Journal Name

ARTICLE TYPE

Cite this: DOI: 10.1039/xxxxxxxxxx

Unravelling the hydrophobicity of urea in water using thermodiffusion: implications for protein denaturation†

Doreen Niether,[a‡] Silvia Di Lecce,[b‡] Fernando Bresme,*[b] and Simone Wiegand*[a,c]

Received Date
Accepted Date

DOI: 10.1039/xxxxxxxxxx

www.rsc.org/journalname

Urea is widely used as protein denaturant in aqueous solutions. Experiments and computer simulation studies have shown that it dissolves in water almost ideally at high concentrations, introducing little disruption in the water hydrogen bonded structure. However, at concentrations of the order of 5 M or higher, urea induces denaturation in a wide range of proteins. The origin of this behaviour is not completely understood, but it is believed to stem from a balance between urea-protein and urea-water interactions, with urea becoming possibly hydrophobic at a specific concentration range. The small changes observed in the water structure makes it difficult to connect the denaturation effects to solvation properties. Here we show that the exquisite sensitivity of thermodiffusion to solute-water interactions allows the identification of the onset of hydrophobicity of urea-water mixtures. The hydrophobic behaviour is reflected in a sign reversal of the temperature dependent slope of the Soret coefficient, which is observed, both in experiments and non-equilibrium computer simulations at $\sim$ 5 M concentration of urea in water. This concentration regime corresponds to the one where abrupt changes in the denaturation of proteins are commonly observed. We show that the onset of hydrophobicity is intrinsically connected to the urea-water interactions. Our results allow us to identify correlations between the Soret coefficient and the partition coefficient, $\log P$, hence establishing the thermodiffusion technique as a powerful approach to study hydrophobicity.

The biological relevance of urea, particularly as protein denaturant, as well as its influence on the structure and dynamics of water has motivated a large number of experimental and computer simulation studies [1–13]. The origin of its unique denaturant activity has not been fully resolved. Computer simulations of proteins in highly concentrated (8 M) urea aqueous solutions suggested that the denaturation process proceeds through two main steps: (1) urea acts as a surfactant, displacing water in the first hydration shell ("dry globule" formation) of the protein. This process is mediated by the van der Waals dispersion interactions, which makes protein-urea contacts more favourable than protein-water ones; and (2) urea binds to the protein, where it can interact with the backbone and with polar and non-polar side chains [14–17].

These results indicate that protein-water and protein-urea interactions play a key role in denaturation. However, the fact that denaturation is observed at a specific urea concentration, indicates that the urea-water interactions play also an important role in defining the onset of conformational changes in the proteins. The aggregation of urea around the protein during the first denaturation step is driven by the energetic balance between urea-water and urea-protein interactions. Indeed, the denaturation has been interpreted, too, in terms of a global change of the solvent properties [18,19]. Despite the large number of studies of urea solutions, a full microscopic picture of the interactions of water and urea is still lacking. While some studies classify urea as structure breaker [2] others concluded that the water network is not influenced or even strengthened by the addition of urea, even at high concentrations [11–13]. Experimental studies of water dynamics have identified two populations of water molecules, which interact weakly or strongly with urea, sharing either one or two hydrogen bonds with urea, respectively [1,6,8]. The stiffening of the hydrogen bond structure at high concentrations, 8 M, is expected to slow down the orientational dynamics of water (about six times with respect to bulk) [1]. Computer simulations showed that the orientational dynamics depend on urea concen-

[a] ICS-3 Soft Condensed Matter, Forschungszentrum Jülich GmbH, D-52428 Jülich, Germany. E-mail: s.wiegand@fz-juelich.de
[b] Department of Chemistry, Imperial College, London SW7 2AZ, United Kingdom. E-mail: f.bresme@imperial.ac.uk
[c] Department für Chemie - Physikalische Chemie, Universität zu Köln, 50939 Cologne, Germany

† Electronic Supplementary Information (ESI) available: pdf file with additional experimental and simulation results. See DOI: 10.1039/b000000x/

‡ D.N. and S.D.L. contributed equally to this work.

tration[2,20], and have reported a variety of structural changes in the solution, which become more evident at high concentrations 4-6 M[8–12]. On the contrary, Nuclear Magnetic Resonance experiments of concentrated solutions, 10 M, have proved inconclusive regarding the slowing down of water dynamics, as well as the possible structural changes of the solution[21]. It is fair to say that although hydration plays a key role in all processes taking place in aqueous solutions, the microscopic picture is unclear. The concept of structure-breaking and -making in water itself is poorly defined and its interpretations in literature can vary significantly.[22] Further, the prediction of hydration properties such as hydrophobicity relies on the extrapolation of empirical data[23], which makes it inaccurate especially for large molecules, where only parts of the surface area are in contact with water. Therefore, there is need for the development of experimental probes to investigate accurately hydration properties and to understand hydrophobicity in aqueous mixtures.

Thermodiffusion, also called the Ludwig-Soret effect, has gained popularity in recent years as an analytical approach (Microscale Thermophoresis (MST)) to monitor binding reactions in biological molecules that are relevant in pharmaceutical applications[24]. The success of this technique relies on the superb sensitivity of thermodiffusion to changes in the hydration layer around a solute. The thermodiffusion response of a solute is quantified by the Soret coefficient, S_T, which is proportional to the concentration gradient that builds as a response to a thermal gradient. A positive Soret coefficient indicates that the solute accumulates on the cold side (thermophobic), while a negative sign denotes drift towards the warm side (thermophilic). The Soret coefficient has proven helpful in the investigation of other urea solutions (urea-pullulan solution), and correlations have been established between the magnitude and sign of the coefficient and the breaking and formation of hydrogen bonds[25]. The increase in the magnitude of the Soret coefficient with temperature has been correlated with the breaking of hydrogen bonds, since the latter is more favourable at high temperature. The temperature dependence of S_T can be modelled with an empirical equation proposed by Iacopini and Piazza[26], which fits accurately the behaviour observed in biological systems[27], aqueous solutions and suspensions. However, deviations from this temperature dependence have been reported in other aqueous solutions such as ethanol[28], ethylen glycol oligomers[29], or formamide at high concentration[30]. We will demonstrate that such deviations are present in the urea-water system, and that they appear in the concentration regime relevant to protein denaturation.

