

Peptide model systems: Correlation between thermophilicity and hydrophilicity

Doreen Niether, Hartmut Kriegs, Jan K.G. Dhont,^{*} and Simone Wiegand[†]

ICS-3 Soft Condensed Matter, Forschungszentrum Jülich GmbH, D-52428 Jülich, Germany

(Dated: July 11, 2018)

Abstract

In recent years, the response of biomolecules to a temperature gradient has been utilized to monitor reactions of biomolecules, but the underlying mechanism is not well understood, due to the complexity of the multicomponent system. To identify some underlying principles, we investigate the thermal diffusion of small amide molecules in water systematically. We re-analyse previous measurements of urea and formamide and compare the results with acetamide, *N*-methylformamide, and *N,N*-dimethylformamide, amides with a lower hydrophilicity. It turns out that the less hydrophilic substances do not show the typical temperature dependence of water soluble macromolecules. Analysing temperature and concentration dependent measurements using an empirical expression originally derived for nonpolar mixtures, we find that the so-called isotope contribution depends strongly on the hydrophilicity of the solute. This can be qualitatively understood by comparing with molecular dynamic simulations of Lennard-Jones fluids. The hydrophobic/hydrophilic balance also influences the structure in the fluid and with that the thermal expansion coefficient, which correlates with the thermal diffusion coefficient. Further, we observe a clear correlation of the temperature and concentration dependence of the Soret coefficient with the hydrophilicity, which can be quantitatively described by the partition coefficient $\log P$.

^{*} Also at Department of Physics, Heinrich-Heine-Universität Düsseldorf, D-40225 Düsseldorf, Germany

[†] Also at Department für Chemie - Physikalische Chemie, Universität zu Köln, 50939 Cologne, Germany

I. INTRODUCTION

In recent years, MicroScale Thermophoresis (MST) has gained popularity as a technique to quantify biomolecular interactions. It is based on the thermophoretic behaviour of biomolecules and its sensitivity to non-covalent binding [1]. While the effect can be used to precisely determine the dissociation constant K_d , which gives access to the change in Gibbs free energy ΔG^0 , the reasons for the change in thermophoretic behaviour of the biomolecules upon binding are not yet understood. For a systematic study of this effect protein systems are rather complicated: they require specific conditions (pH, buffer) that can differ from system to system and often undergo a folding process while the binding takes place so that the chemical groups that are in contact with the solvent change. Therefore, we are looking for smaller building blocks of proteins such as amides, which often serve as model substances for peptide bonds, to identify some underlying principles. An additional advantage is the fact that the molecules are uncharged and allow a systematic variation of the hydrophilicity. While the hydrophilic amides form strong hydrogen bonds with the water molecules the more hydrophobic amides lead to the spontaneous formation of empty water molecule cavities capable of accommodating hydrophobic functional groups [2]. One main objective of this work is to find out how the change in hydrophilicity influences the thermodiffusive behaviour.

While thermophoresis in fluids is not yet fully explained on a microscopic level, the theoretical understanding of non-polar binary systems has become quite comprehensive in recent years [3]. The reason why the thermodiffusion behaviour of a liquid mixture is so difficult to predict is its sensitivity to a large number of parameters [4]. For ionic particles, the thermophoretic behaviour is strongly influenced by the charge, so that the Soret coefficient becomes a function of the Debye screening length [5, 6] and can also be expressed in dependence of the surface potential [7]. For non-polar systems it has been shown that the Soret coefficient depends on mass and moment of inertia, as well as a chemical contribution that is due to non-specific interactions between solute and solvent [8, 9].

In aqueous and therefore most biological systems these interactions are dominated by hydrogen bonds. For the temperature dependence of the Soret coefficient S_T in aqueous

solutions an empirical equation was proposed by Iacopini and Piazza [10],

$$S_T(T) = S_T^\infty \left[1 - \exp\left(\frac{T^* - T}{T_0}\right) \right] . \quad (1)$$

with the adjustable parameters S_T^∞ , T^* and T_0 . S_T^∞ corresponds to the extrapolated Soret coefficient at high temperatures, T^* is the temperature at which S_T changes sign and T_0 describes its temperature sensitivity. It has been shown to hold for a number of biological systems [11], but fails to describe the temperature dependence of diluted solutions of ethanol [12], dimethyl sulfoxide [13] or ethylen glycol oligomers [14] in water. A common feature of these aqueous solutions is the presence of micro heterogeneity in the solution [15, 16]. For formamide it has been shown to hold only for dilute solutions and break down at higher concentrations [17]. Recently, Niether *et al.* [18] related the empirical parameters in Eq.1 with the number and the strength of hydrogen bonds. Using a linear correlation between those parameters, it was possible to reduce the number of adjustable parameters to two for cyclodextrins and their complexes with acetylsalicylic acid.

Wittko and Köhler [19] performed systematic concentration- and temperature-dependent measurement of low molecular weight mixtures. The majority of the systems they discussed were non-polar, but they considered the system ethanol/water in their description. The main outcome was the existence of a temperature-independent fix point at a finite concentration, so they proposed an empirical ansatz to describe the temperature and concentration dependent Soret coefficient $S_T(c, T)$

$$S_T(c, T) = \alpha(c)\beta(T) + S_T^i \quad (2)$$

with polynomial serial expansions for $\alpha(c)$ and $\beta(T)$

$$\begin{aligned} \alpha(c) &= a_0 + a_1c + a_2c^2 + a_3c^3 + \dots, \\ \beta(T) &= 1 + b_1(T - T_0) + b_2(T - T_0)^2 + \dots \end{aligned} \quad (3)$$

In their study they used the molar fraction as concentration variable c , $T_0 = 25^\circ\text{C}$ is an arbitrary reference temperature and S_T^i is a temperature and concentration independent constant, which could be expressed as $S_T^i = a_M\delta M + b_I\delta I$ with the mass and moment of inertia difference δM and δI of two non-polar components and the two adjustable parameters a_M and b_I .

