, in essence, a phenomenon out of thermodynamic equilibrium, which considerably limits the theoretical framework and tools that can be used. While many attempts have been made for thermodiffusion [2,3,14–19] there is currently no “universal”, widely-accepted theoretical model that is able to quantitatively connect the Soret coefficient or the thermal diffusion coefficient to the structural and thermodynamical characteristics of the mixture, especially for molecular-size solutes. We note however that models that contain at least some empirical or phenomenological ingredients were shown to correlate well with the thermophoretic behavior of hard-spheres and Lennard-Jones particles [7], although some limitations were also pointed out [5].

Despite these practical limitations, several appealing models have been proposed that rely on equilibrium considerations to explain the steady-state of a liquid system under a thermal gradient. For example, Eastman derived the following equation (written here with modern notations) for a particle immersed in a fluid [14]:

$$S_T = \frac{1}{k_B T} \frac{dG}{dT} \quad (1)$$

where G is the free energy of the particle in the solution. This idea was later reinforced by a number of experimental results from the Braun group on colloidal suspensions [16,20,21] that were shown to agree with such models, which therefore assume local equilibria along the temperature gradient. In other words, the particles in a solution would, therefore, tend to accumulate in the regions where the system can minimize its free energy; that is, it follows the gradient of solvation free energy. The validity of such a picture has often been questioned, and it was shown that its applicability could depend on the lengthscale of the system considered [19].

Another interesting theory relates the Soret coefficient of particle 2 in a solution with 1 to the differences in the activation energy for single-particle diffusion E_a of the two particle types 1 and 2 in the mixture [17]:

$$S_T = \frac{E_a^2 - E_a^1}{RT^2} \quad (2)$$

In other words, the molecules with the highest activation energies for diffusion would tend to accumulate in the cold regions (positive Soret coefficients).

This equation echoes a phenomenological expression where the diffusion activation energy is replaced by the heat of transport of the species in the mixture. It is important to note that there seems to be a sign typo in the original publication [17], the final equation (7.9 in the original article) not being consistent with the discussion in the text.

Measuring the single-particle diffusion of molecular systems or their solvation free energy is not straightforward in the experiments; the validity of these models has often been discussed based on some measured correlations, for example, with the solute surface area [16,20], or based on theoretical arguments only [14]. For example, one obvious limitation of these models is that their ingredients are mass-independent, i.e. they cannot explain, without further corrections, the mass dependence of the Soret coefficient. In order to test these models more thoroughly, we used molecular dynamics in explicit solvent, replicating experimental thermophoretic settings and equilibrium conditions. We used two types of solutions, starting from binary mixtures of Lennard-Jones (LJ) particles as well as aqueous solutions. We focused on the dilute regime (from the perspective of the solute).

We show that these equilibrium models are not adequate. There is no obvious correlation between the Soret coefficient and the solvation free energy of the dilute particle or the difference between its activation energy for diffusion and that of the solvent. These models fail to reproduce qualitatively the measured trend, and their predictions are sometimes off by one order of magnitude. We finish by discussing some interesting correlations between the Soret coefficient in these mixtures, which suggest that, for very different systems ranging from the binary LJ solutions to several aqueous mixtures and for a wide range of molecular masses that can be artificially tuned in the simulation, the Soret coefficient is directly correlated to the relative difference between the amplitude of the solute motion at short timescales and that of the solvent.

2. Methods

2.1. Systems

The simulations of the LJ binary mixtures considered Argon-like particles [22], with the solvent being

described with $\epsilon = 0.2381$ kcal/mol, $\sigma = 3.405$ Å and $m = 39.94$ u. Solute particles were identical and only differed in their ϵ values and sometimes their mass. The reference temperature was thus $T = \epsilon/k_B = 119.8$ K, with a corresponding $T^* = 1$ in reduced units. As detailed below, simulations were typically performed at reduced temperatures ranging from $T^* = 0.9$ to 1.1 (107.82 to 131.78 K) and reduced pressure $P^* = 0.8$ (33.52 MPa). We have thoroughly checked that the system remains in a liquid phase in these conditions.

The all-atom molecular simulations of the binary mixtures were performed at temperatures ranging from $T = 300$ K to 360 K and a pressure of 1 bar. The force field parameter models for trimethylamine N-oxide (TMAO) [23,24], urea [25], methanol [26], glucose [27] and water [28] were extracted directly from the available literature which fulfill the working temperature and pressure conditions in aqueous solutions.

Solvent particles were randomly integrated into the simulation box through the utilization of the Packmol package [29]. Simultaneously, the solute particles were uniformly inserted along the z -direction of the thermal gradient, leading to an initial configuration characterized by a flat concentration profile. The Coulombic cut-off was 8.5 Å and the LJ cut-off was 9 Å, with particle–particle particle–mesh (PPPM) [30] solver for the long-range electrostatic forces, in the case of dilute aqueous solutions. The particles were randomly generated using LAMMPS software version 07-Aug-19's [31] internal random atom generator for the LJ systems. The cut-off for non-bonded interactions was set to 2.5σ (9.534 Å) for the LJ systems. Timesteps were 1 fs in all cases.

2.2. Thermophoretic simulations

All simulations under a thermal gradient were performed following a protocol described before [32], with 12 steps from preparation to production, using the LAMMPS software version 07-Aug-19 [31] with the eHex algorithm to generate a temperature gradient [33]. These include some energy minimizations, equilibration steps in successive thermodynamic ensembles, activation of a heat exchange algorithm [33] and a subsequent equilibration of the temperature and concentration gradients, and finally a production stage.

