

# Mass effects for thermodiffusion in dilute aqueous solutions

Alejandro Diaz-Marquez and Guillaume Stirnemann\*

*CNRS Laboratoire de Biochimie Théorique, Institut de Biologie Physico-Chimique, PSL  
University, Université de Paris, 13 rue Pierre et Marie Curie, 75005, Paris, France*

E-mail: [stirnemann@ibpc.fr](mailto:stirnemann@ibpc.fr)

## Abstract

Thermodiffusion is the phenomenon by which molecules in a mixture present concentration gradients in response to an imposed temperature gradient. Despite decades of investigation, this effect remains poorly understood at a molecular level. A common, phenomenological approach is to individuate the molecular factors that influence the Soret coefficient, the parameter that quantifies the resulting concentration-gradient. Experimental studies, often performed on organic mixtures, as well as simulations of model particle systems have evidenced that the difference in masses between the mixture components has an important effect on the amplitude of the Soret coefficient. Here, we use molecular dynamics simulations of a thermophoretic setting to investigate the mass dependence of the Soret coefficient in dilute aqueous solutions. An advantage of simulation approaches is that they are not limited in the range of explored molecular masses, which is often limited to isotopic substitutions in the experiments. Our simulations reveal that the mass dependence of the Soret coefficient in these solutions is in agreement with previous experimental and simulation work on molecular-size systems. In particular, it is sensitive to the relative mass difference between the solute and the solvent, but not to their absolute mass. Adjusting the mass of the solvent and of the solute can turn a thermophobic solution into a thermophilic one, where solute accumulation is reversed. This demonstrates that the mass effect can indeed compensate for the other contributions to the Soret coefficient. Finally, we find that changing the molecular moments of inertia has a much more limited impact as compared to a change in the total molecular mass.

## Introduction

When solutions containing at least two constituents are subject to temperature gradients, they become inhomogeneous, i.e., the concentration of the different species are position-dependent. In liquid phases, this effect, usually referred to as thermodiffusion, manifests itself in many different contexts, from molecular mixtures up to micrometer-size particles in

water.<sup>1-3</sup>

It may seem surprising to the physical chemists that there is, as of today, no simple molecular model that is able to explain why thermodiffusion occurs, and no model to quantify the amplitude of the concentration-gradient resulting from a given temperature-gradient (the Soret coefficient, SC) in a system as simple as a binary liquid mixture of molecules A and B. As a consequence, it is in practice impossible to predict, for a given mixture under given thermodynamic conditions, the exact value of the expected SC.<sup>1,3,4</sup> This is particularly striking as the vast majority of the structural, dynamical and thermodynamical properties of the mixture are actually well understood and quantified by a variety of physical models. For example, the individual diffusion coefficients of components A and B are usually well predicted by the Stokes-Einstein equation, which states that the diffusion coefficient is determined by the temperature, the viscosity of the mixture and the hydrodynamic radius of this component. There is no such model for the Soret coefficient.

In order to get a better understanding of the SC, phenomenological descriptions have been used,<sup>1,3</sup> that provide interesting information regarding the origins of the SC. For example, a modification of the diffusion equation has been proposed to account for this additional phenomenon of "thermal diffusion", on top of the regular, Fickian diffusion. For a unidimensional temperature gradient along the  $z$  direction, and using the notations of refs.,<sup>5,6</sup> the total current of solute particles (concentration  $c_s(z)$ ) at a position  $z$  can be written as a sum of the diffusion current  $j_D(z)$  and of the thermal diffusion current  $j_{TD}(z)$ , such that

$$j^s(z) = j_D(z) + j_{TD}(z), \quad (1)$$

with

$$j_D(z) = -D^s(z)\nabla c_s|_z \quad (2)$$

and

$$j_{TD}(z) = -D_T^s(z)c_s(z)\nabla T. \quad (3)$$

Note here that these are written for dilute solutions of solvent molecules  $s$  in a solvent  $w$ , which will be the main focus of the present contribution. Equations are easily modified for concentrated binary mixtures. Other derivations consider the mass currents instead of the particle currents; for dilute solutions, both approaches are equivalent, and non-ideal effects can be neglected as well. We know from experimental and simulation results that a steady-state can be reached, which implies that  $j^s(z) = 0$  at each position. In the steady-state, we thus obtain

$$\left. \frac{d \ln c_s}{dT} \right|_z = -\frac{D_T^s(z)}{D^s(z)}, \quad (4)$$

or,

$$\left. \frac{d \ln c_s}{dT} \right|_x = -S_T^s(z), \quad (5)$$

where  $S_T^s(z) = D_T^s(z)/D^s(z)$  is defined as the Soret coefficient. Again, both experiments and simulations suggest that under "mild" gradient conditions,  $\ln c_s$  is often a linear function of  $T$ , i.e.,  $S_T$  can be considered locally independent of  $T$  and  $c_s$ .

As mentioned before, the Fickian, regular diffusion coefficient  $D^s(z)$  is rather well understood and easily predicted for any given solute in a mixture; however, less is known about the thermal diffusion coefficient  $D_T^s(z)$ . This kind of phenomenological approach does not really allow to understand the molecular bases for the SC.

Another strategy consists in trying to relate the measured values of the SC to the physical properties of the mixture constituents, for example, their molecular mass, their size, or the intermolecular interactions, among other factors.<sup>1,3,4</sup> An intriguing but noticeable feature of the recent experimental and simulations studies is that the corresponding contributions of each of these factors to the SC seem to be additive, at least within the framework of the phenomenological approach where one parameter is varied, the others being fixed. We will focus here on the so-called isotope effect, i.e., the dependence of the SC with the solute and solvent masses.

In order to clearly separate mass from "chemical" effects pertaining to the molecular

interactions in the experiments, the most direct approach is the use of nuclear isotopes, which is typically achieved by deuterating one of the mixture constituents. This mass effect has been recognized in early studies of thermodiffusion in mixtures of normal and heavy water,<sup>7</sup> and later, in mixtures of organic solvents.<sup>8-10</sup> Typically, experiments consider a given molecule A in a mixture with a second constituent B whose mass is altered through deuteration. Quite remarkably, the consequences of this mass change in terms of the SC change are pretty similar among series of different constituents A.<sup>11,12</sup> These experimental studies performed on a variety of solutions point toward a typical change of the SC by  $10^{-4} \text{ K}^{-1}$  for a change in mass of one of the constituent on the order of 1%.

However, deuteration only leads to limited changes in molecular mass. Isotopic substitutions of other nuclei would have even more limited effects. In order to increase the range of investigated masses, while keeping the chemistry unchanged, some authors have for example considered series of chemically analogous molecules, such as halobenzenes, varying the halogen substituent from fluorine to iodine,<sup>13</sup> or mixtures of linear alkanes.<sup>14</sup> Quite noticeably, the obtained data agrees with that obtained with isotope-substitution, with SC changes that are similarly correlated to the changes in mass. But such an approach should be considered with care, since chemical substitution necessarily affects the intermolecular interactions in addition to the mass.

At the other end of the spectrum, molecular dynamics simulation approaches have been trying to decompose the molecular factors that determine the SC. In particular, several seminal studies have highlighted mass effects in mixtures of non-polar simple particles.<sup>15-17</sup> On one hand, a great advantage of the simulations is that the amplitude of mass changes is in principle not limited to isotope-substitution: in the simulation, particle can carry any mass. Therefore, mass effects can easily be exaggerated.

On the other hand, simulations of realistic solutions, and especially of dilute mixtures, are challenging because of the timescales involved to reach the steady-state as well as poor statistics. To the best of our knowledge, mass effects have not been investigated in details

for these systems using MD simulations. The goal of the current contribution is to fill that gap. Here, we specifically focus on dilute aqueous solutions of several small molecular solutes. Building on the recent development of a reliable simulation strategy for these systems, we here investigate whether the simulation results follow the same trend as earlier experimental measurement on isotopically-substituted mixtures as well as simulation results on simple Lennard-Jones fluid mixtures. In particular, is there a "universal" mass-dependence for all solutes in a given solvent, or are there variations depending on the nature of the solute? We find an overall good agreement between our results and those of the previous experimental and simulation studies; in particular, we find that the SC is sensitive to the relative mass difference between the solute and the solvent, but not to their absolute mass. Interestingly, adjusting the mass of the solvent and of the solute can turn a thermophobic solution (where the solute prefers the cold side) into a thermophilic one, where solute accumulation is reversed. This demonstrates that the mass effect can indeed compensate for the other contributions to the SC. Finally, we specifically address the effect of the moment of inertia in the pure water case. We find that, in agreement with previous simulations on model particles, changing the molecular moments of inertia has a much more limited impact on the SC as compared to a change in mass.