Amides belong to the class of associating systems, which have often been investigated by simulations and calculations [13, 20–25]. For these systems strong cross interactions

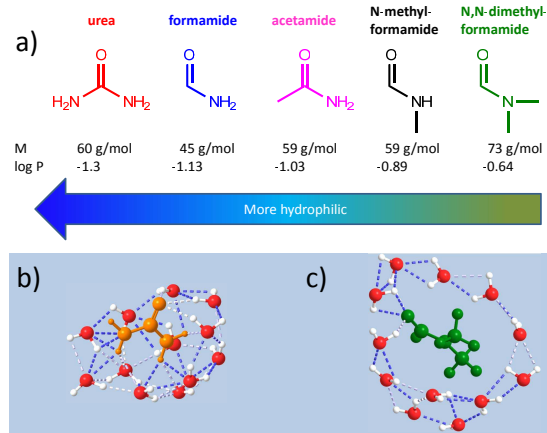


FIG. 1: (a) Chemical structure, molecular weight and $\log P$ values of the amide systems in sequence of their hydrophilicity. (b) schematic view of urea and (c) N,N-dimethylformamide including hydrogen bonds formed with the surrounding water.

between unlike molecules are responsible for the observed sign changes [22]. Systematic simulations [23, 24] show that also the pronounced concentration dependence of those mixtures is strongly related to these cross interactions ϵ_{12} . Artola *et al.* [24] express ϵ_{12} with a modified Lorentz-Berthelot mixing rule $\epsilon_{12} = k_{12} * \sqrt{\epsilon_{11}\epsilon_{22}}$. The simulations demonstrate that with increasing strength of the cross interactions k_{12} compared to the like-interactions the concentration dependent slope of S_T increases. Additionally, they introduce the energy ratio $\psi_\epsilon = \epsilon_{22}/\epsilon_{11}$ as second parameter. It turns out ψ_ϵ has almost no effect on the slope and might be contributing to S_T^i in Eq.2, which accounts for mass or inertia effects [19].

The considered amides are displayed in Fig. 1(a). This substance class is characterized by an $O=C-NR_2$ -group, which can serve as a model for the peptide bond, an important structure in proteins. Due to their polarity, the oxygen as well as the amino-group (NH_2) can act as a hydrogen bond site with water. The investigated amides show a good solubility in water so that a wide concentration range is accessible for measurement. Formamide is the simplest possible amide with no additional groups. Urea has a second amino-group and therefore more potential bond sites. Acetamide (AC), N-methylformamide (NMF) and N,N-dimethylformamide (DMF) can all be viewed as a methylated formamide. While the hydrophobic methyl-group does not change the number of potential hydrogen bond sites in the AC, in NMF and DMF they block the amino-group partially or entirely. The partition

coefficient P of a solute in two immiscible solvents is given by the ratio of the solute in the two phases and commonly expressed as the logarithm $\log P$. The displayed $\log P$ in octanol/water (Fig. 1(a)) is a physical property that is commonly used as a measure of hydrophobicity/hydrophilicity of a substance. It is used as a parameter in the screening for potential pharmaceutical compounds (Lipinski's 'Rule of Five' [26]). Exemplarily, Fig. 1(b) and (c) illustrate schematically the hydration of the most and least hydrophilic solutes urea and N,N-dimethylformamide (DMF). Experimental studies and simulations support a good compatibility of urea with the water network, where urea occupies roughly the volume of 2-3 water molecules and has strong intermolecular hydrogen bonds [27]. DMF has typically only one hydrogen bond [16], so that the water molecules form a clathrate-like structure around DMF. A comparative study of pure FA, N-methylformamide (NMF), and DMF found that the possible number of hydrogen bonds decreases from four per molecule in FA to one in DMF. There have been a few comparative studies of amides [28, 29]. Papamatthaiakis *et al.* [29] find that the isentropic compressibility of almost all amides shows a minimum as function of molar fraction, while in aqueous FA solutions the compressibility decreases monotonically. The authors argue that both FA and water molecules have the ability to form hydrogen bonds, so that in contrast to the other amides the transition is continuous. The compressibility minimum observed for the more hydrophobic amides might be rationalized, by a clathrate-like water structure surrounding the amide molecules at sufficiently high water content [30, 31]. While at higher solute concentrations the number of water molecules is not sufficient to form a clathrate-like cage, so that the compressibility increase monotonously towards the value of the pure solute.

In this work we measured the thermodiffusion behaviour of N-methylformamide (NMF), N,N-dimethylformamide (DMF), and acetamide (AC) in water in a temperature range between 10 and 70°C and a concentration range between 2 and partially up to 70wt% by infrared thermal diffusion forced Rayleigh scattering (IR-TDFRS). Additionally, we re-analyse the previously investigated aqueous solutions of the more hydrophilic amides formamide (FA) [17, 32] and urea [25]. The temperature and concentration dependence of the Soret coefficient S_T and its correlation with the logarithm of the partition coefficient $\log P$ are discussed.