For the LJ mixtures, we started with boxes of $7.65\sigma \times 7.65\sigma \times 50\sigma$, containing 2048 atoms of the solvent and 80 atoms of the solute. The same aforementioned protocol [32] was performed on these systems, adapting the lengths of each step to their reduced time equivalents. The thermal gradient was attained by applying the eHex algorithm of heat exchange [33] in two 4σ -long slabs over the z -axis; kinetic energy was removed from the atoms present in the region between $z = 10.5\sigma$ and 14.5σ and applied to the atoms present in the region between $z = 35.5\sigma$ and 39.5σ . A scheme of the system and of the temperature gradient is shown in Figure 1.

The thermophoretic simulations of the molecular binary mixtures started with boxes with 25 \AA length along the x and y axes and 50 \AA along z . They contain 1024 molecules of water and 40 molecules of the solute. The resulting systems are identical, in terms of construction, to the LJ mixtures shown in Figure 1. The thermostatted regions are positioned along the z -axis at 12.5 \AA and 27.5 \AA with a thickness of 4 \AA , creating two symmetrically located regions of 25 \AA .

The exchanged heat flux in the eHex algorithm was set such that the generated temperature differences between the cold and hot slab were 60 K for the aqueous solutions ($Q = 0.0375 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{fs}^{-1}$) and 0.2 reduced temperature-unit (23.96 K) for the LJ systems ($Q = 5\epsilon/\tau$, or $0.00054 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{fs}^{-1}$).

All presented error bars for the Soret coefficient are calculated from the standard deviation of multiple simulations with different starting configurations.

2.3. Free energy perturbation

We first detail the hydration free energy calculations for the solute molecules in dilute aqueous solutions. The simulation process involved employing cubic simulation boxes of 33.15 \AA -side, filled with solvent molecules inserted randomly without spatial constraints. The system was minimized, the velocities were set to correspond to the target median temperature followed by an isobaric-isothermal NPT equilibration of 200 ps at 100 kPa. The decoupling process involved progressively reducing Coulombic and LJ interactions over 40 steps using soft-core potentials by introducing a tunable parameter (λ). These simulations were performed without an induced temperature gradient.

Simulations were executed using both the LAMMPS software version 07-Aug-19 [31], with its FEP package, and the GROMACS package version 2019.4 [34] to validate the results. LAMMPS simulations were performed at 8.5 \AA Coulombic cutoff, 9 \AA LJ cutoff and using the particle-particle particle-mesh (PPPM) solver for long-range electrostatic forces. The damping parameters for the Nose-Hoover [35,36] barostat and thermostat are $P_{\text{damp}} = 1 \text{ ps}$ and $T_{\text{damp}} = 0.1 \text{ ps}$ respectively, with 3 Nose-Hoover chains and the velocity Verlet algorithm. Then, Coulombic interactions were switched off (λ_{Coul}) over 20 steps ($\delta\lambda_{\text{Coul}} = -0.05$). Once the Coulombic term was completely decoupled, the LJ potential was decoupled in the same way over the remaining windows. During this process, intermolecular non-bonded interactions were also affected, so another decoupling simulation was performed in the gas phase, and the desolvation free-energy was obtained as a difference between these two processes.

GROMACS simulations were performed at 9 \AA Coulombic cutoff, 9 \AA LJ cutoff and the particle-mesh Ewald (PME) solver [37]. The damping parameters for the Parrinello-Rahman [38] barostat and the Nose-Hoover thermostat at 2 ps, and the barostat compressibility were set at $4.46 \times 10^{-5} \text{ bar}^{-1}$, using the leap-frog integrator and 1 Nose-Hoover chain. The Coulombic and LJ interactions were decoupled over 14 steps. First, the Coulombic interactions were switched off (λ_{Coul}) over 4 steps ($\delta\lambda_{\text{Coul}} = -0.25$). Once the Coulombic term was completely decoupled, the LJ potential was decoupled, where (λ_{LJ}) decreased over 10 steps ($\delta\lambda_{\text{LJ}} = -0.1$).

For GROMACS simulations, the Alchemical Analysis tool was employed [39], capable of handling various free-energy methods, including BAR (Bennet acceptance ratio, [40]), MBAR (Multistate Bennett acceptance ratio, [41]) and TI (thermodynamic integration [42]), while LAMMPS simulations utilized a separate tool within the FEP package [43], with our own adaptations made to facilitate compatibility with the Alchemical Analysis tool.

For the LJ binary mixtures, the analyses of free energy perturbation method were done using the FEP-package on the LAMMPS software version 07-Aug-19 [31]. As in the molecular mixtures, the system was initially minimized, then the velocities were then set to the target temperatures. This was followed by a NPT equilibration step of 1000 reduced