## Methods and simulation strategy

Most technical details can be found in a previous work,<sup>18</sup> and only the main aspects are summarized here.

### In silico thermophoretic set-up

Each thermophoretic system consists in a parallelepipedic box of dimensions  $L_x \times L_y \times L_z$ , with  $L_x = L_y = 2.5$  nm and typically  $L_z = 2L_x = 5$  nm. To define a temperature gradient along  $z$ , hot and cold slabs of thickness  $\delta_z$  are symmetrically positioned at  $z_c = -L_z/4$  and

$z_h = L_z/4$ . Thus, the temperature-gradient is established over  $L_z/2$ , and is occurring twice in the box. The temperature-gradient is generated using an enhanced version of the heat-exchange algorithm (eHEX) as implemented in LAMMPS,<sup>19</sup> where, at given time intervals, heat  $-Q$  is "pumped" out of what is decided to be the cold region, and  $+Q$  is injected into the hot region. Heat exchange was performed at each simulation time-step, formally in the micro-canonical ensemble. For example, to impose a 60 K temperature difference, we used  $Q = 0.0375 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{fs}^{-1}$  for the simulations with a regular solvent water mass (18.015 amu). The data used to measure the steady-state behavior of a given aqueous solutions under a set thermal gradient is usually averaged over 20 independent trajectories.<sup>18</sup> After an equilibration procedure,<sup>18</sup> each trajectory is propagated for 10 ns in the micro-canonical ensemble and in the presence of the temperature-gradient. Overall, each SC is thus obtained from 200 ns of trajectory.

## Estimation of diffusion coefficients

We performed separated sets of simulations in order to determine the diffusion coefficients of the solute and solvent molecules. We used in that case cubic boxes with  $L_x = L_y = L_z = 3 \text{ nm}$ , that were first minimized and equilibrated for 5 ns in the NPT ensemble. 5,000 regularly spaced configurations were taken along this long initial run and later propagated for 20 ps each in the microcanonical ensemble. Note that such a strategy is probably much more than necessary for the only determination of diffusion coefficients, but we also used the same runs to analyze activation energy properties that impose this level of statistics, as shown before.<sup>20,21</sup>

## Simulation details and forcefields

We employed the LAMMPS package (version 07Aug19) to perform the simulations, starting from configurations created with Packmol.<sup>22</sup> The chose water forcefield was TIP4P-05.<sup>23</sup> The real space cutoff for Coulombic interactions was  $8.5 \text{ \AA}$ , and the Lennard-Jones cutoff

9 Å (which is larger as compared to the electrostatic cutoff because of the difference in the localization of the negative charge and of the mass of the oxygen atom). The simulation timestep was fixed to 1 fs. The solutes were described using state-of-the-art forcefields (TMAO,<sup>24,25</sup> urea<sup>26</sup> (bonded terms from<sup>27</sup>), methanol<sup>28</sup>). The mass of solvent or solute molecules was directly modified at the beginning of each simulation by rescaling accordingly the mass of all the corresponding atoms, with the only exception of masses smaller than 1 atomic mass unit (amu) that were avoided for numerical stability (see main text).

## Data analysis

We first define slabs along the  $x$  and  $y$  directions that are perpendicular to the direction of the gradient  $z$ . We typically divide  $L_z$  into 15 to 30 slabs. The more slabs, the smoother the data, but at the expense of increased statistical noise, which is problematic when studying properties with poor statistics (such as the solute distributions). We define molality in each slab (indistinctively referred to as concentration here) by dividing the average number of solutes centers of mass found in this slab by the average mass of water solvent molecules in the same slab. It is given in units of molal  $m$ , for moles of solutes by kilogram of solvent. Standard-deviations of the molality profiles were estimated from a set of 20 independent production runs. Soret coefficients (SC) were determined using Equation 5 on the average molality profile. Because individual trajectories lead to very noisy concentration profiles, the determination of the error bars based on the estimation of the SC for each separate trajectory would be meaningless and lead to unrealistic uncertainties. In that case, we therefore used bootstrapping made with 50 samples of 10 trajectories each one to estimate the statistical noise. Fickian diffusion coefficients were estimated by fitting the mean-square displacement of the molecules central atoms, passed an initial ballistic regime region that we took to be equal to 1 ps for all systems. Note that we did not correct for box size effects, which are likely to be very similar among all investigated systems.

# Results

## Rescaling of solute mass for fixed solvent mass

We start by discussing the results of thermophoretic simulations where the mass of the solute has been varied for fixed solvent water masses. As in the rest of this work, a 2-m concentration was chosen as to ensure enough statistics on a timescale of a few hundreds of nanoseconds.<sup>18</sup> While this clearly does not correspond to an infinitely-diluted regime, there are no direct interactions of the chosen solute molecules at this concentration (in particular, no aggregation) such that the solution behavior can be considered as that of a dilute system. Unless otherwise specified, all results discussed below were obtained using temperature differences of  $\Delta T = 60$  K around a median temperature of 330 K.

Increasing the solute mass above its regular value is computationally straightforward, and the mass of all solute atoms was uniformly increased up to 4 to 8 times the regular value. However, smaller masses may result in technical problems. In particular, decreasing the hydrogen atom masses below 1 atomic mass unit (amu) may cause numerical instabilities when using conventional simulation timesteps (1 fs here). As a consequence, when the solute mass was decreased below its reference, natural value, the hydrogen masses were always kept to  $\approx 1$  amu, and other nuclei were uniformly rescaled so that to reach the target molecular mass. Because of the interest in extracting the contribution to the Soret coefficient originating from effects others than that of the mass, an appealing reference is the situation where the solute and the solvent carry the same mass. Therefore, the lower value for the solute mass was set to 18.015 amu, i.e., the molar mass of the solvent water.

In Figure 1, we show the normalized molality profiles on a log-scale as a function of temperature for a dilute TMAO solution (a model amphiphilic solute that was extensively used before<sup>29-33</sup>), the slope of which is equal to  $-S_T$  (Equation 5). In that case, the mass of TMAO was varied between 18.015 amu (the molar mass of water) and 600.88 amu (8 times the molar mass of TMAO). As expected, the concentration gradient, and thus the SC

(inset), increase with the solute mass. However, the effect of a decrease in mass is stronger than that of a mass increase: from its reference value of  $75.11 \text{ g}\cdot\text{mol}^{-1}$ , decreasing the mass by a factor of 4 leads to a 36% decrease of the SC, while multiplying the mass by 4 leads to a 24% increase.

