Thermophoresis of biological and biocompatible systems

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Doreen Niether

aus Eberswalde-Finow

Köln, 2018

Abstract

Thermophoresis, or thermodiffusion, is mass transport driven by a temperature gradient. This work focuses on thermodiffusion in a biological context, where there are two major applications for the effect: accumulation of a component in microfluidic devices through a combination of thermodiffusion and convection, and monitoring of protein binding reactions through the sensitivity of thermodiffusion to complex formation. Both applications are investigated, the first as an accumulation process in the context of origin-of-life theories and the second in light of the question what we can learn from the observed changes in thermodiffusion about modifications of the hydration shell upon complex formation. While thermodiffusion in non-polar liquids can be predicted with reasonable accuracy, the description of aqueous systems is complicated as their concentration and temperature dependence is often anomalous. The underlying goal of this work is to gain a better understanding of the interactions between components in an aqueous mixture and how they influence thermodiffusion.

We find that the temperature dependence of a solute's thermodiffusion correlates with its hydrophilicity and argue that the temperature sensitivity of hydrogen bonds, which dominate the interactions in aqueous solutions, might induce a temperature dependence of the chemical potential. Such a temperature dependence is as of yet not considered in theoretical descriptions of thermodiffusion. Numerical calculations show that the thermophoretic accumulation process, as of yet only considered for the formation of RNA, can accumulate formamide to high concentrations that would allow the formation of prebiotic molecules. A heuristic model is developed to illuminate the mechanism behind the accumulation. Cyclodextrins and streptavidin were investigated as model systems for biological complexes. It is feasible that the exquisite sensitivity of thermodiffusion to interactions with the surrounding solvent allows inferences about changes in the protein's hydration shell upon complex formation. Preliminary measurements on streptavidin-biotin show a decreased hydrophilicity of the complex, which is in qualitative agreement with increased entropy of the hydration shell upon complex formation calculated from calorimetric and neutron scattering experiments.

Kurzzusammenfassung

Thermophorese, oder Thermodiffusion, ist Massentransport, der durch einen Temperaturgradienten hervorgerufen wird. Für Thermodiffusion in einem biologischen Kontext, auf welcher der Fokus dieser Arbeit liegt, gibt es im Wesentlichen zwei Anwendungen: die Akkumulation einer Komponente in mikrofluidischen Systemen durch eine Kombination aus Thermodiffusion und Konvektion und die Detektion von Bindungsreaktionen über die Veränderung in der Thermodiffusion eines Proteins, wenn ein Ligand bindet. Beide Anwendungen werden hier untersucht, Erstere als Anreicherungsprozess im Kontext von Theorien zur Entstehung des Lebens und Letzere im Bezug auf die Frage, inwieweit die beobachteten Änderung der Thermodiffusion Rückschlüsse auf Modifikation der Hydrathülle durch die Komplexbildung zulassen. Während die Thermodiffusion von unpolaren Flüssigkeiten inzwischen im Wesentlichen vorhersagbar ist, wird die Beschreibung von wässrigen Systemen durch Anomalien der Konzentrations- und Temperaturabhängigkeit verkompliziert. Das grundlegende Ziel dieser Arbeit ist deshalb, ein besseres Verständnis der Wechselwirkungen zwischen Komponenten einer wässrigen Lösung und ihres Einflusses auf die Thermodiffusion zu ermöglichen.

Es konnte gezeigt werden, dass es eine Korrelation zwischen der Temperaturabhängigkeit der Thermodiffusion eines gelösten Stoffes und seiner Hydrophilie gibt. Dies kann in der Temperaturempfindlichkeit von Wasserstoffbrückenbindungen begründet sein, welche die Wechselwirkungen in wässrigen Systemen dominieren und zu einer Temperaturabhängigkeit des chemischen Potentials führen können. Eine solche Temperaturabhängigkeit wird bisher in theoretischen Beschreibungen der Thermodiffusion vernachlässigt. Numerische Rechnungen zeigen, dass der thermophoretische Anreicherungsprozess, welcher bislang nur unter dem Gesichtspunkt der Bildung von RNA betrachtet wurde, auch Formamid zu hohen Konzentrationen akkumulieren kann, welche die Bildung von präbiotischen Molekülen zulassen würden. Es wurde ein heuristisches Modell entwickelt, um den Mechanismus der Akkumulation zu erläutern. Cyclodextrine und Streptavidin wurden als Modellsysteme für biologische Komplexe untersucht. Die hohe Empfindlichkeit der Thermodiffusion gegenüber den Wechselwirkungen mit dem umgebenden Lösungsmittel sollte Aussagen über Veränderungen in der Hydrathülle des Proteins durch Ligandenbindung zulassen. Vorläufige Untersuchungen an Streptavidin-Biotin zeigen eine reduzierte Hydrophilie des Komplexes, was in qualitativer Übereinstimmung mit einer Erhöhung des entropischen Beitrags der Hydrathülle ist, wie er aus kalorimetrischen und Neutronenstreu-Experimenten berechnet wurde.