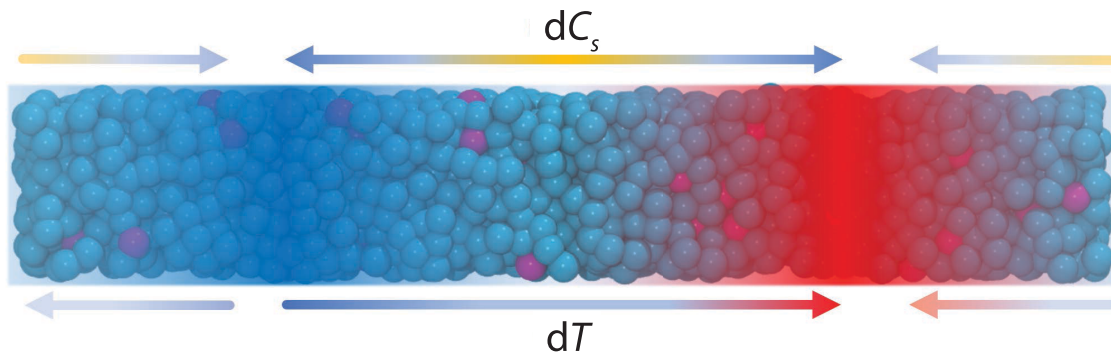


Figure 1. Schematic representation of a thermophoretic simulation (here, for LJ binary mixture). Cold and hot regions are shown in blue and red, respectively. They are positioned at one-fourth and three-fourths of the simulation box along its long z axis so as to generate the temperature gradient “ dT ” twice, thanks to periodic boundary conditions. This in turn generates a concentration gradient “ dC_s ” of solute particles (magenta) in the solvent (blue particles).

time units (2.180 ns) with $T^* = 1.0$ and $P^* = 0.8$. The damping parameters for the Nose–Hoover barostat and thermostat were $P_{\text{damp}}^* = T_{\text{damp}}^* = 1$ reduced unit time. The LJ potential in this method is modified by a soft core to avoid singularities while annihilating atoms [44]. The simulations were performed with a LJ cutoff (r_c) of 2.5σ (9.534 Å). Contrary to the method used for the molecular mixtures, there was no need of two decoupling steps, as there were no Coulombic interactions. The LJ potential was decoupled through the factor λ_{LJ} in 20 steps ($\delta\lambda_{\text{LJ}} = -0.05$). As in the aqueous solution systems, these simulations were performed without an induced temperature gradient. In both the aqueous solution systems and the LJ systems, the presented error bars for the desolvation energies are calculated from the standard deviation of multiple simulations with different starting configurations.

2.4. Mean square displacement calculations

A cubic box (33.15 Å-side) was considered for the molecular binary mixtures. The equilibration part is replicated from thermophoretic simulations [32], which are steps 1 to 8, in order to obtain a system in the microcanonical NVE ensemble with a volume and total energy corresponding to the averages in the NPT ensemble, such that the average pressure and temperature matched the targets of the chosen thermodynamic conditions. The diffusion coefficients were determined by measuring the mean squared

displacements (MSD) of the centers of mass for the solute and the solvent molecules/particles at the median temperature of 330 K. An equivalent protocol was followed for the LJ binary mixtures, with a starting simulation cubic box of size 25σ , at their median temperature of $T^* = 1.0$. As it is well established [45], the diffusion coefficient is expected to be slightly box size-dependent because of periodic boundary effects. However, assuming that the diffusion coefficient of these molecular species can be described by the Debye–Stokes–Einstein model, the size correction is temperature-independent [45], and therefore, the activation energy for diffusion is expected to be size-independent. For the determination of error bars, we used 5 replicas of 5 ns each, analyzed by blocks of 1 ns.

3. Results and discussions

3.1. Determination of Soret coefficients

We first study the thermophoretic behaviour of dilute aqueous solutions and of dilute LJ binary mixtures. In both cases, we systematically vary the solute nature and/or interaction parameters in order to simulate a range of different chemical properties for the solutes while keeping the solvent nature intact. Because these two types of solutions obviously have very different phase diagrams, we adapted the reference temperature, density and temperature gradients to remain in the liquid phase.

In silico thermophoretic settings were modelled as detailed in the Methods section. In short, we used a heat exchange algorithm [33] in the microcanonical ensemble, which pumps kinetic energy in one portion of the simulation box (which becomes colder than the average) and injects it in another region, which becomes the hot section of the system (Figure 1). In between, a temperature gradient is generated. The affected regions are chosen such that the gradient is unidirectional along one axis of the system (typically chosen as the z axis). In general, the temperature-gradient is established after a short equilibration timescale on the order of a few tens of picoseconds for the investigated systems [32]. A subsequent and much longer equilibration time [32] is then required to establish the concentration gradient, and only then is data acquired to determine the Soret coefficient. As shown previously, these timescales correspond to those of the heat and mass transport dynamics, respectively [32]. Simulations were replicated several times (typically 10 or more) in order to determine converged values and statistical uncertainties.

The Soret coefficient is determined after fitting the concentration profile with the following expression:

$$\frac{d \ln C_s}{dT} = -S_T \quad (3)$$

As shown before, under the conditions employed here, the log-scale plot of the solute concentration (defined here as its molality) is typically linear outside the thermostatted regions [32].

In Table 1, we show the values of the simulated Soret coefficients for several dilute aqueous solutions and binary LJ mixtures. The values are typically on the order of 10^{-3} – 10^{-2} K^{-1} . Although Soret coefficients are most often measured for polymer or hydrocarbon mixtures [8,46–48] or large colloidal particles [16,20,21,49], there are a few experimental measurements for aqueous solutions, including urea [11, 12] and ethanol [50–53] (which is expected to be very close in behavior to methanol) that suggest that the Soret coefficient of the solute in these mixtures is indeed close to the values found here. Similar observations can be made for LJ mixtures [15,54]. This comparison thus validates our approach and we thereby consider the simulated values as the reference that will now use to test several theoretical models.