The magnitude of S_T depends in a complex way on the mass, shape, charge and concentration of the solute. Theoretical approaches, such as computer simulations provide an excellent tool to disentangle the impact that these variables have on thermodiffusion. Non-equilibrium molecular dynamics (NEMD) has advanced significantly in recent years.[31] This approach reproduces the general phenomenology of thermodiffusion in fluid mixtures and aqueous solutions, including the temperature dependence of the Soret coefficient on temperature[32–34]. NEMD provides a route to systematically study the impact that specific changes in the solute properties or solute-solvent interactions have on the thermodiffusion response[35]. We will use this approach to understand the microscopic origin of the thermodiffusion of urea-water mixtures.

In this work we exploit both experimental and theoretical techniques to advance our understanding of thermodiffusion of urea in water and to identify the onset of the hydrophobic behavior. We take full advantage of state of the art NEMD methods to interpret our experimental results and investigate the interactions between urea and water as a function of concentration and temperature. Advancing the discussion below, we will show that urea solutions feature both an increase and a decrease of the Soret coefficient with temperature depending on the urea concentration. We argue that this behavior signals a transition from the hydrophilic to the hydrophobic response of the solute.

Results and discussion

Figure 1(a) shows the temperature dependence of S_T at different urea concentrations. The positive sign of the Soret coefficient indicates that the urea solutions are thermophobic for all the concentrations investigated here. Our results agree with early experiments by Story and Turner[36]. Interestingly, urea solutions become more thermophobic as we increase the temperature for concentrations < 5.4 M. This behaviour is well described with Iacopini and Piazza's (IP) empirical equation (2). However, this trend is inverted at high urea concentrations (> 5.4 M), and the solution becomes less thermophobic with increasing temperature, signalling a breakdown of the IP equation. The inversion in the Soret dependence with temperature is reminiscent of the one reported for formamide solutions[30], which has a chemical structure similar to that of urea. In that case the change of slope is observed at a slightly lower concentration, ~ 20 wt% *vs.* ~ 30 wt% for urea.

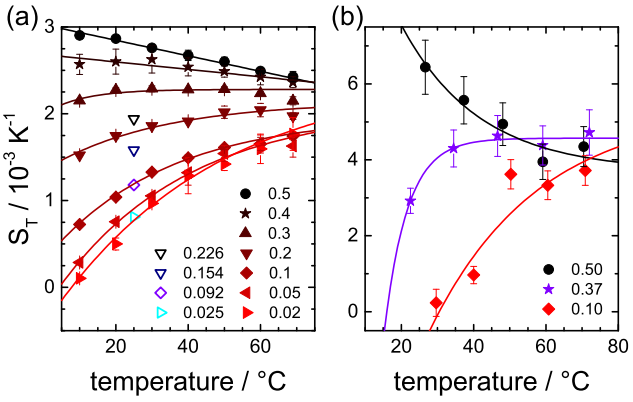

Fig. 1 Soret coefficients of urea-water mixtures obtained from (a) IR-TDFRS experiments and (b) NEMD simulations, as a function of temperature, for different urea concentrations, given in weight fractions (see legend). The open symbols in the left panel were reported by Story and Turner[36] using a different experimental method, thermogravitational columns. The solid lines are fittings of the numerical data (see Eq.2 in the Methods section) for concentrations < 30 wt% (~ 5.4 M) or to the exponential function $S_T = S_T^{\infty} + S_T^0 \cdot exp(-T/T_0)$ for concentrations > 30 wt% (~ 5.4 M). The fitting parameters are S_T^{∞} which represents S_T at high temperatures, S_T^0 which is the difference between S_T at 0^0C and that at high temperature $(S_T(0) - S_T^{\infty})$, and T_0 which indicates the temperature dependence of S_T.

The origin of the inversion behaviour is currently unknown. The concentrations at which we observe the inversion of the thermophoretic response of the urea/water and formamide/water mixture is within the concentration range leading to protein denaturation using co-solvent [37] (see SI for details).

We have performed computer simulations in order to provide a microscopic interpretation of the inversion effect discussed above. Remarkably, the Soret coefficients obtained from the NEMD simulations feature the same qualitative behaviour observed in the experiments (c.f. Figure 1(a) and (b)), even if some quantitative deviations are present. To best of our knowledge, this is the first theoretical demonstration of this inversion effect, hence supporting the generality of the experimental observations. We have performed additional "equilibrium" simulations in the canonical ensemble to investigate the hydrogen bond structure of four solutions: 10 wt%, 37 wt%, 50 wt% and 77 wt%. The Soret coefficient for the lower concentrations, 10 wt% and 37 wt%, increases with temperature, while it decreases for 50 wt%. We then expect that at higher urea concentration, 77 wt%, the thermophobic response is stronger and decreases with temperature. We show in Figure 2 the local solvation environment of urea and water molecules. We have plotted isodensity surfaces enclosing 70% of the hydrogen bonded atoms around the reference molecule. **For instance, W-U indicates the average hydrogen bonds between the reference molecule, water, and urea molecules.** The isodensity surfaces are highly directional, particularly the hydrogen density, signalling the formation of strong hydrogen bonds between urea and water molecules both at low (1.7 M, w.f. $= 0.10$) and high (16.0 M, w.f. $= 0.77$) urea concentration. The general shape of the isodensity surfaces depends little on the concentration of urea. At the highest concentration the structure resembles the one observed at low concentration, even if additional layers of atoms start to become evident (see hydrogen contribution for U-U and W-U in Figure 2(a)). These results confirm the perception that urea does not disrupt significantly the local solvation structure of water [1,4,6,8,12,13].