II. METHODS AND MATERIALS

A. Thermodiffusion

Thermophoresis, or thermodiffusion, is the movement of molecules in a temperature gradient. In a binary fluid mixture it can be described as a flux \vec{J} of the solute molecules,

$$\vec{J} = -\rho D \nabla c - \rho c(1 - c) D_T \nabla T \quad (4)$$

that is the sum of the thermodiffusion along the temperature gradient ∇T with the thermal diffusion coefficient D_T and the Fickian diffusion along the resulting concentration gradient ∇c with the diffusion constant D . In a steady temperature gradient both contributions become equal and the flux becomes zero. In the steady state, the concentration gradient over the temperature gradient is proportional to the ratio D_T over D .

$$S_T \equiv \frac{D_T}{D} \propto \frac{\Delta c}{\Delta T} \quad (5)$$

This ratio is known as the Soret coefficient S_T , a quantity that is used to describe the thermophoretic behaviour of a solute. Note that a positive Soret coefficient signifies a thermophobic solute (movement towards the cold side of the temperature gradient), a negative S_T a thermophilic solute (movement towards the warm side) and the absolute value how pronounced the resulting concentration gradient is.

B. Thermal diffusion forced Rayleigh scattering

To measure the thermodiffusion we used IR-TDFRS, a transient grating technique [33, 34]. Two infra-red laser beams are crossed inside the sample and create a holographic grating that creates a temperature grating due to the inherent absorption of water in that range [35]. Thermodiffusion leads to migration of the particles in that temperature gradient and a superimposed concentration grating. Both, temperature and concentration changes, give rise to changes in the refractive index of the sample. The refractive index grating is probed by Bragg diffraction of a third laser beam, the read-out beam. The heterodyne scattering intensity $\zeta_{het}(t)$ of the read-out beam is measured and fitted with

$$\begin{aligned} \zeta_{het}(t) = & 1 - \exp\left(-\frac{t}{\tau_{th}}\right) - A(\tau - \tau_{th})^{-1} \\ & \times \left\{ \tau \left[1 - \exp\left(-\frac{t}{\tau}\right) \right] - \tau_{th} \left[1 - \exp\left(-\frac{t}{\tau_{th}}\right) \right] \right\}. \end{aligned} \quad (6)$$

With the lifetimes $\tau_{th} = (D_{th}q^2)^{(-1)}$ and $\tau = (Dq^2)^{(-1)}$ of the temperature and concentration grating, respectively. q , D_{th} and D denote the grating vector, the thermal diffusivity and the collective diffusion coefficient. Given that the temperature and concentration dependence of the refractive index n are known (see supplementary for details), the Soret coefficient can be derived from the Amplitude

$$A = \left(\frac{\partial n}{\partial c} \right)_{p,T} \left(\frac{\partial n}{\partial T} \right)_{p,c}^{-1} S_T c (1 - c). \quad (7)$$

C. Sample preparation

The investigated substances, acetamide ($\sim 99\%$), N-methylformamide (99%), and N,N-dimethylformamide ($\geq 99\%$), were purchased from Sigma Aldrich and used without further purification. The solutions were prepared with distilled and deionized water (Millipore). The solutions were filtered through a $0.2 \mu\text{m}$ filter (Whatman Anotop 10) before filling them into an optical quartz cell (Hellma) with an optical path length of 0.2 mm. The concentration dependence of the refractive index was measured for at least 7 concentrations each, within an hour after preparation. The resulting functions were used to determine the concentrations of subsequent samples and to check for changes in older solutions. There were no significant changes of refractive index in the solutions when stored at room temperature, even for long times (2 weeks or more). Slight concentration changes ($< 2\%$) were noted in some samples after measurement, due to evaporation at high temperatures during measurement.

III. RESULTS

A. Concentration dependence

First we use Eq. 2 to describe the temperature and concentration dependence of S_T . Figure 2 shows the concentration dependence of S_T as function of the weight fraction of all considered amide/water systems at various temperatures. Using a simultaneous fit according to Eq. 2 we found a satisfactory description of all studied systems. In the case of urea, NMF, and DMF a cubic description of the concentration dependence was required.

The temperature independent fix point, which corresponds to a zero of $\alpha(c)$ could be observed for urea and FA. In the case of AC it was extrapolated since it is close to the

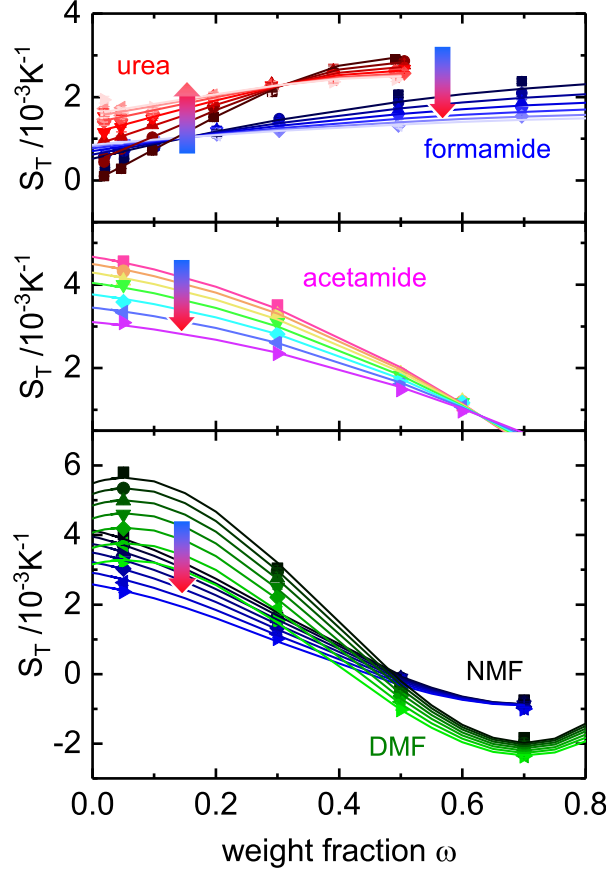


FIG. 2: Soret coefficient for all amide systems as function of weight fraction for different temperatures between 10° and 70° . The arrows mark the direction of increasing temperature. The lines are fits according to Eq. 2 with the weight fraction as concentration variable and 25°C as arbitrary reference temperature.

solubility limit, but should lie only slightly above 0.6 weight fraction. In the case of NMF it is expected to be close to a weight fraction of 0.7, while DMF, similar to the studied system dibrommethane in cyclohexane [19], seems not to have a temperature independent fix point. For both systems the polynomial fit of S_T shows a maximum at low and minimum at high solute concentrations.