Table 1. Molecular masses and Soret coefficients of molecular and LJ solutes

Solute	M ($\text{g}\cdot\text{mol}^{-1}$)	S_T (10^{-3} K^{-1})
Methanol	32.0	1.5 ± 0.7
Urea	60.1	4.4 ± 0.7
TMAO	75.1	5.0 ± 0.9
Glucose	180.2	3.1 ± 1.1
LJ Spheres ($\epsilon = 0.5$)	39.9	-7.6 ± 0.7
LJ Spheres ($\epsilon = 1.5$)	39.9	7.5 ± 2.9

Data for molecular solutes is reported at the median simulation temperature of 330 K in water, while that for LJ solutes is in a binary mixture of LJ particles with solvent particles of the same size and with $\epsilon = 1$ at $T^* = 1.0$.

3.2. Solvation free energy

In order to determine the solvation free energy at several different temperatures, we used standard approaches based on the progressive decoupling of the interactions between the target molecule and its environment. Simulations were performed by different simulation codes. For the aqueous solution simulations, as a natural choice, we first used the same software that was employed for thermophoretic simulations, but we also complemented our study with simulations performed with Gromacs [34]. For each approach, we compared three different algorithms in order to estimate the free energy cost associated with the process: TI, BAR and MBAR. Further details on these can be found in the Methods section.

We first take the example of a dilute aqueous solution of TMAO to compare the results from these different approaches (Figure 2A). Even when using the same raw data from the simulations, we notice that different algorithms give slightly different values, but the employed simulation codes give consistently similar answers. All these variations are typically within error bars, and lies within 0.4 kcal/mol from each other at 300 K and less than 1 kcal/mol at 330 and 360 K. Noticeably, they all quantitatively agree for the observed trend upon increasing temperature, with a ≈ 2 kcal/mol increase in the solvation free energy.

A second important step is to verify that our simulations agree with previous simulations and experimental data when available. Because we were

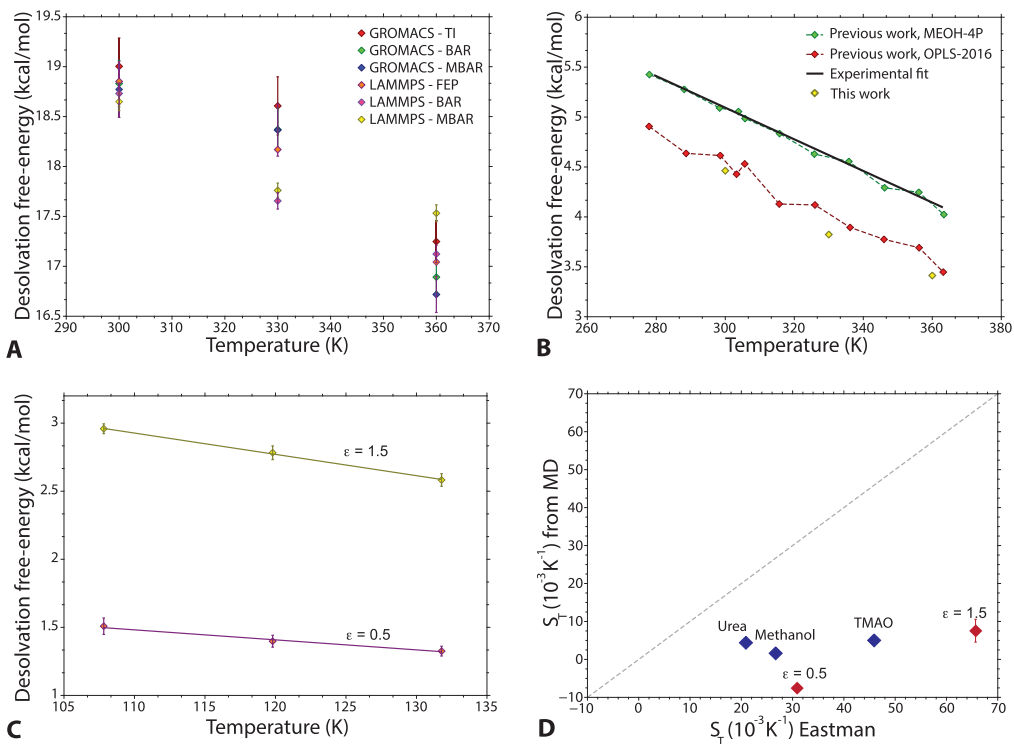


Figure 2. Solvation free energy approach. (A) Desolvation free energy of TMAO as a function of temperature in the dilute aqueous solution. Two different programs, as well as different algorithms, are compared. (B) Desolvation free energy of methanol in the dilute aqueous solution as a function of temperature, comparing results from this work (yellow) to those of previous ones (green and red, using two different forcefields) as well as the experimental estimate (black line). (C) Desolvation free energy for the solute particles as a function of temperature for the binary LJ mixtures. (D) Comparison between the direct estimate of the Soret coefficient in the investigated dilute solutions (vertical axis) and the predictions from the free energy model (horizontal axis).

lacking such data for TMAO, we do this comparison for methanol, using the LAMMPS code as the reference method (Figure 2B). Our simulations are in excellent agreement with previous simulation results using the OPLS force field for methanol [55], and are smaller in absolute value than the experimental values [56] (by about 0.5 kcal/mol). However, a very good agreement with experimental data was reported using a specifically reparametrized force field [55]. Despite these small differences, a critical aspect is that the experimental temperature dependence is very well reproduced by the previous simulations and the current ones. This provides a further validation of our approach.