We show in Figure 2(b) the average number of hydrogen bonds for urea or water as a function of concentration. While pure urea is a solid, we can study in simulations the metastable liquid, given the high activation barrier for nucleation, the liquid will not freeze for typical simulation times (a few ns). The results reported for weight fraction w.f. = 1 correspond to this metastable liquid. The average number of hydrogen bonds (HBs) of water with water (W-W) or urea (W-U) decreases with increasing urea concentration. However, the decrease is not linear, we find a weaker dependence of the number of HBs at weight fractions < 0.4, and a steeper decrease above 0.5 w.f., with a transition region between weight fractions of 0.4 and 0.5. This concentration interval corresponds approximately with the w.f. at which we observe a change in the sign of the slope defining the dependence of the Soret coefficient with temperature (see Figure 1(b)). In the interval 0.4-0.5 w.f. the urea:water ratio changes from 1:5 (7.4 M) to 1:3.33 (9.5 M). At higher concentrations, the hydration shells of individual urea molecules will overlap, so that each water molecule has HBs with more than one urea molecule on average (W-U rises above 1). However, this is not reflected clearly in the three dimensional density profiles for water, which are very similar at all investigated urea concentrations.

Although the three-dimensional density profile offers insight into the local structure around urea an water molecules, it does not provide information on the overall structuring of the

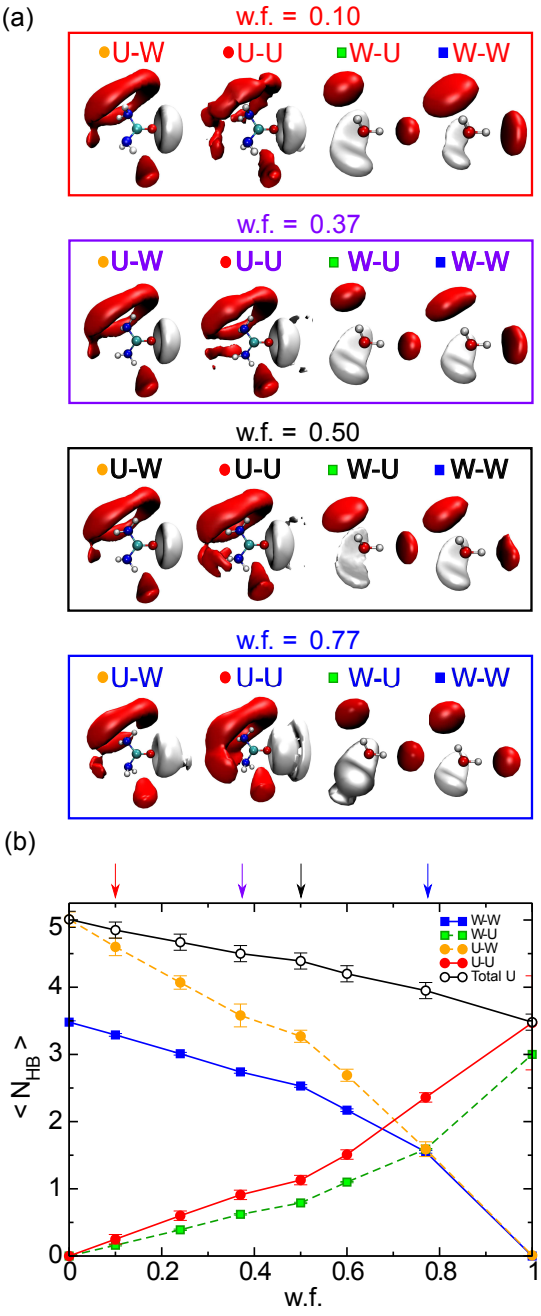

Fig. 2 (a) Isodensity surface enclosing 70% of the hydrogen bonded atoms around the reference molecule, namely urea (U-U and U-W) and water (W-U and W-W). A snapshot of the reference molecule is represented in each case. The red and white clouds indicate the oxygen and hydrogen densities, respectively, of molecules that are hydrogen bonded. (b) Number of hydrogen bonds N_{HB} formed in the urea and water mixture: U-W (yellow circle), U-U (red circle), W-U(violet square), W-W (blue square) and the total number of hydrogen bonds formed by urea (open circle). The arrows indicate the four different weight fractions represented in Panel (a).

molecules in solution. To analyse the urea aggregation, we have computed the number of clusters of urea molecules as a function of urea concentration at 30°C (see Supporting Information). We find few small urea clusters at the lowest concentration considered, 0.1 w.f, while increasing the concentration leads to an increase in the number and size of the clusters, which is consistent with the steeper increase of U-U bonds and the faster decrease of U-W reported in the number of hydrogen bonds (see Figure 2(b)).

We have analysed the thermodynamics of solvation of urea in water by computing the solvation chemical potential of urea as a function of temperature and urea concentration (see Fig. 3). The chemical potential varies in a linear fashion, increasing with temperature, hence indicating a decrease in solubility. Our results show that the variation of the chemical potential with temperature is driven by entropy, as the enthalpic term is essentially constant in the temperature interval investigated here. This behaviour does not match the characteristic dependence observed for hydrophobic solutes in water, *e.g.* methane in water, where both the enthalpy and entropy increase with temperature[38]. Increasing the urea concentration does not change the entropic and enthalpic dependences with temperature, but the chemical potential becomes less negative. The change in the chemical potential is driven mostly by the decrease in the enthalpic contribution, hence suggesting a weakening of the effective interactions of urea at high concentrations with respect to lower concentrations. These results immediately suggest a key role of the solute-solvent interactions in defining the thermodiffusive behaviour of the solutions. To make this connection concrete, we performed additional equilibrium simulations for a solution with w.f. $= 0.50$, and we increased the solvent-solute interactions by a factor of two, as compared with the original interactions. We refer to these new interactions as $U2$ to distinguish it from the original forcefield, $U1$. We expect that increasing the solute-solvent interactions will increase the chemical potential, and make the urea molecules more "hydrophilic", and possibly it will reduce the tendency for aggregation. We expect that a higher affinity towards the water would stall the aggregation and we should recover a behaviour closer to that observed for urea solutions at lower concentrations. At the highest concentration considered, 50 wt%, we observe that the increase on the water-urea interactions (U2) result in a decrease of both size and number of cluster (see Supporting Information). The results for the chemical potential confirm our hypothesis. Indeed, the chemical potential is reduced to values similar to those of the lower concentration, w.f. $= 0.1$ (see Figure 3). Crucially, this change is driven by the enthalpic contribution. As shown in Fig. 3 the entropy does not feature significant changes.