Figure 3 shows b_1 and S_T^i as function $\log P$ for five amide systems (c.f. Eq.2). The numerical values are listed in table ?? of supporting information. The parameter S_T^i , has originally been introduced as a parameter to describe the isotope effect since it does not show any temperature and concentration dependence and the parameter b_1 is related to the temperature dependence in first order. The shaded areas mark the range of the values

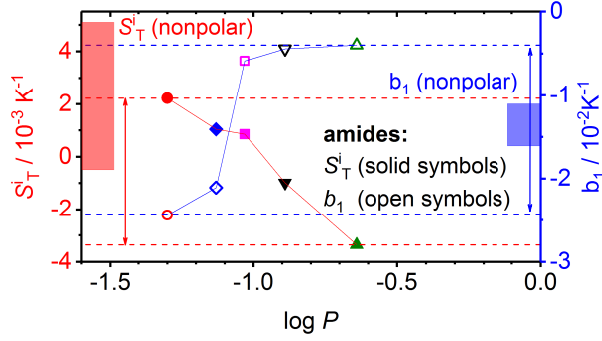


FIG. 3: S_T^i values (solid symbols) and b_1 (open symbols) as function of $\log P$ for the five amides. The red and blue arrows mark the minimum and maximum values of S_T^i and b_1 , respectively. The shaded areas mark the range found for nonpolar mixtures investigated by Wittko and Köhler [19].

determined by Wittko and Köhler [19] for mainly nonpolar mixtures. S_T^i of the amides is more thermophilic (negative values) in comparison with the nonpolar systems, but the spread of the data is similar. S_T^i changes almost linearly with $\log P$ (in the supporting information we discuss also the correlation with the Hildebrandt solubility parameter). This suggests that S_T^i is not solely determined by mass and inertia effects. While the S_T^i -values of the five amides can be related to their differences in mass and moment of inertia an extension to other aqueous ethanol and glucose mixtures fails completely. For nonpolar mixtures the chemical contributions are fully expressed by $\alpha(c)$ and $\beta(T)$ (c.f. Eq.3), but for polar systems this might be not possible due to the strong temperature dependence of the specific interactions. This seems also to be reflected by the much wider distribution of b_1 of the amides compared to the nonpolar mixtures. This hypothesis is additionally supported by the observation that the magnitude of b_1 of the two strong hydrogen bonding solutes urea and FA is four times higher compared to the more hydrophobic amides. A similarity of all mixtures, polar and nonpolar, is the negative sign of b_1 .

B. Temperature dependence

Compared to the non-polar mixtures we observe for the considered amide solutions a larger variation of the temperature dependence, which we would like to scrutinize in more detail. Figure 4 shows the Soret coefficient S_T against temperature at concentrations of 5

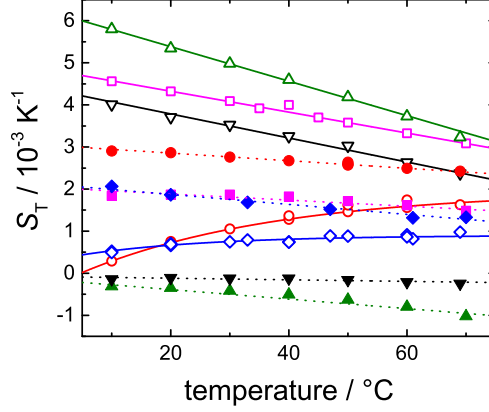


FIG. 4: The Soret coefficient S_T of the investigated amides at 5 (open symbols) and 50wt% (solid symbols) for urea (red circle), formamide (blue diamond), acetamide (magenta square), N-methylformamide (black triangle down), and N,N-dimethylformamide (green triangle up).

and 50 wt% (open and solid symbols, respectively). Looking at the temperature dependence of S_T and its change when going from low to high concentrations, it is evident that the investigated substances can be divided into two groups. Urea and formamide fall into one group. At low concentrations the Soret coefficient increases with rising temperature and Eq. 1 can be used to describe $S_T(T)$. With rising concentration thermophobicity increases and the temperature dependence of S_T weakens until, above a certain concentration, there is a decrease of S_T with rising temperature. AC, NMF and DMF fall into the second group. For these substances we observe a linear decrease of S_T with rising temperatures at low concentrations. At high concentrations, thermophobicity decreases. At 50 wt% (Fig. 4, filled symbols) NMF and DMF show a thermophilic behaviour, the temperature dependence of S_T weakens compared to the low concentration, but still has a negative slope. This difference in behaviour becomes even more obvious when looking at the temperature dependence of the thermal diffusion coefficient D_T (Fig. 5). While D_T is more or less constant over a large concentration range for urea and FA, it decreases with increasing concentration for AC, NMF, and DMF. All investigated substances show an increase of D_T with rising temperature at low concentrations. At 5wt% we can distinguish between two groups which show very similar D_T -values: urea and FA in the first and AC, NMF, and DMF in the second, more thermophobic, group. At high concentrations, this picture changes. While for urea and FA the behaviour of D_T shows almost no change compared to that at low