Contents

1 Introduction

1.1 Introduction to thermodiffusion

Thermodiffusion is mass transport driven by a temperature gradient. It is also known as thermophoresis or Ludwig-Soret-effect and was first reported by Carl Ludwig in 1856 [1]. Over twenty years later, it was systematically investigated in electrolyte solutions by Charles Soret [2]. Soret developed phenomenological equations describing the thermodiffusion based on Fick's diffusion equations. In a binary mixture it can be described as a mass flux $\vec{j}$ along a temperature gradient ∇T with

$$\vec{j} = -\rho D \nabla c - \rho c (1-c) D_\mathrm{T} \nabla T, \tag{1.1}$$

where D_T is the thermal diffusion coefficient, ρ is the mass density and c is the concentration given as mass fraction [3]. Along the concentration gradient ∇c that arises from the thermodiffusion, Fickian diffusion takes place characterised by the diffusion coefficient D. With a stable temperature gradient, a steady state is reached where Fickian and thermodiffusion flux cancel each other out and $\vec{j}$ becomes zero. Then the concentration gradient over the temperature gradient is proportional to a constant value

$$S_\mathrm{T} \equiv \frac{D_\mathrm{T}}{D} = -\frac{1}{c(1-c)} \frac{\Delta c}{\Delta T}, \tag{1.2}$$

which is defined as the Soret coefficient S_T with a unit of K^{-1}. The thermophobic component of the mixture, which enriches at the cold side, has a positive Soret coefficient, while the thermophilic one enriches on the warm side and has a negative S_T.
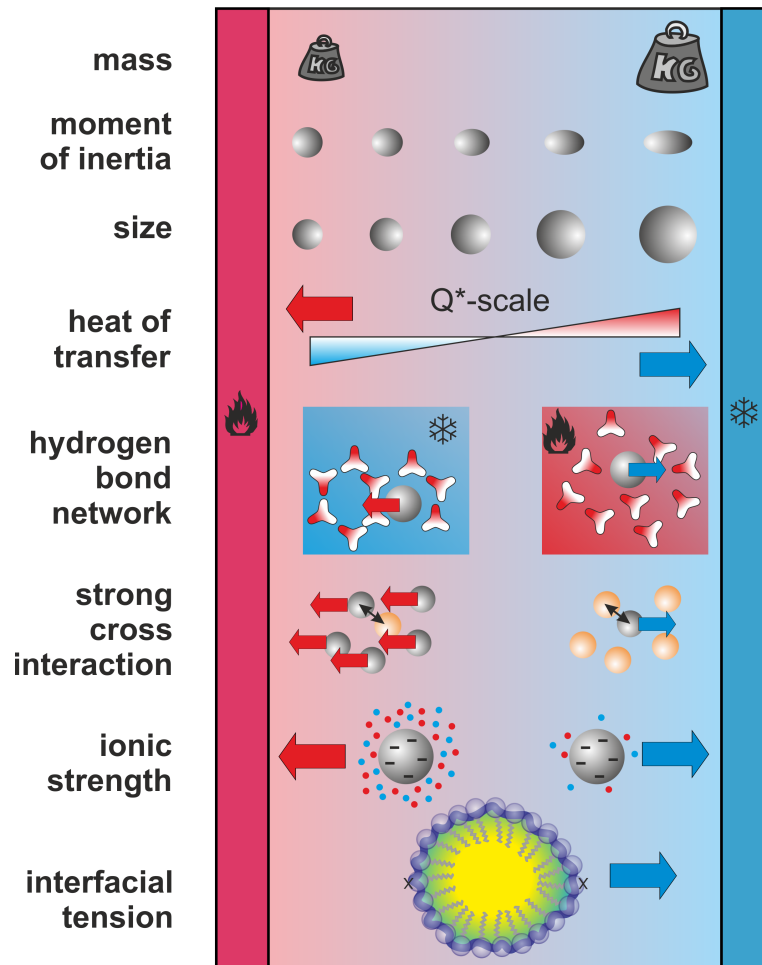
Figure 1.1: *Contributions to the Soret effect.*