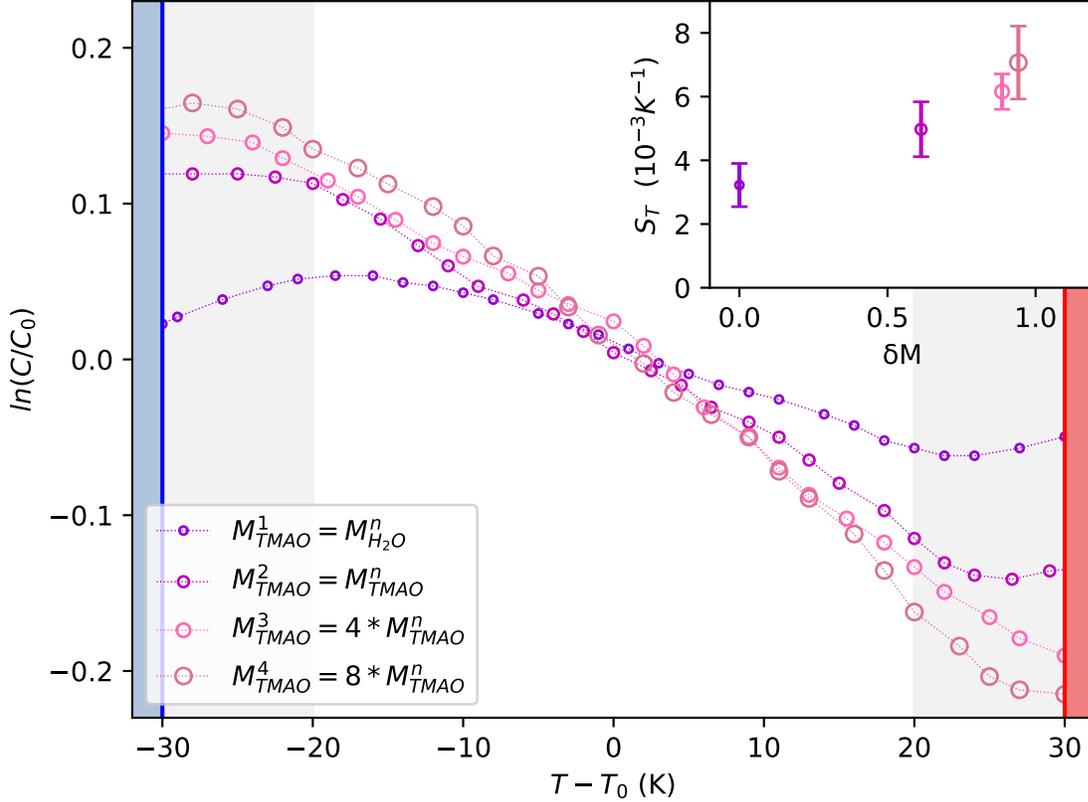


Figure 1: Concentration profiles as a function of temperature (with a reference  $T_0 = 330 \text{ K}$ ), for rescaled TMAO mass and fixed solvent mass ( $M^n$  designates the normal molecule mass). The blue and red bars represent the reservoirs where heat is pumped and injected, respectively. Inset graph: Soret coefficient ( $S_t$ ) vs relative mass difference ( $\delta M$ ). The Soret coefficient values are calculated using the data outside the grey regions.  $\delta M$  is the relative mass difference, equal to  $\frac{M_s - M_w}{M_s + M_w}$

Previous experimental and simulation studies typically found an empirical proportionality between the relative difference in mass between the mixture components (here,  $\delta M = \frac{M_s - M_w}{M_s + M_w}$  with  $M_s$  and  $M_w$  the masses of the solute and that of water) and the Soret coefficient.<sup>1</sup> In Figure 2, we then compare the SC dependence upon  $\delta M$  for TMAO, urea and methanol

solutions. In the last two cases, mass was rescaled following the procedure described in detail above for TMAO. The trend is similar to that observed for TMAO solutions, with an increase of the SC with the mass ratio between the solute and the solvent, but with two important differences. First, this dependence is actually more pronounced for urea and methanol, as seen in the steeper slope in this representation. Second, the lines cross the  $\delta M = 0$  vertical at different positions. This shows both that the SC is not zero even when the two component masses are identical, and that this value is sensitive to the chemical nature of the solute. We will comment on these aspects later in the discussion.

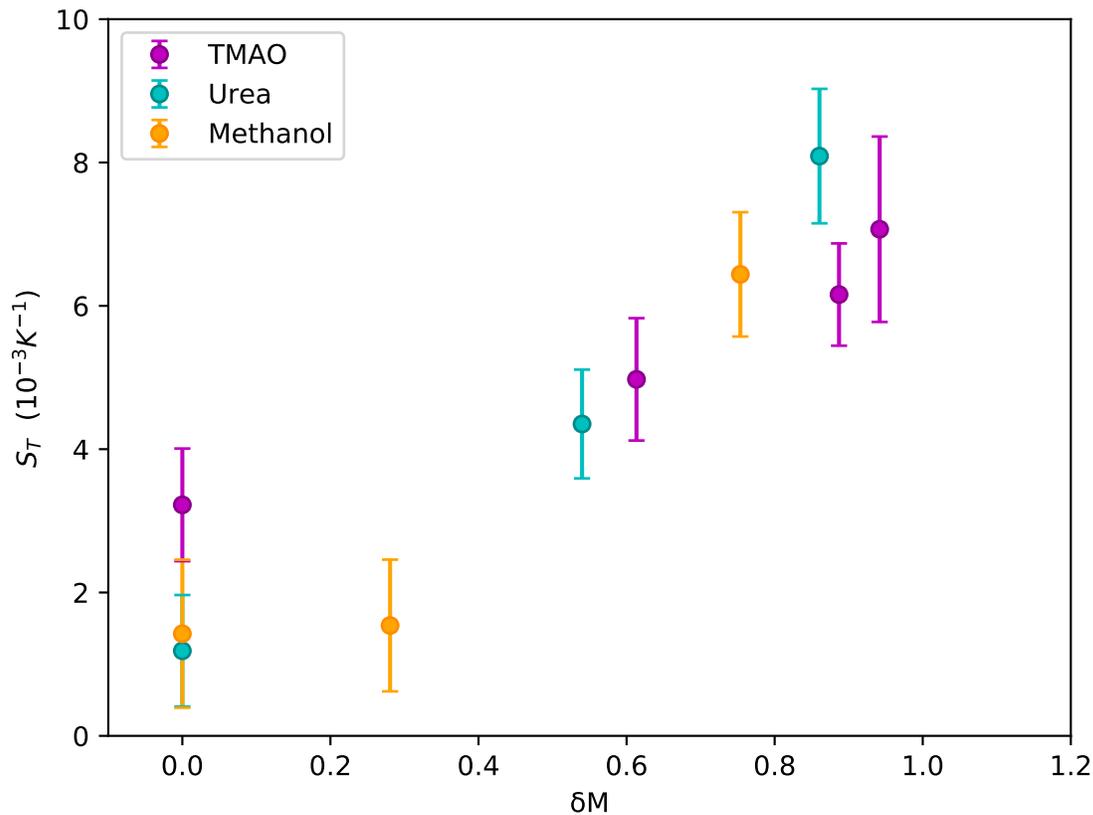


Figure 2: Soret coefficients as a function of the relative mass difference ( $\delta M$ ) between rescaled solutes and unrescaled solvent, for TMAO (magenta), urea (cyan) and methanol (orange) solutions. The error bars are calculated by the bootstrap method.

## Uniform rescaling of solute and solvent masses

The decomposition of Equation 6, or alternative ones, usually imply that the SC is sensitive to any change in the mass ratio between the mixture components, but not to their absolute masses. We now investigate whether this is verified in dilute aqueous solutions, using our model TMAO solution. In a new set of thermophoretic simulations, both the solvent and solute masses were rescaled by half following the procedure described above for the treatment of hydrogen atoms, or multiplied by 4. Whereas a change in the solute mass only does not result in any significant change in its diffusion timescale, rescaling the solvent mass leads to dramatic changes in the viscosity and thus to the diffusion timescale of all molecules (solvent and solute) in the system. This implies that the timescale required for establishing a concentration-gradient is longer, which requires to pay a specific attention to definition of the steady-state. In our case, the employed value of 2 ns is conservative enough so that a steady-state is established at this timescale even when solvent molecules are as heavy (and thus slowly diffusing) as the ones employed here.

As shown in Figure 3, the obtained concentration gradients and SC were the same within error bars, for the two different solutions. This suggests that, at least in the range of investigated masses, the SC is indeed sensitive to the asymmetry between the two component masses, and not to the individual mass of one component.

## Rescaling the solvent mass for a fixed solute mass

All the solutions studied so far are thermophobic, i.e., the solute molecules tend to accumulate in the cold region. Since the SC was observed to scale with the relative difference of masses between the solute and the solvent, solute molecule with low molecular masses exhibit SC values that are close to zero. This is for example the case for urea once rescaled to have the molecular mass of water, with a corresponding SC of  $1.2 \cdot 10^{-3} \text{ K}^{-1}$ . We now wonder whether negative values of the SC could be obtained, so that to render the solute thermophilic. For technical reasons, further scaling down the solute mass would be challeng-