We now examine whether the change in urea-water interactions can explain the change in slope of the Soret coefficient with temperature. We show in Fig. 4(a) the dependence of the Soret coefficient with temperature for the two solutions, $U1$ and $U2$ at w.f. $= 0.50$. Clearly, stronger interactions between urea and water favour positive slopes, $dS_T/dT > 0$, as observed in solution with low urea concentrations, and hence we recover the behaviour predicted by the Iacopini-Piazza equation. These results highlight

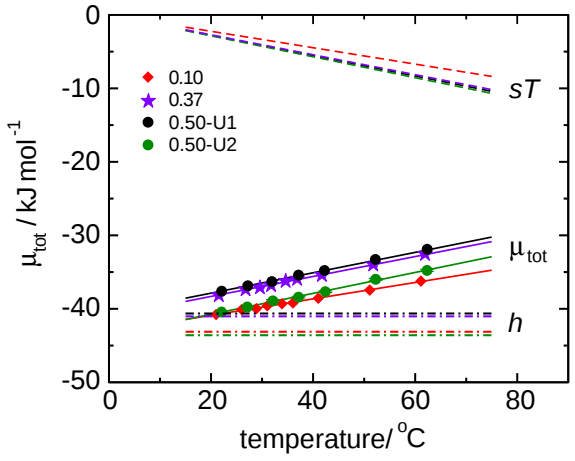

Fig. 3 Solvation free energy (μ_{tot}), enthalpy (h) and entropy (Ts) of urea in water, as a function of temperature, for three different urea weight fractions: 0.1, 0.37 and 0.50. For the highest urea concentration we increased the strength of the urea-water interactions by a factor of two.

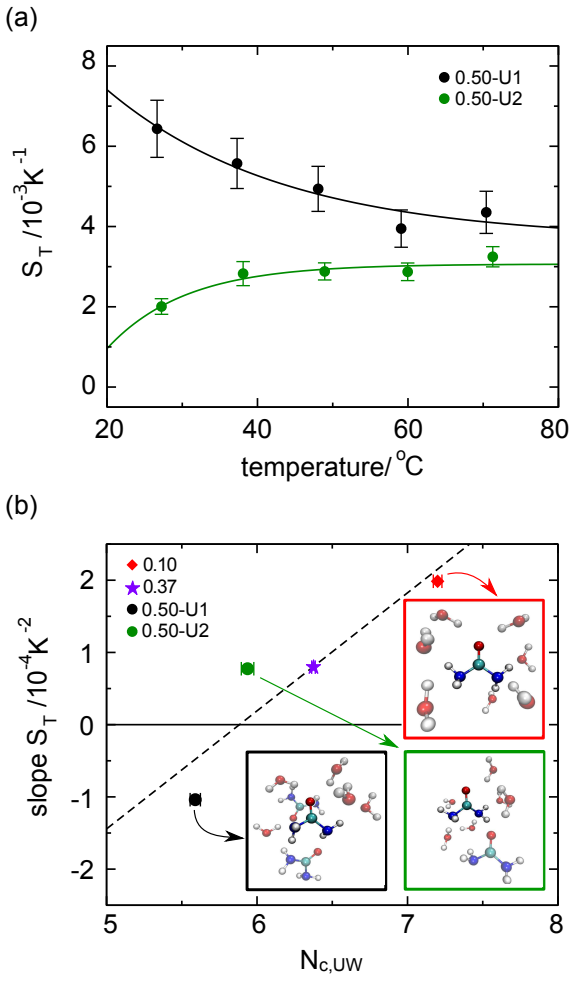

Fig. 4 (a) Dependence of the Soret coefficient of urea with temperature for two different urea-water interactions strengths. $U2$ is twice as strong as $U1$. (b) Slope of the Soret coefficient, dS_T/dT, as a function of the number of water molecules in the first solvation shell of the urea molecules. All the results correspond to a temperature of 30°C.

the key role that urea-water interactions play in determining the variation of the Soret coefficient with temperature. Our tentative conclusion is that a negative slope in the S_T vs T function is connected to unfavourable interactions between solute and water. We have made this hypothesis more concrete by computing the coordination number (see Methods) of the urea molecules in water as a function of the urea concentration. We find that low concentrations resulting in positive slopes, $dS_T/dT > 0$, involve larger urea-water coordination numbers, $N_{c,UW} = 7$. An increase in temperature makes the solution more thermophobic. The water is tightly bound to the urea molecules and the Soret coefficient follows the Iacopini-Piazza behaviour. Increasing the urea concentration reduces the coordination number to values approaching 5, for the higher concentrations, 50 wt%. This pentacoordination with water is consistent with the solvation structures inferred from pump-probe spectroscopy studies of solution at 8 M concentration[1]. At a comparable concentration, 50 wt% = 9.4 M, the system with stronger urea-water interactions ($U2$) features a higher coordination number and one can see from Figure 4(b) that $dS_T/dT > 0$ is consistently shifted to slightly positive values, similar to those obtained for a solution with lower concentration that has approximately the same coordination number for urea-water correlations. Hence, we conclude that hydrophilic solutes become more thermophobic at high temperatures.