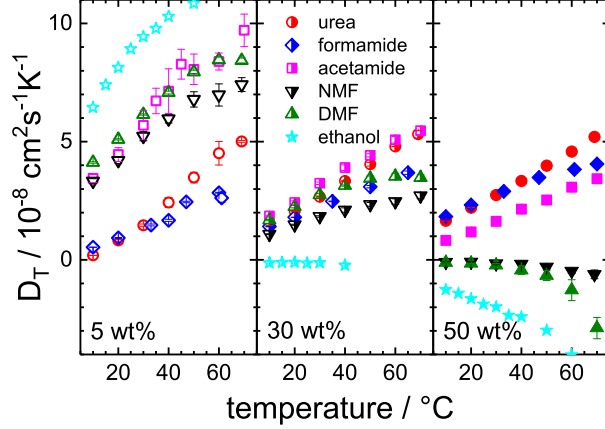


FIG. 5: Thermal diffusion coefficient D_T of the investigated amides at 5 (open symbols), 30 (half filled symbols), and 50 wt% (solid symbols). Same symbols as in Fig.4, ethanol data (cyan stars) reproduced from [12].

concentrations (there is a slight shift to higher values at cold temperatures), the methylated compounds AC, NMF, and DMF shift strongly to lower values, indicating an increasing thermophilicity of the solutes. For comparison we display also the D_T -values of ethanol in water [12]. Ethanol is less hydrophilic than DMF and the change in the temperature dependent slope is even more pronounced.

The diffusion coefficient D and its temperature dependence is rather similar for all investigated substances at low concentrations (shown in SI). As expected from their molecular masses (see Fig. 1), urea, AC and NMF have the same diffusion coefficients, while formamide diffuses slightly faster and DMF slightly slower. At a concentration of 50wt% the variance between the amides is more marked. While we observe no significant change in the diffusion of formamide, the other systems show a shift to lower D -values, especially at low temperatures.

IV. DISCUSSION AND CONCLUSION

Although the temperature and concentration dependence of the amide systems has a higher complexity than the nonpolar mixtures, a description of the experimental results with Eq.2 is possible. In order to understand the changes in the slope let us consider the two parameters introduced by Artola and Rousseau [24]: k_{12} and ψ_ϵ , the strength of the cross interactions and the energy ratio, respectively. The observed concentration dependence

for the different amides (c.f. Fig.2) suggests that k_{12} increases from values below 1 for the hydrophilic compounds urea and formamide to values above 1 for the more hydrophobic compounds. While for nonpolar systems ψ_ϵ is mainly determined by mass and moment of inertia effects, this does not hold for polar substances. As mentioned in the results section, the S_T^i found for polar systems cannot be expressed as a function of mass and moment of inertia, but correlates with $\log P$. Therefore we might introduce a third term characterizing the interaction level,

$$S_T^i = a_M \delta M + b_I \delta I + c_\epsilon \Delta \psi_\epsilon. \quad (8)$$

Due to the similar masses and shapes of the amides the first two terms are almost constant. Considering only the amides $c_\epsilon \Delta \psi_\epsilon$ can be expressed by an equation of the form $a + b \cdot \log P$, whereas the intercept contains mass and inertia effects. While it is possible to determine a_M and b_I for amides, there is no solution which also covers aqueous solutions of ethanol and glucose. Therefore, we assume that non-ideal structural changes (e.g. excess effects) of the polar systems with temperature and concentration are not fully covered by this simple approach.

To clarify the observed difference between the amide systems we would like to scrutinize those structural and interaction changes. Only two amides, FA and urea, show the typical temperature dependence of S_T [11]. As mentioned in the introduction, a common feature of aqueous solutions, which deviate from this behaviour, is the presence of micro heterogeneity in the solution. This hypothesis holds also for the three additionally investigated amides. Recently, it was shown in simulations that AC forms ring like structures similar to those formed by ethanol in water [36]. A comparative study of formamide and NMF shows that the additional methyl group disrupts the almost homogeneous encasement by oxygen observed in aqueous formamide solutions [37]. Aqueous DMF mixtures are controversially discussed: while earlier works favour micro heterogeneities [16, 38] an improved force field leads to a better prediction of bulk parameters, improves the miscibility and leads to less heterogeneities on a molecular level [39]. In order to gain a better understanding of the underlying mechanisms in the considered mixtures we want to look into two aspects: first the structural changes of the mixture by considering the thermal expansion coefficient and second the governing interactions especially the hydrogen bonding, which we try to quantify by its hydrophilicity.

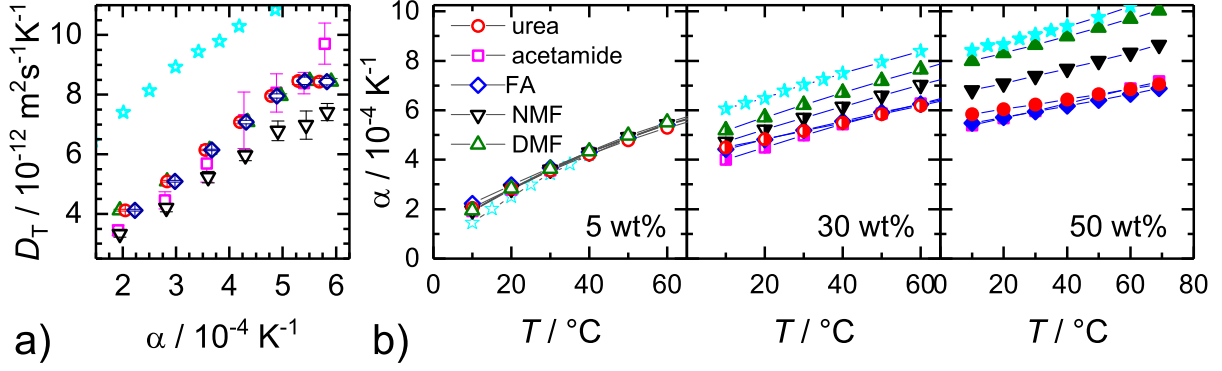


FIG. 6: (a) Thermal diffusion coefficient D_T of amides as function the experimentally determined thermal expansion coefficient α at a weight concentration of 5%. (b) Experimental thermal expansion coefficient of urea, formamide and DMF at 5, 30 and 50 wt% as function of temperature. Ethanol data (cyan stars) are calculated from Wittko and Köhler [12]. We used the same symbols as in Fig. 5.