We now turn to the determination of the temperature dependence of the solvation free energy for

a variety of solutes. Figure 2C shows the temperature dependence for LJ solutes as a function of temperature. In addition to TMAO and methanol, we also performed simulations on aqueous solutions of urea. Overall, we notice that the solvation free energy becomes favorable as temperature increases, i.e., the solvation entropy is negative. In aqueous solutions, this is typically what is observed in the experiments [56] and in previous simulation work [57–59], as there is an entropic cost to create a cavity accommodating the solute, and also a reduced entropy due to sometimes strong interactions between the water molecules and some of the solute chemical groups (for example, H-bond acceptors).

The positive slope of the free energy as temperature increases is thus compatible with the

thermophobic nature of these solutes; however, testing Equation (1) reveals that the amplitude of these variations with temperature (Figure 2D) would lead to Soret coefficients at least one order of magnitude higher to what is measured in the thermophoretic simulations. We do not observe strong variations of the results by changing the algorithm used to estimate the free energies, and similar results are obtained for a range of solutes. In addition, we checked that performing the simulations with two different programs gave the same qualitative answer. Therefore, such a model does not seem appropriate to explain thermodiffusion and the value of the Soret coefficient.

3.3. Temperature dependence of single-particle diffusion

We then examine, for the same solutions, the temperature dependence of the translational single-particle diffusion coefficient of both the solute and the solvent particles or molecules. As done for the solvation free energy calculations, we simulated the different systems under equilibrium conditions at several temperatures (see Methods). Diffusion coefficients were obtained from mean square displacement calculations. Their temperature dependence was very Arrhenius-like, enabling the determination of the corresponding activation energies to test the Equation (2) model.

Starting with the aqueous solutions, we find that the activation energy for the solute diffusion is generally higher than that of water. As typical solutes display positive Soret coefficients, i.e., they accumulate in the cold regions: this is consistent with the original idea of Prigogine's work that a higher activation energy for diffusion would imply that molecules would be more easily trapped in the cold regions [17]. For the investigated solutes, the differences in the diffusion activation energy with the solvent are on the order of RT , which gives rise to the right orders of magnitude for S_T . However, as evidenced in Figure 3, we fail to find any correlation between the magnitude of the difference between the activation energies for the diffusion of the solute and of the solvent and the Soret coefficient, and Equation (2) therefore does not appear to be valid for these solutions. Using the same methodology, we reach similar conclusions for the investigated LJ dilute mixtures (Fig-

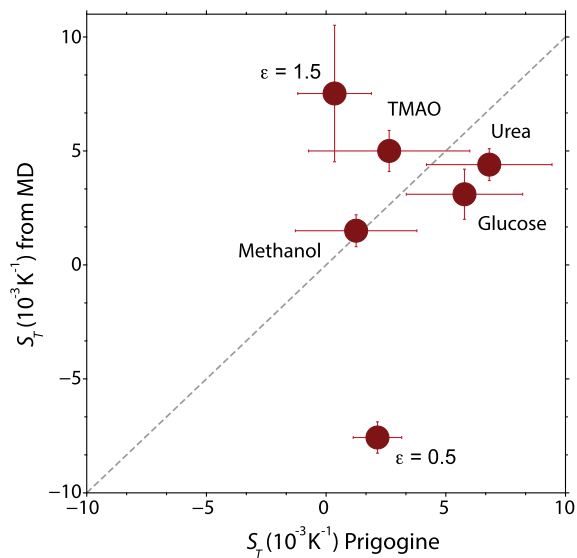


Figure 3. Activation energy for diffusion approach from Equation (2). Comparison between the direct estimate of the Soret coefficient in the investigated dilute solutions at $T = 330$ K (molecular solutes) and $T^* = 1$ (LJ solutions) (vertical axis) and the predictions from the activation energy for diffusion model (horizontal axis).

ure 3), albeit in obviously very different temperature conditions.

As suggested before, Prigogine's model is not able to explain the effects of mass on thermophoresis [18]. In particular, diffusion coefficients of a dilute particle in a given solvent are in principle mass-independent (note, however, that changing the solvent's mass could alter the solvent's viscosity and thus the diffusion coefficients of both the solute and the solvent). However, observations made from experimental and simulation measurements suggest that the mass-dependencies of thermophoresis follow phenomenological relationships in which one component of the Soret coefficient is proportional to the relative mass difference between the components of the binary mixtures [2]. The asymmetry between a change of mass in the solvent and a change of mass in the solute from the perspective of the diffusion coefficient (the first one having a small effect, the later one no effect) is a first hint that the Prigogine model would not be able to

explain the phenomenological mass-dependencies. Moreover, the variations of the diffusion coefficient temperature-dependence are too small to match the large effects of mass found before.

In order to account for the mass dependencies, a correction to the original Prigogine model has been suggested and tested on binary LJ mixtures [18]. The Soret coefficient of particle 2 in a mixture with 1 was written as [18]:

$$S_T = \frac{E_a^2 - E_a^1}{RT^2} + \frac{M^2 - M^1}{M^1 + M^2} \frac{E_a^2 + E_a^1}{RT^2} \quad (4)$$

(we note that Prigogine's original expression seems to be written with the wrong sign, but was taken as such when the derivation of this expression was made, we also corrected their expression when writing it here). This model was again tested for our solutions and largely overestimated the measured Soret coefficients. Indeed, for molecular solutes that are typically several times the mass of a water molecule, the mass correction simply becomes the sum of the activation energies divided by RT^2 . However, since the activation energies for diffusion are often several times the thermal energy, the correction typically amounts to several times $1/T$, whereas S_T is on the order of $1/T$ for the dilute molecular solutes studied here.