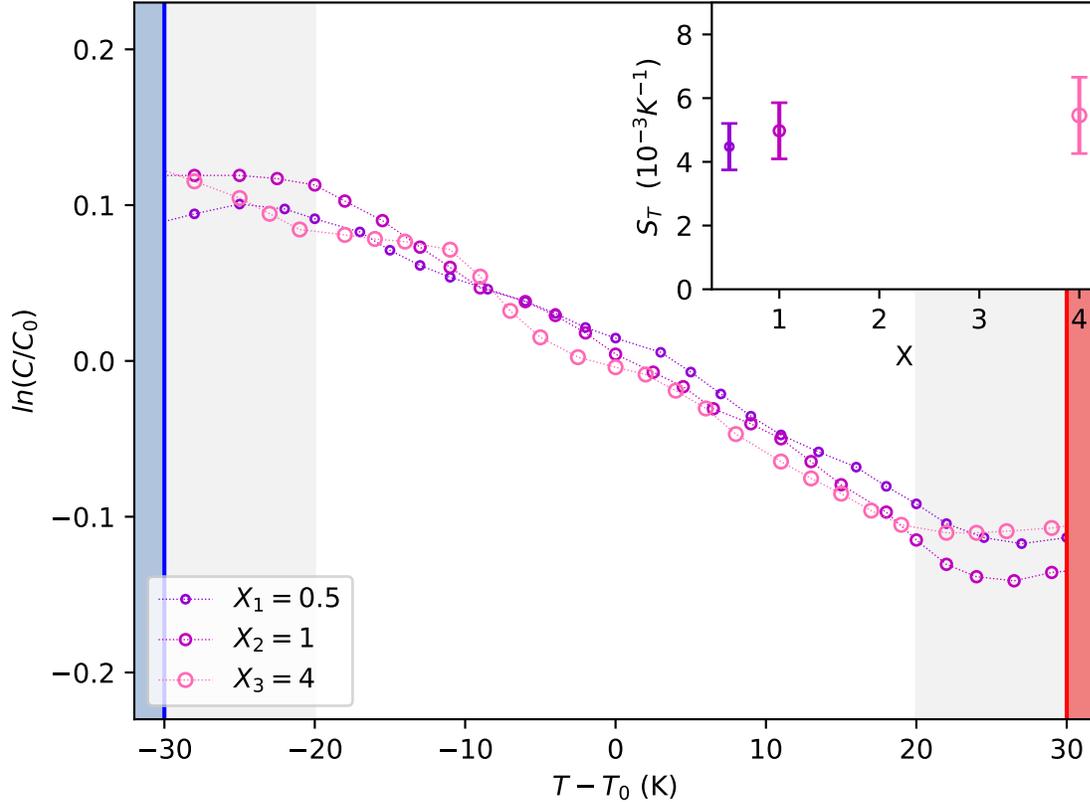


Figure 3: Concentration profiles as a function of temperature (with a reference  $T_0 = 330$  K) for uniformly rescaled TMAO and solvent masses by a factor  $X$ . The blue and red bars represent the reservoirs where heat is pumped and injected, respectively. Inset graph: Soret coefficient ( $S_t$ ) vs relative mass difference ( $\delta M$ ), the Soret coefficient values are calculated using the data outside the grey regions (same code color).

ing, with instabilities arising from the the existence of harmonic bonds between very light atoms. Instead, negative values of  $\delta M$  can be obtained by now increasing the mass of the solvent up to 8 times its natural mass.

The obtained concentration profiles (Figure 4) unambiguously demonstrate that the solution becomes thermophilic for very negative values of  $\delta M$ , and light urea in very heavy water thus tend to accumulate on the hot side, whereas regular urea in regular water is clearly thermophobic. The log-scale concentration profiles also exhibit significant deviations from linearity, which is only observed in these peculiar cases. For urea, the slope close to the

thermostatted regions can even switch sign, especially when the masses of urea and water are close and when the SC is close to zero. It is not clear whether this particular behavior is due to the simulation strategy where only the velocities of the solvent molecules are exchanged by the heat-exchange algorithm. As the solvent mass increases, we also note larger fluctuations among several replicas of the system, which leads to larger error bars in the SC. As detailed above, increasing the solvent mass directly affects the solution viscosity and the equilibration timescale, but not beyond values that are accessible by the current simulation setup, although they clearly lead to larger fluctuations. Another possibility for this behavior could be a temperature-dependence of the SC in this regime.

Since the SC corresponds to the slope of a linear fit of the concentration-gradient, the obtained value is sensitive to the width of the region that is chosen for the fit. We typically exclude regions close to the thermostats, but given the concentration profiles shown in Figure 4, changing the excluded zones will to some extent affect the SC, which already suffers from large fluctuations among the different replicas anyway. Therefore, the exact values should be taken with care; however, the trend for increasing solvent mass is unambiguous and when the solvent is much heavier than the solute, the solution is very clearly thermophilic.

In order to confirm these observations on a system where the solute mass had not been artificially reduced, the same approach was employed for the methanol solution, which readily exhibits a low SC value with unrescaled masses ( $1.5 \cdot 10^{-3} \text{ K}^{-1}$ ). The water mass was systematically increased and the resulting SC were estimated and compared to that of the urea solution (Figure 5), exhibiting qualitatively similar results. In particular, the solution becomes thermophilic for largely negative values of the relative mass difference between the solute and the solvent. These results thus suggest that the mass contribution can actually compensate for and overcome the chemical components, even for large interacting solutes that are naturally thermophobic.

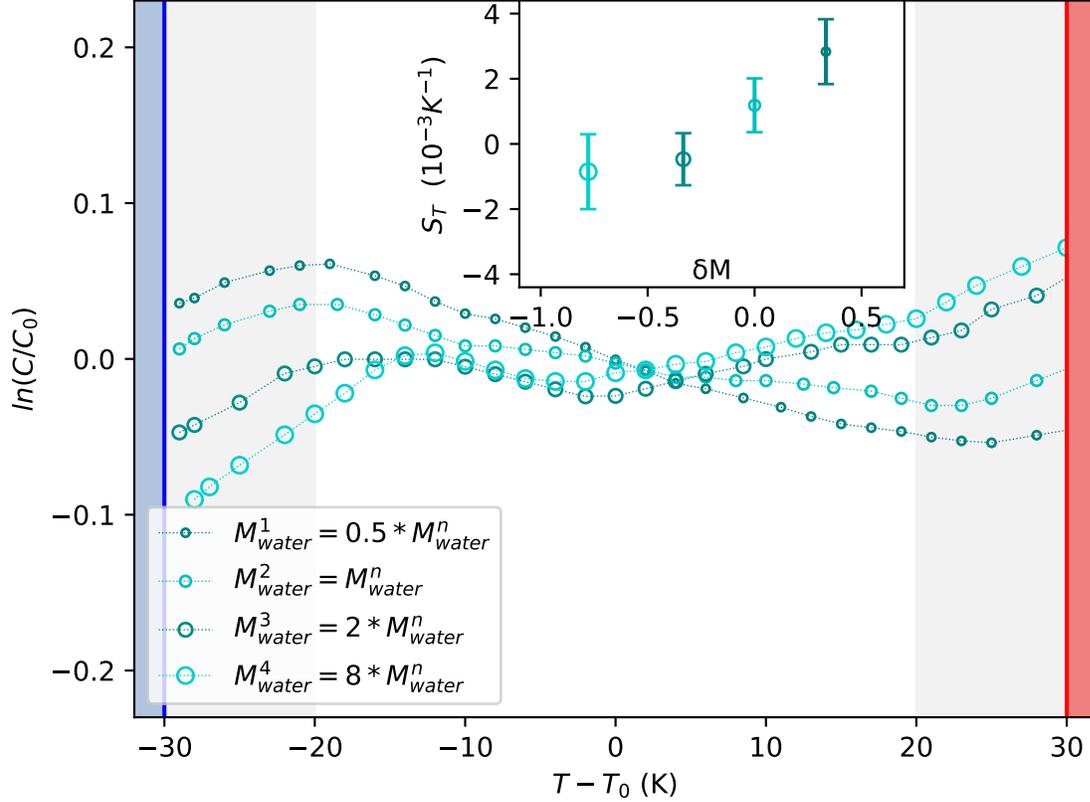


Figure 4: Concentration profiles as a function of temperature (with a reference  $T_0 = 330$  K), with rescaled solvent mass for fixed urea mass ( $M^n$  designates the normal molecule mass). The blue and red bars represent the reservoirs where heat is pumped and injected, respectively. Inset graph: Soret coefficient ( $S_t$ ) vs relative mass difference ( $\delta M$ ), the Soret coefficient values are calculated using the data outside the grey regions.

### Isotope substitution of water in water and moment of inertia effect

We then consider isotopic substitution in pure water. In that case, a fraction of the water molecules were made heavier or lighter as compared to the "solvent" water molecules, at concentrations similar to those employed for real solutes. By doing so, we can directly measure the effect of the differences in masses at fixed chemistry, i.e., for two molecules for which there would normally be no thermodiffusion.