1.1.1 Contributions to the Soret effect

Thermodiffusion has been observed in mixtures of any kind: in homogenous gases, liquids and solids as well as in heterogeneous mixtures like gas bubbles in liquids, colloids and aerosols [4,5]. Experimental findings show that the thermodiffusion behaviour of a substance is sensitive to a large number of parameters. The main factors are illustrated in Fig. 1.1.

The influence of a particle's mass, size and moment of inertia are the only parameters that play a role in an ideal gas. These three factors contribute to the so-called *isotope Soret effect*. The name derives from experiments that investigated isotope mixtures in order to observe species of different mass that are chemically identical [6, 7]. The meaning of the term has been widened since, as it could be shown that there is a fixed contribution to the

Soret coefficient that is only dependent on the differences of mass M_i and moment of inertia I_i between the components i of a non-polar mixture [8,9]. However, to avoid confusion, the *isotope Soret effect* will be termed mass-inertia contribution to the Soret coefficient S_T^m in the following text. Equations to characterise S_T^m show some variations [4], most common is

$$S_T^m = a_M \frac{M_1 - M_2}{M_1 + M_2} + b_I \frac{I_1 - I_2}{I_1 + I_2}, \tag{1.3}$$

which is derived from descriptions of gaseous mixtures and adjusted for liquids by the factors a_M and b_I [10].

In non-ideal gases and liquids, the interactions between particles have to be considered, which demands complex theoretical models. Especially in the case of aqueous systems, strong hydrogen bonds cause an anomalous concentration and temperature dependence of molecular interactions that impedes predictions of the thermodiffusion behaviour. These interaction contributions, because they depend on the chemical nature of the components, are called the chemical contribution to the Soret effect S_T^{chem} [8]. The observed Soret coefficient is a sum of these two contributions:

$$S_T = S_T^m + S_T^{chem}. \tag{1.4}$$

Even for mixtures of noble gases a temperature dependence of the thermodiffusion can be observed [11]. The thermal diffusion factor $\alpha = S_T \cdot T$ is then theoretically described by

$$\alpha = \alpha_{hard} \cdot R_T, \tag{1.5}$$

where α_{hard} is the theoretical value for hard spheres and R_T a correction factor ($R_T < 1$) that contains the temperature dependence by its proportionality to an interaction potential. Using a Lennard-Jones potential, the theoretical description fits the experimental data qualitatively. It has been pointed out that this temperature dependence might be one of the strongest tests for interaction potentials [12].

Note that Grew *et al.* [11] found for the noble gas mixtures that α increases with rising temperature. Prigogine *et al.*, on the other hand, found for alcohols in organic solvents that $S_T \propto 1/T^2$, implying $\alpha \propto 1/T$ [13]. Obviously, the investigated systems in both cases are vastly different: in the case of the noble gases S_T^{chem} is caused by a certain permeability of the

gas atoms, in the organic mixtures the chemical contribution is due to specific interactions caused by polarity of the molecules. This just illustrates that, as already implied by Fig. 1.1 the chemical contribution S_{T}^{chem} can be a sum of different effects specific to the investigated system.

In the following section a short overview of thermodiffusion models is presented.