The absolute value of S_T depends on a number of parameters making difficult a direct comparison between different compounds. However, as noted above, the temperature dependence of S_T is determined by the interactions between the solute and water. The correlation between hydrophilicity and the sign of $dS_T/dT < 0$ is consistent with empirical correlations between the Soret coefficient's *temperature sensitivity*, $\Delta S_T = S_T(50^\circ C) - S_T(20^\circ C)$, which defines the sign of dS_T/dT, and the degree of hydrophilicity, through $\log P$, of different solutes at similar concentrations (see Figure 5). Note, however, that the partition coefficient $\log P$ is defined only for a dilute regime, where solute-solute interactions are negligible.[39] The concentration of the solute molecules in these measurements was fairly low with 1 wt% (cyclodextrins), 5 wt% (urea and formamide), 10 wt% (oligossacharides) and 20 wt%(glucose). Figure 5 shows clearly an increase of the temperature sensitivity, $\Delta S_T > 0$, with increasing hydrophilicity (corresponding to a more negative $\log P$) of the solute molecules. For lower hydrophilicity, the increase of ΔS_T with decreasing $\log P$ is roughly linear, while for strongly hydrophilic compounds, the temperature sensitivity of S_T saturates. This behaviour can be understood, if we assume that the solute-solvent interactions approach a maximum at high hydrophilicities and a further reduction of $\log P$ does not influence ΔS_T as strongly. We show in Figure 5 that urea is at the bottom of the temperature sensitivity scale, hence indicating a higher sensitivity to changes in the solution composition. This notion is consistent with our simulation results, namely, the sign of dS_T/dT can be tuned by modifying the water-urea interactions.

Conclusion

Thermal diffusion measurements of urea/water solutions feature distinctive changes in the temperature dependence of the Soret coefficient: positive slopes at low concentrations ($dS_T/dT > 0$), and negative slopes at high concentrations ($dS_T/dT < 0$). The transition between these two regimes is observed in urea solutions at a concentration around 30 wt% (5.4 M). Our computer simulations indicate that the concentration defining the transition is determined by the solvent-solute interactions. We expect that strong interactions will shift the transition to higher concentrations. The reversal in the sign of dS_T/dT is also observed in thermal diffusion experiments of formamide, although it appears at lower concentrations (around 20 wt%)[30]. This observation can be rationalized using the Soret/partition coefficient correlation discussed in this work (see Figure 5), which shows that formamide is less hydrophilic than urea, and consistently features a transition at lower concentrations. Consistently moving toward lower $\log P$ values we expect that the transition will be increasingly difficult to observe, in fact it should not be observed for very hydrophilic substances. Microscopically, the simulations show that the reversal in the sign of dS_T/dT in urea is driven primarily by the urea-water interactions, and this phenomenon is mostly enthalpic in origin, with a lower influence of the entropic contribution.

We have shown that thermodiffusion provides an extremely sensitive probe to investigate hydration of solutes in water. Our finding that the hydrophilicity of urea decreases with urea concentration may be relevant in the context of the “dry globule" formation, as the first step of the denaturation mechanism proposed by Hua *et al.*[14]. These authors reported an increase in the concentration of urea in the first solvation shell of lysozyme, from 8 M up to 13 M, which results in the subsequent unfolding of the protein. It has been shown that the solvation free energies of non-polar amino acids decrease when moving the solute from pure water to urea solutions[43], and as a consequence the weaker hydrophobic interactions in urea solutions could contribute to the protein denaturation effect[44]. Our observation that the hydrophilicity of urea decreases with rising concentration suggests that there is a threshold concentration of urea where its hydrophilicity sinks to a point where it becomes prone to aggregation around a protein, because of favourable van der Waals interactions. Our observations prompt us to propose a model to explain the abrupt denaturation of proteins in a relatively narrow concentration range: as long as the urea is present at low concentration and it is relatively hydrophilic, it will remain well dissolved in water. However, when the concentration rises, the hydrophilicity decreases down to a specific threshold where the aggregation around the protein would become more favourable for urea. We propose that this threshold can be monitored using thermodiffusion, and that this technique allows to identify the transition from hydrophilic to hydrophobic behaviour. The specific concentration at which the denaturation is observed will depend also on the protein, which might explain why the denaturation concentration of urea changes for different proteins. Thermodiffusion provides a route to test these ideas and further investigations are needed along these lines to validate this picture.

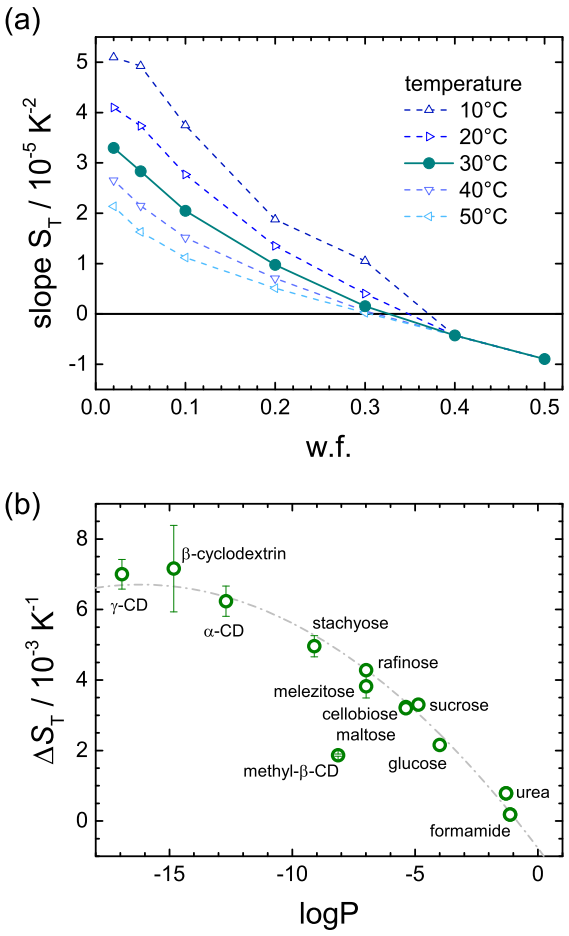

Fig. 5 (a) Slope of the Soret coefficient, dS_T/dT, as a function of the urea weight fraction for different temperatures, as specified in the legend. (b) $\Delta S_T = S_T(50°C) - S_T(20°C)$ against $\log P$ for cyclodextrins[40], oligosaccharides[41,42], formamide[30] and urea at low concentrations (see text for details).