A. Thermal expansion

Different theoretical models can be found in the literature relating the thermophoretic mobility or thermal diffusion coefficient with the thermal expansion coefficient α , and viscosity. For instance, in the model developed by Semenov and Schimpf [40] for dilute solutions of dissolved or suspended components, the thermophoretic mobility (D_T) is proportional to the ratio between the thermal expansion and the dynamic viscosity. Likewise, Brenners [41] kinematic model of thermal diffusion for liquid and gases establishes that D_T is proportional to the thermal expansion coefficient. This linear correlation has also been observed for sugar solutions [42, 43] and for alkanes in organic solvent [44].

In Fig. 6(a) we show the thermal diffusion coefficients of diluted solutions (5wt%) of all amides versus the thermal expansion coefficient $\alpha = (1/V) \cdot (\partial V / \partial T)$, whereas the effective volume, V , has been estimated using the Clausius-Mosotti equation from the refractive index measurements, leading to thermal expansion coefficient

$$\alpha(T) = -\frac{6 \cdot n(T) \cdot (\partial n / \partial T)_{p,c}}{n(T)^4 + n(T)^2 - 2} \quad (9)$$

At the low concentration we find for all amides a similar linear dependence of D_T as function of α (c.f. Fig. 6(a)). This linear correlation between D_T and alpha is indeed one, which holds for polar as well as for non-polar mixtures. Compared to the other amides NMF

shows the largest deviations just outside the uncertainty. For comparison we display also the D_T -values of ethanol in water. The thermal expansion dependent slope of D_T of ethanol is similar to the amides, but the values are roughly two times higher. This can not be explained by different viscosities [45], but might be related to the two times lower dipole moment of ethanol compared to the amides, which influences the hydrogen bonding [46].

In the results section we have seen (c.f. Fig. 5) that the temperature dependence of the thermal diffusion coefficient of urea and FA does not vary much with concentration, while the more hydrophobic amides show a strong temperature dependence change of D_T with concentration. To investigate this further, we calculated the thermal expansion coefficient of the five amides at different concentrations as a function of temperature (c.f. Fig. 6(b)). At low concentrations all systems show the same thermal expansion, which indicates that α agrees with water at this low solute concentration. At higher concentrations α increases for all systems, but the effect is much stronger for DMF and NMF than it is for AC, FA, and urea, which remain almost constant.

The physical reason for this behaviour is probably the better compatibility of urea and formamide with the water network, so that an increase of the solute concentration requires less rearranging of the fluid. For these more hydrophilic amides, due to the fact that at all concentrations hydrogen bonds are the dominating interactions, the behaviour of D_T seems to be solely determined by the change of α . This results in the weak concentration dependence shown in Fig. 5. In contrast, increasing the concentration of the more hydrophobic amides leads to a change of the fluid structure. As illustrated in Fig. 1 the water molecules form a clathrate structure around the molecule inducing a void, which will not be filled by water molecules leading to a higher thermal expansion with increasing concentration and hydrophobicity (c.f. Fig. 6). At first glance this increase of α seems to contradict the drop of the thermal diffusion coefficient with increasing concentration of the hydrophobic solute. To clarify this point we need to regard the dominating interactions in the mixture.

B. Correlation with $\log P$

A correlation is observed between the temperature and concentration dependence of the Soret coefficient, $\Delta S_T(\Delta T) = S_T(50^\circ\text{C}) - S_T(10^\circ\text{C})$ and $\Delta S_T(\Delta c) = S_T(50\text{wt}\%) - S_T(5\text{wt}\%)$ respectively, and the partition coefficients of the investigated substances, $\log P$. This is

shown in Fig. 7. The system ethanol/water has been investigated in a broad temperature and concentration range [12] so that we can include it in our analysis. Note that there is no direct correlation to the absolute value of S_T , since that contains also the mass and inertia contribution. But ΔS_T is assumed to be proportional to the chemical contribution, because it is the portion of S_T that is dependent on temperature or concentration.

The partition coefficient $\log P$ is used to quantify the balance of hydrophobic and hydrophilic groups in a molecule, which determines its interactions with the surrounding solvent and thus, influences its thermodiffusion behaviour. When a substance is dissolved in two immiscible solvents, it will not distribute equally between the phases, but will reach a characteristic concentration ratio in the respective solvents. This ratio is given by the partition coefficient P . Since the concentration difference in 1-octanol and water arises from Fickian diffusion along the solute’s activity gradient between the phases, $\log P$ is closely related to the activity coefficient γ of the solute. The partition coefficient is measured in equilibrium, when the activities of the solute in the two phases are equal ($a_{octanol} = a_{water}$). This can also be expressed as $\gamma_{octanol} \cdot c_{octanol} = \gamma_{water} \cdot c_{water}$, where $\gamma_{solvent}$ and $c_{solvent}$ are the solute’s activity coefficient and concentration in the respective solvent, so that

$$\log P = \log \left(\frac{c_{octanol}}{c_{water}} \right) = \log \gamma_{water} - \log \gamma_{octanol}, \quad (10)$$

where $\log \gamma_{octanol}$ is constant within a homologous group and shows slight variation between chemical species. It was shown to be in the order of $\log \gamma_{octanol} \approx 1$ [47]. This implies that the partition coefficient $\log P$ of a given solute is proportional to its activity coefficient in water $\log \gamma_{water}$ and can be used as a measure of solute-solvent interactions in aqueous solutions.