3.4. Correlation with short-term molecular motion

As we failed to quantitatively explain the Soret coefficients in dilute LJ mixtures and aqueous solutions with some of the previous theoretical equilibrium models, we tried to investigate possible correlations between S_T and other equilibrium observables at a given temperature. Keeping in mind Prigogine's conceptual ideas, i.e., that an imbalance in the temperature dependence of the molecular motion of molecules/particles would explain the onset of a concentration gradient, but seeking for quantities that would depend on mass, we focused on the initial regime of the molecular motion. For example, we looked at the average root mean squared displacement at a time interval corresponding to the onset of the diffusive regime after the initial ballistic region. In order to generate more data points, we (artificially) varied the solute masses and repeated the simulations for several solutes. As seen in Figure 4A, S_T is inversely correlated to the amplitude of

this motion, with molecules faster molecules associated with a lower S_T value.

By taking a different perspective, we can also look at the average time it takes for a molecule to travel a certain distance. Because this picture probably makes more sense for particles of similar sizes, we performed these measurements on simulations of binary LJ mixtures for which we systematically varied the masses and the interaction energies of the solute particles. As shown in Figure 4B, we again find a clear correlation between S_T and the timescale of a short-range initial motion covering 1σ . Interestingly, the crossover between positive and negative S_T values quantitatively agrees with the timescale of this initial motion being slower or faster as compared to that of the solvent particles. While a correlation does not imply that two quantities are directly connected to each other, we found that these observations, made here for very different types of dilute solutions, are encouraging and would deserve future investigations.

4. Conclusions

In this work, we investigate the validity of some of the models that have been proposed to explain thermophoresis based on equilibrium considerations, working here exclusively on dilute binary mixtures (where the dilute particle is considered as a solute and the concentrated one as the solvent). The first of these models is based on an idea first put forward by Eastman in the 1940s [14], and which has gained considerable attention in the last 15 years thanks to important results from the Braun group [16,20,21]. Based on experimental measurements on small colloidal particles, it has been suggested that the Soret coefficient could be explained in terms of the temperature dependence of the solute's solvation free energy. The second investigated model connects the Soret coefficient to the temperature dependence of the diffusion coefficients of the solute and of the solvent molecules and was originally suggested by Prigogine [17], with some later refinements [18].

We use molecular dynamics to compute both non-equilibrium properties (such as thermodiffusion) and equilibrium ones (such as translational dynamics and solvation free energies) and to compare them. As already shown in our own work on dilute aqueous solutions [32], as well as in previous contributions focusing on LJ particles mixtures [15,54],

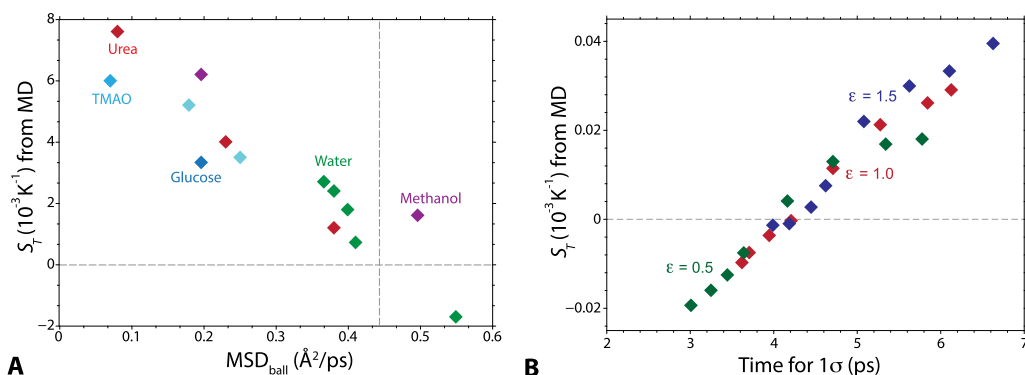


Figure 4. Correlation between the measured Soret coefficients for the investigated systems (varying both the chemical nature and the relative mass differences of the solute particles/molecules) and the short-term motion, seen from two different perspectives. (A) Correlation between S_T at a median temperature of 330 K with the mean square displacement of the solute at 330 K after 170 fs (corresponding to a change in sign in the second derivative of the MSD, indicative of a change of regime, that was observed to be almost system-independent for the dilute aqueous solutions shown here). Data for “water” solutes in water, with masses corresponding to 0.5, 1.5, 2.0, 4.0 and 8.0 times that of a regular water mass; for other solutes (except glucose, where the only data point corresponds to its natural mass): mass of a water molecule, regular mass, or 4 times the regular mass. (B) Soret coefficient at a median temperature of $T^* = 1.0$ vs the typical time to move by 1σ at the same temperature. These values were obtained through the MSD measurement of simulations as described in Section 2.4, tested over three values of ϵ for the solute (0.5, 1.0 and 1.5) and eight values of solute mass (0.33, 0.5, 0.75, 1.00, 2.00, 4.00, 6.00 and 8.00).

we first demonstrate how out-of-equilibrium simulations with an active heat exchange scheme can generate temperature gradients in molecular dynamics systems, which, in turn, leads to the establishment of concentration-gradients whose amplitude is compatible with experimental measurements.