As anticipated, when the solute water molecules are lighter than usual, they exhibit a negative SC (Figure 6), while the Soret coefficient is positive and proportional the relative

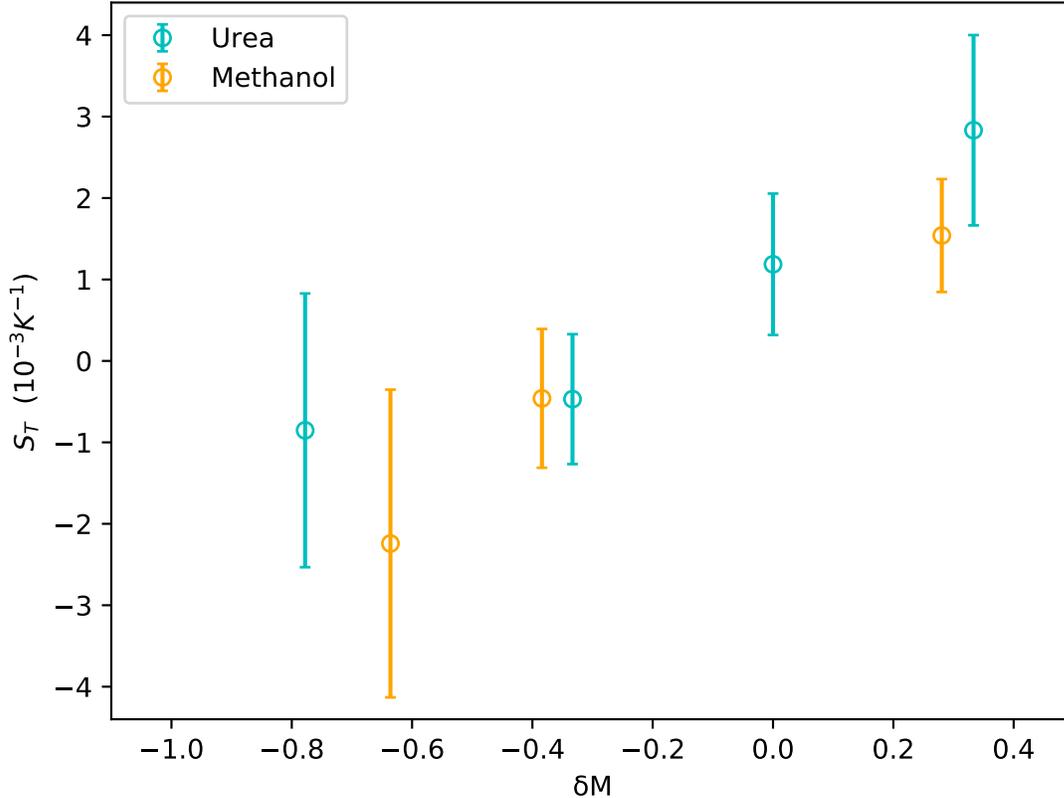


Figure 5: Soret coefficients as a function of the relative mass difference ( $\delta M$ ) for rescaled solvent mass for urea and methanol fixed mass solutions. Urea (cyan) and methanol (orange) solutions. The error bars are calculated by the bootstrap method.

increase in mass when solute water molecules are heavier than in bulk.

We use the same pure water system to investigate the effect of changes in molecular moments of inertia. Based on experiments performed on benzene derivatives,<sup>9-11</sup> the mass effect has been decomposed into an *intrinsic* mass contribution, due to changes in the total molecular mass, and a contribution stemming from a change in the moment of inertia, due to changes in the distribution of masses in the molecule. For water, we thus considered an unperturbed solvent and dilute "solute" molecules for which the oxygen mass was varied from 16 amu (regular mass) down the 2 amu, the total mass of 18.015 amu being constant, and the remaining mass being distributed homogeneously on the two hydrogen atoms. For

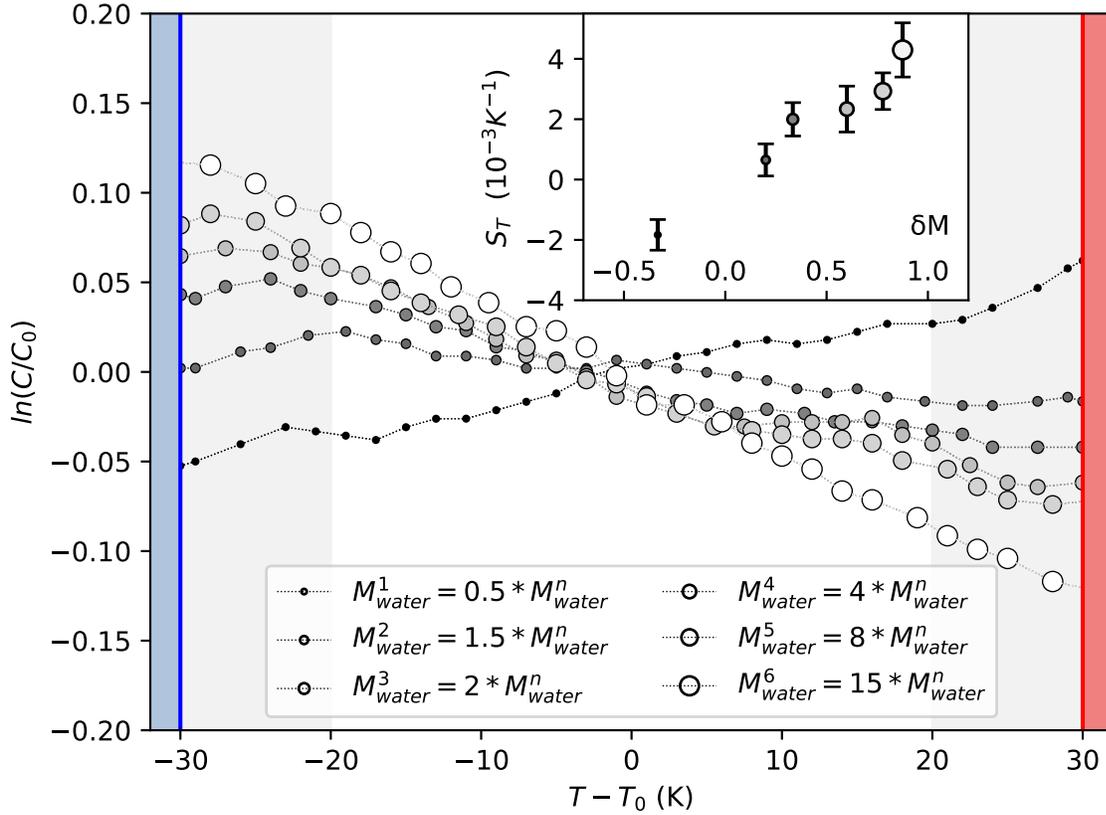


Figure 6: Concentration profiles as a function of temperature (with a reference  $T_0 = 330$  K) for rescaled water solute mass in a system of bulk (unscaled) water. The blue and red bars represent the reservoirs where heat is pumped and injected, respectively. Inset graph: Soret coefficient ( $S_t$ ) vs relative mass difference ( $\delta M$ ), the Soret coefficient values are calculated using the data outside the grey regions (same code color).

a water molecule, the moment of inertia has three  $I_{xx}$ ,  $I_{yy}$  and  $I_{zz}$  components, whose root mean squared sum was considered as the total, average moment of inertia. In the spirit of previous phenomenological approaches, the resulting Soret coefficient is plotted against  $\frac{I_s - I_w}{I_s + I_w}$ , where  $I_s$  is the moment of inertia of the solute water molecule and  $I_w$  that of the unscaled water solvent (Figure 7).

Scaling the mass of the water oxygen down results in an increase of all three components of the moment of inertia; yet, quite surprisingly, this leads to decreasing, negative values of the Soret coefficient, while previous experimental studies on very different organic mixtures,<sup>9-11</sup>

as well as simulations on model particles,<sup>16</sup> noticed an opposite trend. In our simulations, the average concentration profiles clearly point toward such negative SC values, but the overall effect remains very limited as compared to the total mass effects discussed so far, even on the same water system (Figure 6).