Methods and Materials

Theoretical background of thermophoresis

Thermodiffusion is a physical effect whereby thermal gradients induce concentration gradients and particle motion[45]. The thermodiffusion of a binary liquid mixture can be described in terms of the mass flux $\vec{j}$ with a contribution $\sim -D_T\vec{\nabla}T$ from the thermodiffusion along the temperature gradient and $-D\vec{\nabla}c$ from the Fickian diffusion along the resulting concentration gradient:

$$\vec{j} = -D\vec{\nabla}c - c(1-c)D_T\vec{\nabla}T\ . \qquad (1)$$

In equilibrium, a steady state is reached with $\vec{j}=0$. The ratio of the concentration gradient that occurs over the applied temperature gradient is proportional to the ratio of the thermal diffusion coefficient D_T and the mass-diffusion coefficient D and characterized by the Soret coefficient $S_T = D_T/D$. The sign of S_T indicates the direction of the concentration gradient ($S_T > 0$: particles move from high to low temperature), a larger amplitude implies a larger concentration gradient resulting from a given temperature gradient.

Iacopini and Piazza[46] proposed an empirical equation to model the temperature dependence of the Soret coefficient,

$$S_T(T) = S_T^{\infty}\left[1 - \exp\left(\frac{T^* - T}{T_0}\right)\right]\ . \qquad (2)$$

The parameters S_T^{∞}, T^*, and T_0 can be adjusted to fit experimental data. Note that S_T^{∞} represents the limiting value of S_T at high temperatures, and T^* the temperature at which the Soret coefficient changes sign, *i.e.* $S_T = 0$. Equation (2) describes accurately the thermodiffusion of macromolecules in dilute aqueous solutions[26,27,47], but fails with low-molecular-weight mixtures, usually at high concentrations[30], and also in the dilute regime for specific systems, such as ethanol/water[28].

IR-TDFRS measurements

Thermodiffusion was measured using the Infra-Red Thermal Diffusion Forced Rayleigh Scattering technique (IR-TDFRS). This is a laser-induced transient grating technique that has been described in detail elsewhere[48,49]. Urea ($\geq$ 99%, Fluka, Sigma-Aldrich Steinheim, Germany) was dissolved in Millipore water and filtered through an 0.8/0.2 μm membrane filter (PALL Acrodisc PF). Measurements were done at several concentrations in the range from 2 to 50 wt% at temperatures between 10 and 70°C. Details on the refractive index contrast measurements necessary to evaluate the TDFRS data are provided in the Supporting Information.

Non Equilibrium Molecular Dynamics simulations

Thermal gradients were applied in urea aqueous solutions using Non-Equilibrium Molecular Dynamics simulations (NEMD) using the method discussed in references[34,50]. We used a modified version of the code GROMACS v. 4.6.3[51]. A typical snapshot of the system is reported in Figure 6 in the Supporting Information, along with a representative temperature profile. Two thermostat regions (hot and cold) were defined in the center and at the edges of the simulation box. The hot and cold temperatures were set by rescaling every time step the velocity of the molecules of water using the v-rescale algorithm[52]. At the start of the simulation the molecules lying in the thermostatted layers were restrained in the z-direction (the heat flux direction) using a harmonic potential, while the molecules were free to rotate and translate in the xy plane. The molecules in between the hot and cold thermostatted regions were not thermostatted directly. These molecules move freely through the simulation box and exchange momentum with the restrained molecules, hence allowing the set up of the thermal gradient. The simulation cell consisted of a prismatic box with vectors $\{L_x, L_y, L_z\}/L_x = \{1, 1, 5\}$, with $L_x = 2.78$ nm. Three different urea concentrations were simulated, 10 wt% (1.7 M), 37 wt% (6.7 M) and 50 wt% (9.4 M) urea weight fractions, using 3305-1000 molecules of water and 112-325 molecules of urea. The simulation trajectories were integrated using the *Leap-Frog* algorithm with a time step of 2 fs.

A typical simulation involved a 5 ns equilibration in the NPT ensemble at 1 bar and temperature, $T = (T_{COLD} + T_{HOT})/2$, where T_{COLD} and T_{HOT} are the target temperatures in the NEMD simulation. The equilibrated configurations were then employed in

the NEMD computations. Hot and cold layers of width $\simeq 0.1$ nm were set in the center and at the edges of the simulation cell. The molecules in the HOT and COLD layers were thermostatted every time step. An equilibration period of 5 ns with the applied thermal gradient was performed to allow the system to reach the stationary state. We ensured that the pressure in the NEMD simulation matched 1 bar, by removing when necessary water and urea molecules to keep the desired concentration. The production runs involved simulation times of 2.7μs. The trajectories were employed to construct the temperature, density and concentration profiles, by dividing the simulation box in 100 sampling volumes of width ~ 0.14 nm, along the direction of the thermal gradient, z. The temperature was calculated using the equipartition principle by sampling the velocities of molecules,

$$T_k(R) = \frac{1}{N_t N_{df} N_R} \sum_{i=1}^{N_t} \sum_{k}^{N_R} \sum_{j=1}^{N \in R} \frac{m_{ikj} v_{ikj}^2}{k_B} \tag{3}$$

where N the number of molecules a given sub-region, R, N_R the total number of sub-regions and N_t the total number of configurations analysed, k_B the Boltzmann constant, and N_{df} is the number of degrees of freedom; 6 and 17 for a water and a urea molecule, respectively.

The Soret coefficient was computed from the simulated number and temperature profiles x_u:

$$S_T = \frac{-1}{x_u(1-x_u)} \left(\frac{\nabla x_u}{\nabla T} \right)_{J=0}. \tag{4}$$

The Soret coefficient is computed locally along the simulation box at the stationary state, namely when the mass flux $\vec{j} = 0$ (see *e.g.* ref.[53] for a test of the calculation of local properties). We report in the supplementary information representative particle number profiles for water and urea. We employed thermal gradients of the order of $\nabla T \sim 12.4\ K\ nm^{-1}$. For this thermal gradient the system is in the linear regime.