1.1.2 Thermodynamic thermodiffusion models

In the case of ideal gases the kinetic theory of gases can be applied to predict thermodiffusion of the heavier component towards the cold side [14]. Thermodiffusion in solids can be described by considering the heat of transport for the jump from a neighbouring lattice position into a vacancy caused by crystal defects [15]. For the case of liquids and non-ideal gases, where inter-particle interactions are neither zero, nor as well-defined as in a crystalline solid, the microscopic understanding of thermodiffusion is still lacking. Existing theoretical models are based in kinetic theory, non-equilibrium thermodynamics and combinations of both. They use interaction potentials or heats of transport to account for interactions, but due to inaccuracy of these expressions they are not reliable in predicting thermodiffusion behaviour, especially for polar systems.

As the last paragraph already suggests, there are two principle approaches to describe the thermodiffusion of liquids: (1) one can start from an ideal gas and add terms to include molecular interactions, or (2) one can start from the non-equilibrium framework that describes crystals and try to define some analogy to the jumps between lattice positions that might apply to a liquid in order to arrive at a heat of transport. The interaction strength is estimated by correlation with equilibrium or dynamic characteristics of the mixture.

As thermodiffusion in liquids has kinetic and thermodynamic contributions, the success of either approach depends on the system under consideration. While the solvent in colloidal mixtures can be treated as a continuum, such assumptions might lead to large errors in molecular mixtures, where the size of components is comparable. For very weak molecular interactions kinetic theory approaches might hold with some corrections, while the activation energy needed for mass transport becomes a dominating factor when interactions are strong.

For dilute colloidal systems the kinematical model of Brenner [16] gives reasonably good

predictions for the thermal diffusion coefficient with

$$D_T = \lambda D_S \beta, \tag{1.6}$$

where D_S is the self-diffusivity of the solvent, β its thermal expansion coefficient and λ a factor to adjust for non-ideality. Interestingly, D_S and β are exclusively dependent on the solvent and λ is taken to be only weakly dependent on the solute, making D_T in dilute colloid solutions dependent only on solvent properties. The correlations of D_T with D_S [17] and with β [18–20] have been confirmed experimentally.

A number of models that are based only on equilibrium properties of the mixture under consideration are based on the works of Haase, who drew an analogy between mass transport due to pressure and due to a temperature gradient [21]. Kempers provided the groundwork in non-equilibrium thermodynamics and proposed a *Gedankenexperiment* in which two bulbs are held at different temperatures and connected by a valve that allows mass transport. His expression for the thermal diffusion factor α contains a thermodynamic and a kinetic contribution and yields good agreement with experimental data on gases and some organic solvents [22].

Models following the second approach are based on the works of Denbigh [23] and Rutherford and Drickamer [24]. Strongly influenced by the concept of activation energies in chemical processes, they propose that the heat of transfer must be proportional to an activation energy for molecular motion. There are then different approaches to approximate this value with the activation energy of viscous flow [13, 25–27] or with the activation energy of self-diffusion calculated in equilibrium molecular dynamics simulations [28, 29].

Inspired by the additivity found in case of the mass-inertia contribution, Hartmann *et al.* [25, 26] found an additivity for the heats of transport of the pure components for binary mixtures of organic solvents. Good agreement with experimental data was achieved by calculating the Soret coeffient with

$$S_T = \frac{Q_1 - Q_2}{RT^2 \left[1 + (\partial \ln \gamma_1 / \partial \ln x_1)_{p,T}\right]}, \tag{1.7}$$

where R is the gas constant, γ_1 is the activity coefficient and x_1 the concentration of component 1. Q_1 and Q_2 are the molar absolute reduced heats of transport of the respective pure

components, values that have been calculated from experimental data relative to an arbitrary offset ($Q_{\text{tetralin}} = 0$) and give a scale of thermophobicity for the investigated compounds. While this approach works well for non-polar liquids, predictions of S_{T} for mixtures with a polar component are not as successful.

There is a special difficulty in describing with reasonable accuracy associating mixtures, meaning mixtures that show strong interactions such as the hydrogen bonds formed in alcohol-water mixtures. One proposed model [27] uses the activation energy of viscous flow, but takes into account the anomalous concentration dependence observed in the viscosity of associating mixtures. Microscopically, the effect can be understood by the formation of clusters with varying size and alcohol content, that lead to varying activation energies of molecular motion depending on alcohol/water ratio.