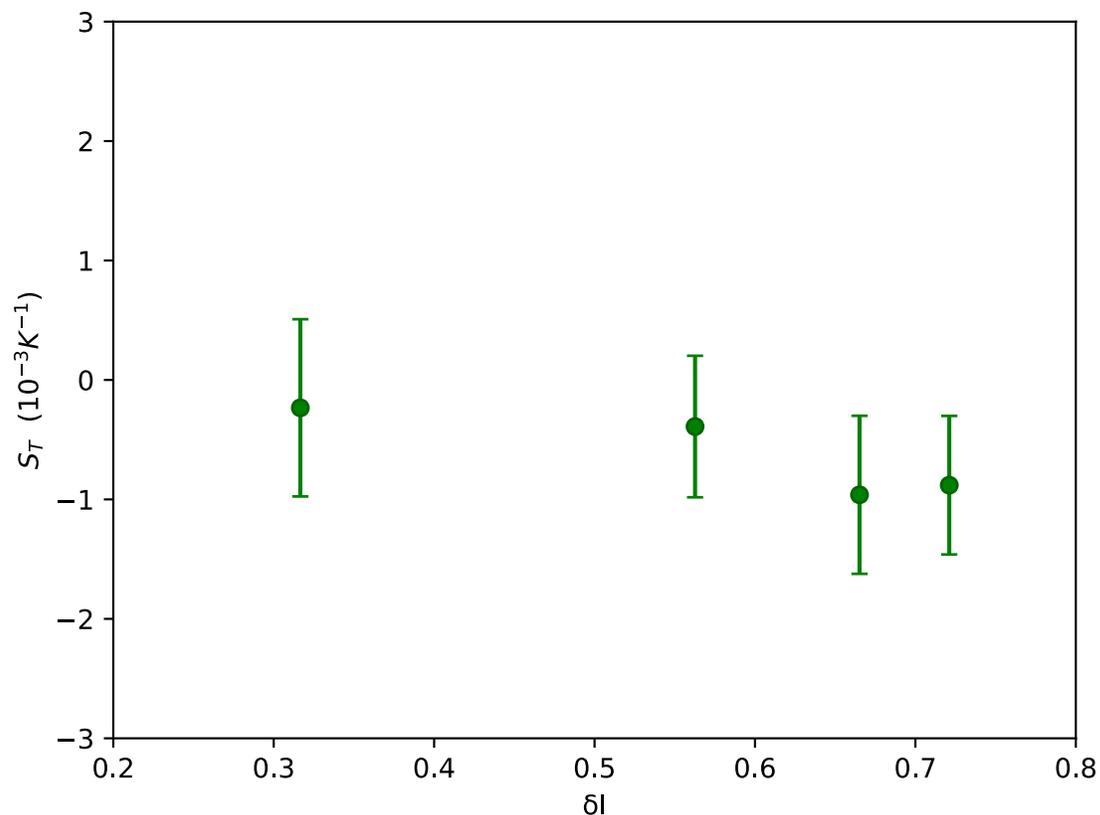


Figure 7: Soret coefficients for rescaled moment of inertia for a dilute fraction of water molecules in a pure water system, represented as a function of relative moments of inertia difference ( $\delta I$ ).

## Discussion and comparison with previous work

Here, we have systematically varied the mass of solute, the mass of solvent, or the mass of both solute and solvent, in dilute aqueous solutions. Changing the particle masses in the simulation is rather straightforward, and the range of covered relative mass changes  $\delta M$

is large compared to what can be achieved in the experiments, which are mostly limited to isotope substitutions. A plot encompassing all SC data discussed so far, varying the solute or solvent masses, is shown in Figure 8. The SC values are necessarily slightly noisy, but their dependence with  $\delta M$  can reasonably be fitted by a linear function

$$S_T = S_T^{iso} + S_M \delta M \quad (6)$$

whose values are given in Table 8. The fact that data corresponding either to the rescaling of the solute, of the solvent, or both, all align in such a representation, further validates the decomposition of the SC in terms of an isotope-independent contribution  $S_T^{iso}$  and a mass-dependency through a slope  $S_M$ , as extensively done before in the literature.

**Table 1: Constant values of the linear fit of Equation 6 from Figure 8 for the water, methanol, urea and TMAO molecules.**

Solute	$S_T^{iso}$ ( $10^{-3}\text{K}^{-1}$ )	$S_M$ ( $10^{-3}\text{K}^{-1}$ )	Coefficient of determination ( $R^2$ )
Water	0.0	4.5	0.96
Methanol	1.1	5.9	0.97
Urea	1.5	5.3	0.92
TMAO	3.2	3.4	0.94

The effect of mass is dictated by the amplitude of the slope  $S_M$ , which has been directly or indirectly measured in several experimental and simulation studies.<sup>1</sup> Experimentally, this can be achieved by several ways. In isotope-substitution studies, the SC can be determined in a liquid where a fraction of the molecules have been deuterated. This has been done on several molecular liquids, with reported data that ranges from  $S_M = 0.0013 \text{ K}^{-1}$  for water,<sup>7</sup> with a very limited concentration dependence,  $S_M = 0.007 \text{ K}^{-1}$  for equimolar mixtures of deuterated and regular methanol,<sup>8</sup> and typically  $S_M = 0.015\text{--}0.020 \text{ K}^{-1}$  for equimolar mixtures of benzene and several deuterated derivatives.<sup>8–10</sup> In slightly different experiments (still based on isotope substitution), the SC was measured in binary mixtures and compared for two different systems: first, a regular mixture of A and B; and second, a mixture of A

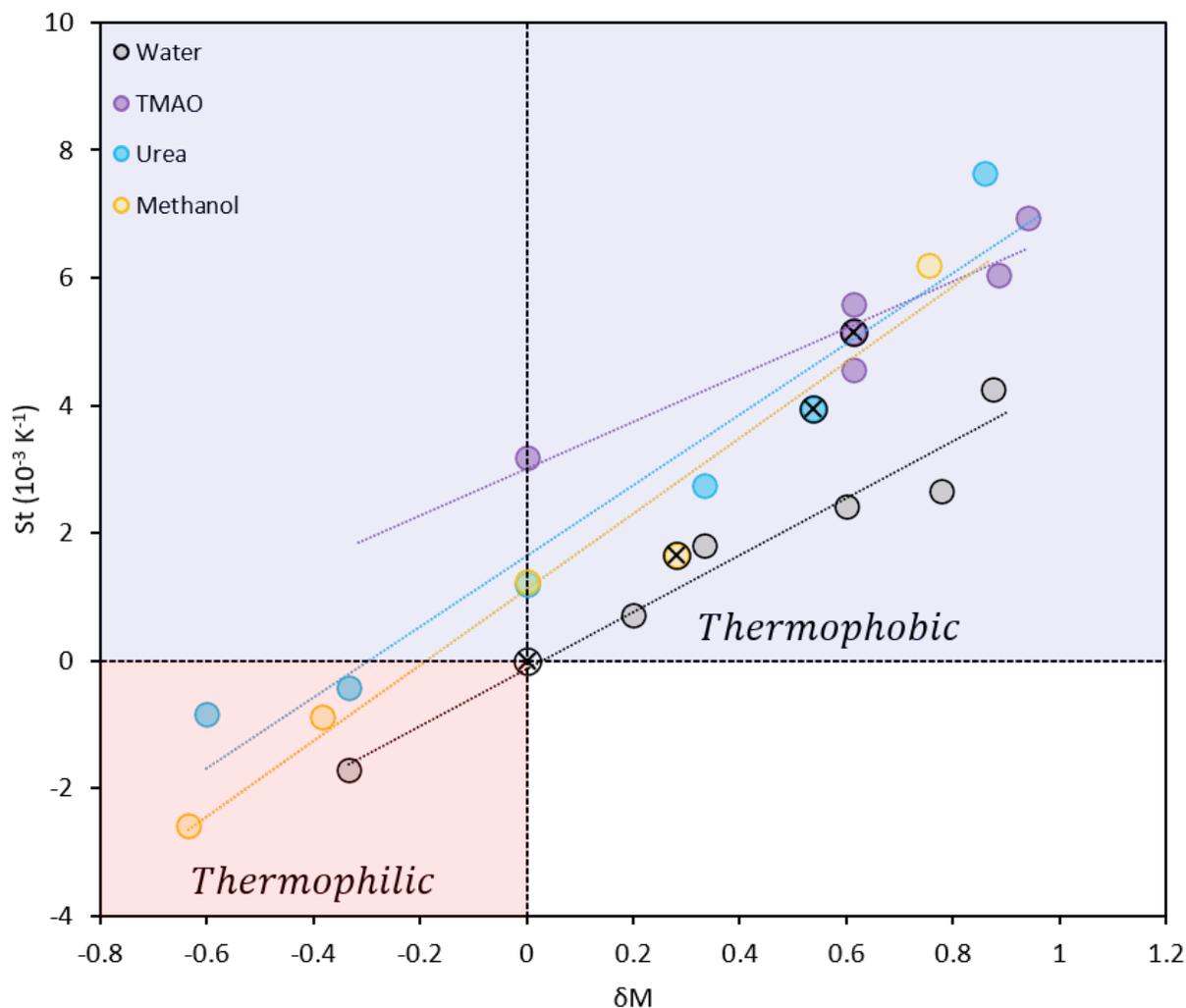


Figure 8: Summary of the mass effect on the Soret coefficient for dilute aqueous solutions, with methanol (yellow), ethanol (olive), urea (cyan), TMAO (magenta) and water (grey). The blue and red zones represent thermophobic and thermophilic regions. The cross dots are the simulations at natural mass for both solute and solvent. The dashed lines are linear regressions of the equation 6.