The water molecules were modelled using the extended simple point charge model (SPC/E)[54]. For the urea molecules we used the GROMOS 53A6 parameter set. We tested the accuracy of our forcefields by computing the density if urea solutions for all the concentrations considered in this paper. We found good agreement with both experimental and computed data (see Figure 4 in Supporting Information). The Lennard-Jones interactions were truncated at a cut-off radius of $r_c = 1.0$ nm and the long range electrostatic interactions were computed using the particle-mesh Ewald method (PME) with a mesh width of 0.12 nm.

We computed the coordination number counting the average number of molecules of urea in the first urea hydration shell of the water molecules, in a radius of about 0.42 nm from the center of mass of the water molecules.

Free energy computations

The chemical potentials obtained here represent the work associated to adding a molecule of urea in the aqueous solution, at a specific urea concentration. We follow the procedure used in reference[34] to calculate the chemical potential of ions. The chemical potential is defined by the sum of an ideal and excess terms. The ideal term is defined by $\mu_{id} = k_B T \ln\left(\frac{N_u k_B T}{P\langle V\rangle}\right)$, which corresponds to the ideal gas contribution containing N_u urea molecules at pressure, P, temperature, T and average volume $\langle V\rangle$. The excess chemical potential was obtained using a perturbative approach that involves the computation of the free energy needed to "grow" one urea molecule in a solution with a specific weight fraction. The perturbation method interpolates between two states, *State 1* ($\mathscr{S}1$), where the extra urea molecule is absent, and *State 2* ($\mathscr{S}2$) where it is present. We divide the path between $\mathscr{S}1$ and $\mathscr{S}2$ in different sub-states. The van der Waals interactions are tuned on progressively using 20 coupling parameters between $\lambda_{vdw} = 0 \cdots 1$, where 0 and 1 correspond to zero and full van der Waals interactions, respectively. Subsequently we charged the molecules using 20 coupling parameters, $\lambda_c = 0 \cdots 1$, where again 1 correspond to the fully charged molecule. We computed the van der Waals and Coulombic interactions in two states. First we obtained the excess in chemical potential of adding a Lennard-Jones particle with no charge μ_{vdw}, followed by the computation of the chemical potential of charging it, μ_c. The total chemical potential was computed by adding the van der Waals and Coulombic contributions, $\mu_{ex} = \mu_{vdw} + \mu_c$. For each λ, the simulations were performed at constant temperature and pressure using a time step of 2 fs. A typical simulation involved a 5 ns equilibration period, followed by 20 ns of production. We discarded the first 1 ns of the trajectories before computing the averages. The temperature was controlled using the v-rescale thermostat with a time constant of 0.1 ps while the pressure was maintained at 1 bar using the Parrinello-Rahman, with time constant 1 ps. The solvation free energy computations were performed for three different aqueous solutions with weight fractions, 10 wt% (1.7 M), 37 wt% (6.7 M) and 50 wt% (9.4 M) and for different temperatures in the range 22 and 62 ^{0}C. The number of urea and water molecules varied between 15–92 and 280–524, respectively, in order to model the desired weight fraction. These simulations were performed in cubic boxes with volumes 14.95-23.19 nm^3.

The entropic contribution to the solvation free energy of urea μ_{tot} was obtained from the temperature derivative of the chemical potential,

$$s(T) = -\left(\frac{\partial \mu_{tot}(T)}{\partial T} \right)_P, \tag{5}$$

and the enthalpic contribution to the solvation free energy of urea from the equation,

$$h(T) = \mu_{tot}(T) + T s(T). \tag{6}$$

To compute the solvation entropy and enthalpy, we fitted our chemical potential to a linear equation which accurately reproduce the computed data. The fitting parameters are reported in the Supporting Information.

Hydrogen bonds

Water-water (W-W), urea-urea (U-U) and water-urea (W-U, U-W) molecules were deemed to be hydrogen bonded if the oxygen-oxygen distances were shorter than 0.35 nm, 0.21 nm or 0.25 nm, respectively, and if the angle between the vector connecting the

oxygen atoms of the two molecules and the vector connecting the oxygen and the hydrogen in the same molecule was smaller than 30°.

Acknowledgement

We thank Jan Dhont for his generous support and interest of our (D.N. and S.W.) work. We thank the EPSRC-UK (EP/J003859/1) for financial support and acknowledge the Imperial College High Performance Computing Service for providing computational resources. Calculator Plugins were used for structure property prediction and calculation of $\log P$, Marvin 16.5.2.0, 2016, ChemAxon (http://www.chemaxon.com).

Conflict of interest

There are no conflicts to declare.