In these systems the intermolecular interactions are dominated by hydrogen bonds. The hydrogen bond strength between the solute and water, as well as the strength of the hydrogen bond network of the water are expected to be sensitive to temperature changes, resulting in changes in enthalpy and entropy of the system. Assuming that the chemical contribution must decrease with rising temperatures, we see that it can be negative, leading to a rising $S_T(T)$, or positive, leading to a decreasing temperature dependence. This indicates different mechanisms that make solute-solvent-interactions attractive or repulsive. A hydrophilic group of the solute molecule can form hydrogen bonds with the surrounding water, which leads to a decrease of free energy. Hydrophobic parts of the molecule cannot form hydrogen

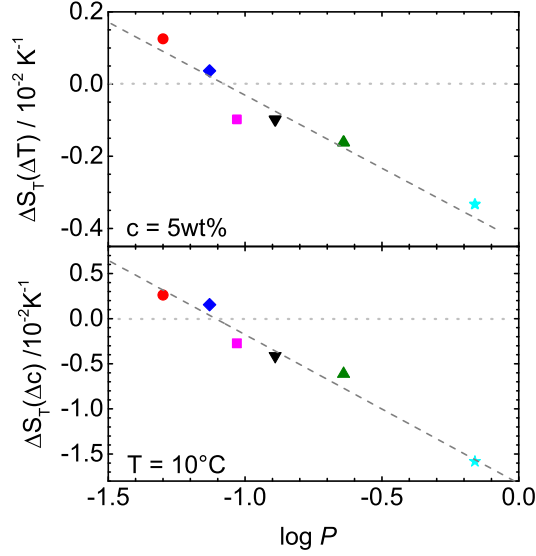


FIG. 7: The change of S_T with temperature $\Delta S_T(\Delta T)$ (upper graph) and concentration $\Delta S_T(\Delta c)$ (lower graph) is correlated with the hydrophilicity $\log P$. Same symbols as in Fig 4.

bonds. The HB network of water is broken and has to be rearranged into clathrate-like cages to accommodate their volume [48]. The breaking of hydrogen bonds as well as the distortion of the network raises the free energy of the system.

The concentration and the temperature dependence of S_T both correlate with $\log P$. In both cases the sign change of ΔS_T is at $\log P \approx -1.1$ or, if we look at the corresponding activity coefficient, at $\gamma_{water} \approx 1$. This means that, when the mixture behaves like an ideal solution, the chemical contribution is zero. This correlation indicates a proportionality between the chemical contribution to the Soret coefficient and deviation from ideal behaviour expressed by the activity coefficient.

Microscopically, the underlying physical reason for the correlation between $\log P$ and the temperature sensitivity of the Soret coefficient might be related to the solute's capability to form hydrogen bonds. It is reasonable to assume that a compound with a low $\log P$ value has more hydrogen bonds to the surrounding water. Those bonds are weakened when the temperature is raised which leads to a changed of the chemical contribution of the Soret coefficient. Therefore we assume that the change in S_T is larger. The correlation with $\log P$ is probably limited to fairly small molecules, which are too small to coil or to fold, so that all chemical groups are accessible by the solvent. This shows that thermodiffusion is

extremely sensitive to probe hydration of solutes in water. In contrast to $\log P$ it can also be determined at temperatures other than RT and it can be used to describe the hydrophilicity of complexes [18].

V. SUPPLEMENTARY MATERIAL

Shows the diffusion coefficients for AC, NMF and DMF in water at 5 and 50wt%, the refractive index contrast factors with temperature and the refractive index values as function of temperature and concentration. Additionally the correlations with the so-called Hildebrandt parameter are shown. In the last section additional the thermal expansion coefficients of FA, DMF and urea are shown, which have been calculated using published density data.

ACKNOWLEDGMENTS

Part of the experimental data presented were obtained with funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No 731019 (EUSMI) which is gratefully acknowledged. DN acknowledges support by the International Helmholtz Research School of Biophysics and Soft Matter (IHRS BioSoft). Calculator Plugins were used for structure property prediction and calculation (of $\log P$), Marvin 16.5.2.0, 2016, ChemAxon (<http://www.chemaxon.com>).

-
- [1] M. Jerabek-Willemsen, T. Andr, W. Wanner, H. Roth, S. Duhr, P. Baaske, and D. Breitsprecher, *J. Mol. Struct.* **1077**, 101 (2014).
 - [2] Y. A. Mikheev, L. N. Guseva, E. Y. Davydov, and Y. A. Ershov, *Russian Journal of Physical Chemistry A* **81**, 1897 (2007).
 - [3] W. Köhler and K. I. Morozov, *J. Non-Equil. Thermody.* (2016).
 - [4] S. Wiegand, *J. Phys.: Condens. Matter* **16**, R357 (2004).
 - [5] J. K. G. Dhont, S. Wiegand, S. Duhr, and D. Braun, *Langmuir* **23**, 1674 (2007).
 - [6] J. K. G. Dhont and W. J. Briels, *Eur. Phys. J. E* **25**, 61 (2008).
 - [7] K. A. Eslahian, A. Majee, M. Maskos, and A. Wurger, *Soft Matter* **10**, 1931 (2014).
 - [8] C. Debuschewitz and W. Köhler, *Phys. Rev. Lett.* **87**, 055901 (2001).