and deuterated B. Experiments performed on various organic mixtures, typically in regular or deuterated cyclohexane, led to  $S_M$  values that are little sensitive to the nature of the first component and typically on the order of  $0.016 \text{ K}^{-1}$ .<sup>11,12</sup> Such isotope effects have also been reported for aqueous solutions of ethanol ( $S_M = 0.019 \text{ K}^{-1}$ ) and DMSO ( $S_M = 0.012 \text{ K}^{-1}$ ).<sup>34</sup>

Isotope substitution can provide very valuable information, but it typically leads to moderate changes in the relative mass. Other experimental studies could enhance these changes

by considering chemical substitutions or additions that are assumed not to alter the molecular interaction and thus the chemical contribution to the SC. For example, mixtures of linear alkanes were investigated.<sup>14</sup> In that case, estimation of  $S_M$  can be achieved by varying the size of one component while keeping the other component fixed. Typically, this led to values around  $S_M = 0.01 \text{ K}^{-1}$ . Other studies focused on the series of monohalobenzenes in organic solvents,<sup>13</sup> leading to  $S_M = 0.017 \text{ K}^{-1}$ . Finally, simulations performed on uncharged, model particle mixtures could also isolate the mass contributions to the SC by systematically varying the mass of one of the component.<sup>15,16</sup> Surprisingly, the values reported on rather analogous systems range from  $S_M = 0.004 \text{ K}^{-1}$  in ref.<sup>16</sup> to  $S_M = 0.03 \text{ K}^{-1}$  in ref.<sup>15</sup>

The results of our simulations can be compared with those of these earlier measurements. Our values range from  $S_M = 0.004 \text{ K}^{-1}$  to  $S_M = 0.006 \text{ K}^{-1}$  depending on the solute. While they all fall in the same range, the slopes  $S_M$  thus slightly differ among the investigated systems.

We first note that these values are on the same order as those mentioned above. In particular, they are close to those for mixtures of deuterated methanol and methanol ( $S_M = 0.007 \text{ K}^{-1}$ ) as well as those obtained on model particles ( $S_M = 0.004 \text{ K}^{-1}$ ). They are 2–3 times below the values reported for larger molecular systems such as benzene derivatives or long alkanes in organic solvents. While there is no obvious trend indicating that  $S_M$  should increase with solute size, it is not clear whether an agreement with systems that are significantly different from dilute aqueous solutions should actually be expected. In particular, the system composition (the molar fraction of solute to solvent) and density are known to affect the SC values, and our current simulations lie in a regime that is pretty different from that explored in most experimental studies. A larger mismatch is observed with experiments performed on water and heavy water mixtures, which led to  $S_M$  values 3 times smaller than the one observed in our simulations,<sup>7</sup> or with those of ethanol in water, which exhibits values typically 3 times larger.<sup>34</sup> However, we note that experimental results themselves often lead to largely different values on similar systems, and the overall good

agreement on the amplitude of the mass effect in our simulation compared to this vast body of experimental and simulation results further reinforce our simulation approach.

We also note that our results are in agreement with an empirical correlation between  $S_M$  and the system density derived before based on MD simulations of model particles.<sup>35</sup> When rewriting Equation (2) of this reference, and rescaling the density by the critical point density, we find that the infinite dilution limit of  $S_M$  is expected to be  $0.0038 \text{ K}^{-1}$ , which is in very good agreement with our simulation results, although it does not explain why  $S_M$  slightly varies depending on the solute. This could be due to specific interactions between the solute and the solvent, not taken into account in such a simple model, or to effects due to the moments of inertia.

Indeed, inspired by theories on gaseous mixtures, it has been proposed to decompose the mass contribution of Equation 6 into three contributions and to write the following phenomenological equation:<sup>9-11</sup>

$$S_M = a_M \frac{M_s - M_w}{M_s + M_w} + b_M \frac{I_s - I_w}{I_s + I_w} \quad (7)$$

that is, to decompose the effect of mass in an intrinsic effect of the total molecular mass, and that of its moment of inertia. Such a decomposition was mainly motivated by studies on benzene derivatives with the same masses but slightly different thermophoretic response.<sup>9,10</sup> It has also been verified for mixtures of simple diatomic model molecules in the simulations.<sup>16</sup> In our case, changing the mass of all the atoms in a molecule, even homogeneously, will not only affect its molecular mass, but also the components of its moments of inertia. We thus note that the intrinsic effect of mass and that of the moment of inertia both contribute to the observed  $S_M$ , and they are not easily separated. In particular, since the moments of inertia of a molecule are not the same depending on the direction of its main axis for reorientation, it is not clear which component(s) would contribute to the moment of inertia in Equation 7. Originally, this formulation was proposed for molecules of much higher symmetry, such as

benzene, for which the definition of  $I$  is less ambiguous, and for which the specific ring geometry leads to pretty large values of  $I$ . A systematic simulation study indicated that the Soret coefficient was in fact much less sensitive to a change in the moments of inertia in the mixtures than it was to change in masses.<sup>16</sup> Our own investigation here on the pure water system indeed confirms that the moment of inertia effect is at least 3-4 times more moderate than that of the total molecular mass, with very limited changes in the SC that contrast with important changes with the total solute/solvent mass.

Finally, while the detailed investigation of the molecular origins for such a mass dependence of the Soret coefficient in the condensed phase goes beyond the scope of the current manuscript, we briefly comment on the connection between the effect of mass on  $S_T$  and on its diffusion and thermal diffusion,

$$S_T^s = D_T^s/D^s. \tag{8}$$

Not surprisingly, given the symmetry of Equation 6, the effect of a decrease in mass of the solute would be totally opposite to that of an increase in mass of the solvent. Yet, in dilute solutions, these have very distinct effects on the solute and solvent diffusion coefficients. Indeed, the diffusion coefficient of a dilute solute is not sensitive to a change in molecular mass, as expected from stochastic equations and as verified in our simulations. For example, the diffusion coefficient of bulk water at 298 K in a  $\approx 3$ -nm side cubic box was measured to be  $2.15 \times 10^{-5} \text{ cm}^2/\text{s}$ . When the mass of a dilute fraction (corresponding to a 2-m system) of water molecules is doubled, this results in a diffusion coefficient of both the "solute" and solvent water that is barely affected and equal to  $2.10 \times 10^{-5} \text{ cm}^2/\text{s}$ . However, now doubling the mass of solvent water molecules while leaving the solute one unaffected, leads to a diffusion coefficient of both types of molecules equal to  $0.67 \times 10^{-5} \text{ cm}^2/\text{s}$ , very close to the value of a pure bulk system with a doubled mass,  $0.62 \times 10^{-5} \text{ cm}^2/\text{s}$ . Despite these very different consequences on the diffusion coefficients, we find that indeed the decomposition of

Equation 6 can describe both cases, with  $S_T = 2.210^{-2} \text{ K}^{-1}$  with heavy solute in light solvent and  $S_T = -2.210^{-2} \text{ K}^{-1}$  for light solute in heavy solvent. Another striking observation is that when the solute mass is now multiplied by a factor of 8, the diffusion coefficient of both the solute and solvent water is barely affected and equal to  $1.75 \times 10^{-5} \text{ cm}^2/\text{s}$ . At the same time, the Soret coefficient is multiplied by a factor of 2 (Figure 6).

These observations thus suggest, on one hand, that upon changes in the mass of one of the components, very large variations can be observed for the SC with almost no variation of the diffusion coefficient of both species, and, on the other hand, that similar SC (in absolute value) can be measured while the diffusion coefficients dramatically vary. Following Equation 8, this means that the (phenomenological) thermal diffusion coefficient  $D_T$  would significantly vary in the first scenario, and totally compensate for the variations of  $D$  in the second one. Unfortunately, a molecular theory that would be able to directly predict the variations of  $D_T$  is currently not available to be compared with the results of this indirect estimation.

## Concluding remarks

Here, we have performed molecular dynamics simulations of dilute aqueous solutions under thermal gradients. Following our previous work that demonstrated the success of such a strategy for a variety of molecular solutes in water,<sup>18</sup> addressing in particular the challenge of dilute regimes that require very long simulation timescales, we investigate here the effect of mass on the amplitude of the concentration-gradient of the solute that is observed in response to the temperature-gradient. For different solutes, including heavy/light water, in normal water, the mass of the solute, the mass of the solvent, or both, were systematically varied. In a way, this echoes experimental isotope-substitution experiments, with the noticeable difference that simulations are not limited to naturally-accessible chemical isotopes. As a consequence, a much larger spectrum of mass differences can be explored.