We then run and analyse molecular dynamics trajectories on the same systems but in equilibrium conditions. By using different software and algorithms, we show that we can obtain reliable and converged estimates of the solute solvation free energies, and we examine their temperature dependence. Both for the aqueous solutions as well as for the LJ mixtures, we find that these cannot explain the Soret coefficients and that the Eastman model is not valid for the investigated molecular systems and predicts values that are one order of magnitude different from the steady-state measurements in non-equilibrium simulations. We cannot rule out that it would be correct for different systems, but we unambiguously show that this is not the case here. The same conclusions apply to Progogine’s model and its variant to include mass dependencies. We find that the Soret co-

efficient does not correlate with the imbalance in the activation energy for diffusion between solute and solvent molecules.

These conclusions could have perhaps been expected. Experiments as well as simulations have evidenced the critical mass effects on the amplitude of the concentration gradient [2]. However, free energy considerations, as well as single-particle diffusion properties, are in principle mass-independent. A decomposition of the Soret coefficient into different terms corresponding to the intrinsic so-called chemical contribution, mass contribution, moment of inertia component, or, as recently discussed, a mass dipole contribution, is an interesting phenomenological perspective but lacks theoretical grounds; hence, a model encompassing all these effects at once is still lacking. We finish our discussion with some interesting correlations we found between the asymmetry of the short-term motion of solvent and solute molecules and the amplitude of the Soret coefficient in these solutions, which we believe would deserve further investigations as they cover a wide range of chemistry and masses in very different

solutions. In particular, the effect of different molecular chemistries on water dynamical properties and H-bond exchange kinetics [60–63] could maybe be related to the thermal transport properties of the molecular solutes in aqueous solutions.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

Funding

The research leading to these results has received funding from the European Research Council under the European Union's Eighth Framework Program (H2020/2014-2020)/ERC Grant Agreement No. 757111 (GS). This work was also supported by the "Initiative d'Excellence" program from the French State (Grant "DYNAMO", ANR-11-LABX-0011-01 to GS).

Acknowledgments

The simulations presented here benefited from a local computing platform administered by G. Letessier, and benefited from the HPC resources of TGCC under the allocation A0070811005 made by GENCI (Grand Equipement National de Calcul Intensif).

References

- [1] M. A. Rahman, M. Z. Saghir, *Int. J. Heat Mass Transf.*, 2014, **73**, 693-705.
- [2] W. Köhler, K. I. Morozov, *J. Non-Equil. Thermodyn.*, 2016, **41**, 151-197.
- [3] D. Niether, S. Wiegand, *J. Phys. Condens. Matter*, 2019, **31**, article no. 503003.
- [4] W. Köhler, A. Mialdun, M. M. Bou-Ali, V. Shevtsova, *Int. J. Thermophys.*, 2023, **44**, article no. 140.
- [5] N. E. Zimmermann, G. Guevara-Carrion, J. Vrabec, N. Hansen, *Adv. Theory Simul.*, 2022, **5**, article no. 2200311.
- [6] O. R. Gittus, F. Bresme, *Phys. Chem. Chem. Phys.*, 2023, **25**, 1606-1611.
- [7] H. Hoang, G. Galliero, *Eur. Phys. J. E*, 2022, **45**, article no. 2200311.
- [8] W. M. Rutherford, *J. Chem. Phys.*, 1987, **86**, 5217-5218.
- [9] W. M. Rutherford, *J. Chem. Phys.*, 1989, **90**, 602-603.
- [10] O. R. Gittus, F. Bresme, *J. Chem. Phys.*, 2023, **159**, article no. 114503.
- [11] D. Niether, S. Di Lecce, F. Bresme, S. Wiegand, *Phys. Chem. Chem. Phys.*, 2018, **20**, 1012-1020.
- [12] D. Niether, H. Kriegs, J. K. Dhont, S. Wiegand, *J. Chem. Phys.*, 2018, **149**, article no. 044506.
- [13] A. Einstein, *Ann. Phys.*, 1905, **322**, 549-560.
- [14] E. D. Eastman, *J. Am. Chem. Soc.*, 1926, **48**, 1482-1493.
- [15] P.-A. Artola, B. Rousseau, *Mol. Phys.*, 2013, **111**, 3394-3403.
- [16] S. Duhr, D. Braun, *Phys. Rev. Lett.*, 2006, **96**, article no. 168301.
- [17] I. Prigogine, L. de Brouckere, R. Buess, *Physica*, 1952, **18**, 915-920.
- [18] P. A. Artola, B. Rousseau, G. Galliéro, *J. Am. Chem. Soc.*, 2008, **130**, 10963-10969.
- [19] A. Würger, *C. R. Mec.*, 2013, **341**, 438-448.
- [20] S. Duhr, D. Braun, *Proc. Natl. Acad. Sci. USA*, 2006, **103**, 19678-19682.
- [21] M. Reichl, M. Herzog, A. Götz, D. Braun, *Phys. Rev. Lett.*, 2014, **112**, article no. 198101.
- [22] L. Rowley, D. Nicholson, N. Parsonage, *J. Comput. Phys.*, 1975, **17**, 401-414.
- [23] K. M. Kast, J. Brickmann, S. M. Kast, R. S. Berry, *J. Phys. Chem. A*, 2003, **107**, 5342-5351.
- [24] C. Hölzl, P. Kibies, S. Imoto, R. Frach, S. Suladze, R. Winter, D. Marx, D. Horinek, S. M. Kast, *J. Chem. Phys.*, 2016, **144**.
- [25] S. Weerasinghe, P. E. Smith, *J. Phys. Chem. B*, 2003, **107**, 3891-3898.
- [26] D. Gonzalez-Salgado, C. Vega, *J. Chem. Phys.*, 2016, **145**, article no. 034508.
- [27] S. H. Jamali, T. van Westen, O. A. Moulto, T. J. H. Vlugt, *J. Chem. Theory Comput.*, 2018, **14**, 6690-6700.
- [28] J. L. Abascal, C. Vega, *J. Chem. Phys.*, 2005, **123**, article no. 234505.
- [29] L. Martínez, R. Andrade, E. G. Birgin, J. M. Martínez, *J. Comput. Chem.*, 2009, **30**, 2157-2164.
- [30] E. Polak, G. Ribiere, *Esaim Math. Model. Numer. Anal.*, 1969, **3**, 35-43, http://www.numdam.org/item/M2AN_1969__3_1_35_0/.
- [31] S. Plimpton, *J. Comput. Phys.*, 1995, **117**, 1-19.
- [32] A. Diaz-Marquez, G. Stirnemann, *J. Chem. Phys.*, 2021, **155**, article no. 174503.
- [33] P. Wirsberger, D. Frenkel, C. Dellago, *J. Chem. Phys.*, 2015, **143**, article no. 124104.
- [34] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *SoftwareX*, 2015, **1-2**, 19-25.
- [35] S. Nosé, *J. Chem. Phys.*, 1984, **81**, 511-519.
- [36] W. G. Hoover, *Phys. Rev. A*, 1985, **31**, 1695-1697.
- [37] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.
- [38] M. Parrinello, A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182-7190.
- [39] P. V. Klimovich, M. R. Shirts, D. L. Mobley, *J. Comput. Aided Mol.*, 2015, **29**, 397-411.
- [40] C. H. Bennett, *J. Comput. Phys.*, 1976, **22**, 245-268.
- [41] M. R. Shirts, J. D. Chodera, *J. Chem. Phys.*, 2008, **129**, article no. 124105.
- [42] D. Frenkel, B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 1st ed., Academic Press, Inc., USA, 1996.
- [43] R. W. Zwanzig, *J. Chem. Phys.*, 2004, **22**, 1420-1426.