References

1 Y. Rezus and H. Bakker, *Proc. Natl. Acad. Sci. U.S.A.*, 2006, **103**, 18417–18420.
2 A. Idrissi, M. Gerard, P. Damay, M. Kiselev, Y. Puhovsky, E. Cinar, P. Lagant and G. Vergoten, *J. Phys. Chem. B*, 2010, **114**, 4731–4738.
3 P. Jedlovszky and A. Idrissi, *J. Chem. Phys.*, 2008, **129**, 164501.
4 A. Kuffel and J. Zielkiewicz, *J. Chem. Phys.*, 2010, **133**, 035102.
5 K. Nakanishi, *Chem. Soc. Rev.*, 1993, **22**, 177–182.
6 V. Agieienko and R. Buchner, *Phys. Chem. Chem. Phys.*, 2016, **18**, 2597–2607.
7 C. A. Bertran, J. A. J. V. Cirino and L. C. G. Freitas, *J. Braz. Chem. Soc.*, 2002, **13**, 238 – 244.
8 Y. Hayashi, Y. Katsumoto, S. Omori, N. Kishii and A. Yasuda, *J. Phys. Chem. B*, 2007, **111**, 1076–1080.
9 L. B. Sagle, Y. J. Zhang, V. A. Litosh, X. Chen, Y. Cho and P. S. Cremer, *J. Am. Chem. Soc.*, 2009, **131**, 9304–9310.
10 N. Samanta, D. Das Mahanta and R. K. Mitra, *Chem. Asian J.*, 2014, **9**, 3457–3463.
11 M. C. Stumpe and H. Grubmüller, *J. Phys. Chem. B*, 2007, **111**, 6220–6228.
12 D. Bandyopadhyay, S. Mohan, S. K. Ghosh and N. Choudhury, *J. Phys. Chem. B*, 2014, **118**, 11757–11768.
13 A. K. Soper, E. W. Castner and A. Luzar, *Biophys. Chem.*, 2003, **105**, 649–666.
14 L. Hua, R. Zhou, D. Thirumalai and B. Berne, *Proc. Natl. Acad. Sci. U.S.A.*, 2008, **105**, 16928–16933.
15 M. Auton, L. Holthauzen and D. Bolen, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**, 15317–15322.
16 A. Das and C. Mukhopadhyay, *J. Phys. Chem. B*, 2009, **113**, 12816–12824.
17 R. Zhou, J. Li, L. Huan, Z. Yang and B. Berne, *J. Phys. Chem. B*, 2011, **115**, 1323–1326.
18 B. Bennion and V. Daggett, *Proc. Natl. Acad. Sci. U.S.A*, 2003, **100**, 5142–5147.
19 D. Tobi, R. E. R and D. Thirumalai, *Biopolymers*, 2003, **68**, 359–369.
20 P. Astrand, A. Wallqvist and G. Karlstrom, *J. Phys. Chem.*, 1994, **98**, 8224–8233.
21 A. Shimizu, K. Fumino, K. Yukiyasu and Y. Taniguchi, *J. Mol. Liq.*, 2000, **85**, 269–278.
22 P. Ball, *Nature*, 2008, **108**, 74–108.
23 T. Cheng, Y. Zhao, X. Li, F. Lin, Y. Xu, X. Zhang, Y. Li and R. Wang, *J. Chem. Inf. Model.*, 2007, **47**, 2140–2148.
24 M. Jerabek-Willemsen, T. André, W. Wanner, H. Roth, S. Duhr, P. Baaske and D. Breitsprecher, *J. Mol. Struct.*, 2014, **1077**, 101–113.
25 B. W. R. Sugaya and R. Kita, *Biomacromolecules*, 2006, **7**, 435–440.
26 S. Iacopini, R. Rusconi and R. Piazza, *Eur. Phys. J. E*, 2006, **19**, 59–67.
27 Y. Kishikawa, S. Wiegand and R. Kita, *Biomacromolecules*, 2010, **11**, 740–747.
28 A. Königer, B. Meier and W. Köhler, *Philos. Mag.*, 2009, **89**, 907–923.
29 K. Maeda, N. Shinyashiki, S. Yagihara, S. Wiegand and R. Kita, *J. Chem. Phys.*, 2015, **143**, 124504.
30 D. Niether, D. Afanasenkau, J. K. G. Dhont and S. Wiegand, *Proc. Natl. Acad. Sci. U.S.A.*, 2016, **113**, 4272–4277.
31 F. Bresme, A. Lervik and J. Armstrong, in *Experimental Thermodynamics Volume X: Non-equilibrium Thermodynamics with Applications*, The Royal Society of Chemistry, 2016, pp. 105–133.
32 C. Nieto-Draghi, J. Avalos and B. Rousseau, *J. Chem. Phys.*, 2005, **122**, 114503.
33 F. Römer, Z. Wang, S. Wiegand and F. Bresme, *J. Phys. Chem. B*, 2013, **117**, 8209–8222.
34 S. Di Lecce, T. Albrecht and F. Bresme, *Sci. Rep.*, 2017, **7**, 44833.
35 S. Di Lecce, T. Albrecht and F. Bresme, *Phys. Chem. Chem. Phys.*, 2017, **19**, 9575–9583.
36 M. J. Story and J. C. R. Turner, *Trans. Faraday Soc.*, 1969, **65**, 1810–1811.
37 A. Tischer and M. Auton, *Protein Sci.*, 2013, **22**, 1147.
38 H. Ashbaugh and L. Pratt, *Rev. Mod. Phys.*, 2006, **78**, 159.
39 J. Sangster, *Octanol-Water Partition Coefficients: Fundamentals and Physical Chemistry*, Wiley, 1997.
40 K. Eguchi, D. Niether, S. Wiegand and R. Kita, *Eur. Phys. J. E*, 2016, **39**, 86.
41 P. Blanco and S. Wiegand, *J. Phys. Chem. B*, 2010, **114**, 2807–2813.
42 P. Blanco, H. Kriegs, B. Arlt and S. Wiegand, *J. Phys. Chem. B*, 2010, **114**, 10740–10747.
43 Y. Nozaki and C. Tanford, *J. Biol. Chem.*, 1963, **238**, 4074.
44 Z. Su and C. L. Dias, *J. Mol. Liq.*, 2017, **228**, 168.
45 S. de Groot and P. Mazur, *Non-Equilibrium Thermodynamics*, New York: Dover, 1984.
46 S. Iacopini and R. Piazza, *Europhys. Lett.*, 2003, **63**, 247–253.
47 Z. Wang, H. Kriegs, J. Buitenhuis, J. Dhont and S. Wiegand, *Soft Matter*, 2013, **9**, 8697–8704.

48 S. Wiegand and W. Köhler, *LNP Vol. 584: Thermal Nonequilibrium Phenomena in Fluid Mixtures*, 2002, **584**, 189–210.
49 P. Blanco, H. Kriegs, M. P. Lettinga, P. Holmqvist and S. Wiegand, *Biomacromolecules*, 2011, **12**, 1602–1609.
50 F. Römer, A. Lervik and F. Bresme, *J. Chem. Phys.*, 2012, **137**, 074503.
51 B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.
52 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.
53 F. Bresme and J. Armstrong, *J. Chem. Phys.*, 2014, **140**, 016102.
54 H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *J. Phys. Chem.*, 1987, **91**, 6269–6271.