Quite remarkably, we find that the SC typically varies linearly with the relative mass difference between the solute and the solvent, in agreement with previous experimental studies as well as computer simulations on simple model particles. However, a similar rescaling of both the solvent and the solute molecules does not lead to noticeable variations of the SC, as expected. The slope of the mass dependence upon the relative mass difference were seen to slightly vary among the solutes, but were always on the order of  $10^{-2} \text{ K}^{-1}$ . These values are in good agreement with available experimental and simulation results on a variety of solutions, including organic, aqueous, and model mixtures. This suggests that molecular dynamics simulations on real dilute aqueous systems are able to capture the mass dependence of thermodiffusion. We also found that differences in the moments of inertia of the solute and solvent molecules has a much more limited effect on the Soret coefficient as compared to changes in the total molecular mass. Finally, we show that the large variations of the Soret coefficient upon a change in mass of the solute necessarily come from large variations of its thermal diffusion component, because the diffusion coefficient is largely insensitive to the mass of dilute species. Provided with an approach and a set of simulations that reliably reproduce experimental data, we are currently working on the molecular interpretation of these mass effects.

## Acknowledgement

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 (G.S.). This work was also supported by the "Initiative d'Excellence" program from the French State (Grant "DYNAMO", ANR-11-LABX-0011-01 to GS). The simulations presented here benefited from a local computing platform administered by G. Letessier, and was granted access to the HPC resources of TGCC under the allocation A0070811005 made by GENCI (Grand Equipement National de Calcul Intensif).

## Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## References

- (1) Köhler, W.; Morozov, K. I. The Soret Effect in Liquid Mixtures - A Review. *J. Non-Equil. Thermodyn.* **2016**, *41*, 151–197.
- (2) Rahman, M. A.; Saghir, M. Z. Thermodiffusion or Soret effect: Historical review. *Int. J. Heat Mass Transf.* **2014**, *73*, 693–705.
- (3) Niether, D.; Wiegand, S. Thermophoresis of biological and biocompatible compounds in aqueous solution. *J. Phys.: Condens. Matter* **2019**, *31*, 503003.
- (4) Piazza, R. Thermophoresis: moving particles with thermal gradients. *Soft Matter* **2008**, *4*, 1740.
- (5) Duhr, S.; Braun, D. Thermophoretic depletion follows Boltzmann distribution. *Phys. Rev. Lett.* **2006**, *96*, 168301.
- (6) Würger, A. Is Soret equilibrium a non-equilibrium effect? *C. R. - Mec.* **2013**, *341*, 438–448.
- (7) Prigogine, I.; de Brouckere, L.; Buess, R. Recherches Sur La Thermodiffusion En Phase Liquide. Thermodiffusion De L'eau Lourde. *Physica* **1952**, *18*, 915–920.
- (8) Ning-Yuan, R. M.; Beyerlein, A. L. The  $2H$  thermal diffusion isotope effect in benzene and methanol. *J. Chem. Phys.* **1983**, *78*, 7010–7011.
- (9) Rutherford, W. M. Effect of mass distribution on the isotopic thermal diffusion of benzene. *J. Chem. Phys.* **1987**, *86*, 5217–5218.

- (10) Rutherford, W. M. Effect of carbon and hydrogen isotopic substitutions on the thermal diffusion of benzene. *J. Chem. Phys.* **1989**, *90*, 602–603.
- (11) Debuschewitz, C.; Köhler, W. Molecular origin of thermal diffusion in benzene+cyclohexane mixtures. *Phys. Rev. Lett.* **2001**, *87*, 55901.
- (12) Wittko, G.; Köhler, W. Universal isotope effect in thermal diffusion of mixtures containing cyclohexane and cyclohexane- d 12. *J. Chem. Phys.* **2005**, *123*, 014506.
- (13) Hartmann, S.; Köhler, W.; Morozov, K. I. The isotope Soret effect in molecular liquids: A quantum effect at room temperatures. *Soft Matt.* **2012**, *8*, 1355–1360.
- (14) Alonso De Mezquia, D.; Mounir Bou-Ali, M.; Madariaga, J. A.; Santamaría, C. Mass effect on the Soret coefficient in n -alkane mixtures. *J. Chem. Phys.* **2014**, *140*, 084503.
- (15) Reith, D.; Müller-Plathe, F. On the nature of thermal diffusion in binary Lennard-Jones liquids. *J. Chem. Phys.* **2000**, *112*, 2436.
- (16) Galliéro, G.; Duguay, B.; Caltagirone, J. P.; Montel, F. Thermal diffusion sensitivity to the molecular parameters of a binary equimolar mixture, a non-equilibrium molecular dynamics approach. *Fluid Ph. Equilibria* **2003**, *208*, 171–188.
- (17) Artola, P. A.; Rousseau, B. Microscopic interpretation of a pure chemical contribution to the soret effect. *Phys. Rev. Lett.* **2007**, *98*, 125901.
- (18) Diaz-Marquez, A.; Stirnemann, G. In silico all-atom approach to thermodiffusion in dilute aqueous solutions. *J. Chem. Phys.* **2021**, *155*, 174503.
- (19) Wirnsberger, P.; Frenkel, D.; Dellago, C. An enhanced version of the heat exchange algorithm with excellent energy conservation properties. *J. Chem. Phys.* **2015**, *143*, 124104.
- (20) Piskulich, Z. A.; Laage, D.; Thompson, W. H. Using Activation Energies to Elucidate Mechanisms of Water Dynamics. *J. Phys. Chem. A* **2021**, *125*, 9941–9952.

- (21) Piskulich, Z. A.; Mesele, O. O.; Thompson, W. H. Removing the barrier to the calculation of activation energies: Diffusion coefficients and reorientation times in liquid water. *J. Chem. Phys.* **2017**, *147*, 1–6.
- (22) Martinez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* **2009**, *30*, 2157–2164.
- (23) Abascal, J. L.; Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*.
- (24) Kast, K. M.; Brickmann, J.; Kast, S. M.; Berry, R. S. Binary Phases of Aliphatic N-Oxides and Water: Force Field Development and Molecular Dynamics Simulation. *J. Phys. Chem. A* **2003**, *107*, 5342–5351.
- (25) Hölzl, C.; Kibies, P.; Imoto, S.; Frach, R.; Suladze, S.; Winter, R.; Marx, D.; Horinek, D.; Kast, S. M. Design principles for high-pressure force fields: Aqueous TMAO solutions from ambient to kilobar pressures. *J. Chem. Phys.* **2016**, *144*, 144104.
- (26) Weerasinghe, S.; Smith, P. E. A Kirkwood-Buff Derived Force Field for Mixtures of Urea and Water. *J. Phys. Chem. B* **2003**, *107*, 3891–3898.
- (27) Chiba, S.; Furuta, T.; Shimizu, S. Kirkwood–Buff Integrals for Aqueous Urea Solutions Based upon the Quantum Chemical Electrostatic Potential and Interaction Energies. *J. Phys. Chem. B* **2016**, *120*, 7714–7723.
- (28) Gonzalez-Salgado, D.; Vega, C. A new intermolecular potential for simulations of methanol: The OPLS/2016 model. *J. Chem. Phys.* **2016**, *145*, 034508.
- (29) Stirnemann, G.; Duboué-Dijon, E.; Laage, D. Ab Initio Simulations of Water Dynamics in Aqueous TMAO Solutions: Temperature and Concentration Effects. *J. Phys. Chem. B* **2017**, *121*, 11189–11197.

- (30) Mondal, J.; Stirnemann, G.; Berne, B. J. When Does Trimethylamine N-Oxide Fold a Polymer Chain and Urea Unfold It? *J. Phys. Chem. B* **2013**, *117*, 8723–8732.
- (31) Canchi, D. R.; García, A. E. Cosolvent effects on protein stability. *Annu. Rev. Phys. Chem.* **2013**, *64*, 273–293.
- (32) Stirnemann, G.; Sterpone, F.; Laage, D. Dynamics of water in concentrated solutions of amphiphiles: Key roles of local structure and aggregation. *J. Phys. Chem. B* **2011**, *115*, 3254–3262.
- (33) Stirnemann, G.; Hynes, J. T.; Laage, D. Water hydrogen bond dynamics in aqueous solutions of amphiphiles. *J. Phys. Chem. B* **2010**, *114*, 3052–3059.
- (34) Niether, D.; Kriegs, H.; Dhont, J. K.; Wiegand, S. Peptide model systems: Correlation between thermophilicity and hydrophilicity. *J. Chem. Phys.* **2018**, *149*, 044506.
- (35) Galliero, G.; Bugel, M.; Duguay, B.; Montel, F. Mass effect on thermodiffusion using molecular dynamics. *J. Non-Equilib. Thermodyn.* **2007**, *32*, 251–258.