- [44] T. C. Beutler, A. E. Mark, R. C. van Schaik, P. R. Gerber, W. F. van Gunsteren, *Chem. Phys. Lett.*, 1994, **222**, 529-539.
- [45] I.-C. Yeh, G. Hummer, *J. Phys. Chem. B*, 2004, **108**, 15873-15879.
- [46] D. Alonso De Mezquia, M. Mounir Bou-Ali, J. A. Madariaga, C. Santamaría, *J. Chem. Phys.*, 2014, **140**, article no. 084503.
- [47] G. Wittko, W. Köhler, *J. Chem. Phys.*, 2005, **123**, article no. 014506.
- [48] G. Wittko, W. Köhler, *Eur. Phys. J. E*, 2006, **21**, 283-291.
- [49] J. K. Dhont, S. Wiegand, S. Duhr, D. Braun, *Langmuir*, 2007, **23**, 1674-1683.
- [50] S. Wiegand, H. Ning, H. Kriegs, *J. Phys. Chem. B*, 2007, **111**, 14169-14174.
- [51] A. Königer, B. Meier, W. Köhler, *Phil. Mag.*, 2009, **89**, 907-923.
- [52] P. Kolodner, H. Williams, C. Moe, *J. Chem. Phys.*, 1988, **88**, 6512-6524.
- [53] K. J. Zhang, M. E. Briggs, R. W. Gammon, J. V. Sengers, *J. Chem. Phys.*, 1996, **104**, 6881-6892.
- [54] D. Reith, F. Müller-Plathe, *J. Chem. Phys.*, 2000, **112**, 2436-2443.
- [55] M. Martínez-Jiménez, H. Saint-Martin, *J. Chem. Theory Comput.*, 2018, **14**, 2526-2537.
- [56] J. Staudinger, P. V. Roberts, *Chemosphere*, 2001, **44**, 561-576.
- [57] D. L. Mobley, C. I. Bayly, M. D. Cooper, M. R. Shirts, K. A. Dill, *J. Chem. Theory Comput.*, 2009, **5**, 350-358.
- [58] M. M. Kubo, E. Gallicchio, R. M. Levy, *J. Phys. Chem. B*, 1997, **101**, 10527-10534.
- [59] E. Gallicchio, M. M. Kubo, R. M. Levy, *J. Phys. Chem. B*, 2000, **104**, 6271-6285.
- [60] D. Laage, G. Stirnemann, J. T. Hynes, *J. Phys. Chem. B*, 2009, **113**, 2428-2435.
- [61] F. Sterpone, G. Stirnemann, J. T. Hynes, D. Laage, *J. Phys. Chem. B*, 2010, **114**, 2083-2089.
- [62] G. Stirnemann, J. T. Hynes, D. Laage, *J. Phys. Chem. B*, 2010, **114**, 3052-3059.
- [63] G. Stirnemann, F. Sterpone, D. Laage, *J. Phys. Chem. B*, 2011, **115**, 3254-3262.