- [9] S. Hartmann, W. Köhler, and K. I. Morozov, *Soft Matter* **8**, 1355 (2012).
- [10] S. Iacopini, R. Rusconi, and R. Piazza, *Eur. Phys. J. E* **19**, 59 (2006).
- [11] Y. Kishikawa, S. Wiegand, and R. Kita, *Biomacromolecules* **11**, 740 (2010).
- [12] A. Königer, B. Meier, and W. Köhler, *Philos. Mag.* **89**, 907 (2009).
- [13] P. Polyakov and S. Wiegand, *J. Chem. Phys.* **128**, 034505 (2008).
- [14] K. Maeda, N. Shinyashiki, S. Yagihara, S. Wiegand, and R. Kita, *J. Chem. Phys.* **143**, 124504 (2015).
- [15] O. Gereben, *Journal of Molecular Liquids* **211**, 812 (2015).
- [16] Y. Lei, H. Li, H. Pan, and S. Han, *The Journal of Physical Chemistry A* **107**, 1574 (2003).
- [17] D. Niether, D. Afanasenkau, J. K. G. Dhont, and S. Wiegand, *Proceedings of the National Academy of Sciences* **113**, 4272 (2016).
- [18] D. Niether, T. Kawaguchi, J. Hovancova, K. Eguchi, J. K. G. Dhont, R. Kita, and S. Wiegand, *Langmuir* **33**, 8483 (2017).
- [19] G. Wittko and W. Köhler, *Epl* **78**, 46007 (2007).
- [20] I. Prigogine, L. Debrouckere, and R. Amand, *Physica* **16**, 851 (1950).
- [21] H. Ning and S. Wiegand, *J. Chem. Phys.* **125**, 221102 (2006).
- [22] J. Luettmmer-Strathmann, *International Journal of Thermophysics* **26**, 1693 (2005).
- [23] B. Rousseau, C. Nieto-Draghi, and J. Avalos, *Europhys. Lett.* **67**, 976 (2004).
- [24] P. Artola and B. Rousseau, *Phys. Rev. Lett.* **98**, 125901 (2007).
- [25] D. Niether, S. Di Lecce, F. Bresme, and S. Wiegand, *Physical Chemistry Chemical Physics* **20**, 1012 (2018).
- [26] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, *Adv. Drug Delivery Rev.* **64**, 4 (2012).
- [27] A. K. Soper, E. W. Castner, and A. Luzar, *Biophys. Chem.* **105**, 649 (2003).
- [28] P. C. Schoester, M. D. Zeidler, T. Radnai, and P. A. Bopp, *Zeitschrift Fur Naturforschung Section a-a Journal of Physical Sciences* **50**, 38 (1995).
- [29] D. Papamatthaiakis, F. Aroni, and V. Havredaki, *Journal of Chemical Thermodynamics* **40**, 107 (2008).
- [30] E. Baumgartner and G. Atkinson, *J. Phys. Chem.* **75**, 2336 (1971).
- [31] G. Onori, *J. Chem. Phys.* **89**, 4325 (1988).
- [32] D. Niether and S. Wiegand, *Entropy* **19**, 33 (2017).

- [33] S. Wiegand and W. Köhler, LNP Vol. 584: Thermal Nonequilibrium Phenomena in Fluid Mixtures **584**, 189 (2002).
- [34] P. Blanco, H. Kriegs, M. P. Lettinga, P. Holmqvist, and S. Wiegand, Biomacromolecules **12**, 1602 (2011).
- [35] S. Wiegand, H. Ning, and H. Kriegs, J. Phys. Chem. B **111**, 14169 (2007).
- [36] J. A. Aguilar-Pineda, G. A. Mendez-Maldonado, E. Nunez-Rojas, and J. Alejandre, Molecular Physics **113**, 2716 (2015).
- [37] A. K. H. Weiss, T. S. Hofer, B. R. Randolph, A. Bhattacharjee, and B. M. Rode, Physical Chemistry Chemical Physics **13**, 12173 (2011).
- [38] C. Caleman, P. J. van Maaren, M. Y. Hong, J. S. Hub, L. T. Costa, and D. van der Spoel, Journal of Chemical Theory and Computation **8**, 61 (2012).
- [39] V. Vasudevan and S. H. Mushrif, Journal of Molecular Liquids **206**, 338 (2015).
- [40] S. Semenov and M. Schimpf, Phys. Rev. E **69**, 011201 (2004).
- [41] H. Brenner, Phys. Rev. E **74**, 036306 (2006).
- [42] P. Blanco and S. Wiegand, J. Phys. Chem. B **114**, 2807 (2010).
- [43] P. Blanco, H. Kriegs, B. Arlt, and S. Wiegand, J. Phys. Chem. B **114**, 10740 (2010).
- [44] P. Blanco, M. M. Bou-Ali, J. K. Platten, P. Urteaga, J. A. Madariaga, and C. Santamaria, J. Chem. Phys. **129**, 174504 (2008).
- [45] M. Hartung, J. Rauch, and W. Khler, J. Chem. Phys. **125**, 214904 (2006).
- [46] Y. Marcus, The Properties of Solvents, Vol. 4 (John Wiley & Sons, Chichester, 1999).
- [47] U. Avico, E. C. Signoretti, and P. Zuccaro, Farmaco-Edizione Scientifica **35**, 590 (1980).
- [48] F. Jimenez-Angeles and A. Firoozabadi, ACS Cent. Sci. (2018).