1.1.3 Simulations

One of the problems in designing experiments is that thermodiffusion is sensitive to many parameters of the investigated molecules, some of which cannot be altered in reality without altering others as well. Here lies one major advantage of simulations as parameters can be changed systematically and independently. The second advantage is the possibility to get microscopic insight, which can help to characterise and understand molecular interactions. The main simulation methods are equilibrium and non-equilibrium molecular dynamics (EMD and NEMD, respectively) [30, 31].

The additivity of chemical and mass-inertia contribution in molecular mixtures has been corroborated by simulation results and it has been shown that the concentration dependence of S_{T} is due to the chemical contribution [33]. The impact of paramter variation on the mass-inertia contribution has also been studied in detail [34]. Simulations for colloidal systems have been done with NEMD [35] and on a mesoscopic scale [36].

In Ch. 4 of this work NEMD simulations [32] have been used to test the influence of solute-solvent interactions on the temperature dependence of the Soret coefficient and to obtain a microscopic picture of the molecular mixture at different concentrations and temperatures.

Another aim of simulations is the numerical determination of the Soret coefficient, but apart from very non-polar mixtures [37] only qualitative agreement can be reached when simulations of real systems are compared to experimental data as of yet [38, 39]. Here too, the

high sensitivity of thermodiffusion to even small deviations of the simulated from the real system is a problem. However, the fast development in recent years due to better numerical techniques and more accurate molecular force fields makes the simulation approach promising.

1.1.4 Thermodiffusion in aqueous systems

From a phenomenological point of view, the striking difference in the thermodiffusion behaviour of non-polar and polar mixtures is the strong temperature and concentration dependence of the latter ones. Often, a sign change of the Soret coefficient, *i.e.* a change from thermophobic to thermophilic behaviour, can be observed [40, 41].

Since aqueous systems are of great importance for biotechnological applications, many such systems have been investigated experimentally [41–49]. For a large number of aqueous solutions a common temperature dependence of the Soret coefficient S_T is observed [41]: at low temperatures, there is a rapid increase of S_T that flattens out at high temperatures. Often, the solute is thermophilic ($S_T < 0$) at low temperatures and there is a sign change to a positive S_T (thermophobic behaviour) at a certain temperature T^*, which is specific to the mixture.

A qualitative explanation of this behaviour has been attempted by Wang *et al.* [42]: At low temperatures, where the enthalpy contribution dominates, a minimisation of free energy is realised by a largely undisturbed hydrogen-bond network of water, leading to a preference for the solute to be on the warm side. At higher temperatures, the entropy contribution dominates and the high translational and orientational entropy of water leads to enrichment on the warm side.

Iacopini and Piazza [50] have proposed the empirical equation

$$S_T(T) = S_T^\infty \left[1 - exp\left(\frac{T^* - T}{T_0} \right) \right], \tag{1.8}$$

which is able to characterise experimental results that follow the typical behaviour with three fitting parameters: S_T^∞ is the Soret coefficient that is approached at high temperatures, T^* the temperature at which the sign change occurs, and T_0 characterises the slope of the curve and therefore the temperature sensitivity of S_T. It was suggested that the last two

parameters are not independent from each other and $T^*/T_0 = \mathrm{const.}$, at least for chemically similar systems [51].

While this work focuses on uncharged systems it should be noted that thermodiffusion is strongly influenced by charge. In the case of charged colloids in water the Soret coefficient becomes a function of the Debye screening length [52, 53] and can also be expressed in dependence of the surface potential [54]. Ion-specific effects consistent with the Hofmeister series have been observed when electrolytes are added even in low concentrations to nano particles [55]. This demonstrates the sensitivity of thermophoresis to changes in the particle-solvent interactions, but has to be investigated in more detail.

There is no clear microscopic picture based on these observations, but a rise of S_T is generally associated with a weakening or breaking of hydrogen bonds [56]. Attempts to correlate the number of hydrogen-bond sites with S_T have yielded a linear dependency for solute molecules of a homologous series [57]. These observations will be discussed further in Ch. 5-7.

There are, however, some mixtures that do not follow the typical behaviour described above. This is mainly the case for polar molecular mixtures at high concentrations, but can also be observed for some compounds, such as ethanol/water [58], at low concentrations. This behaviour has been connected with microstructural heterogeneities and is discussed in detail in Ch. 4 of this work.

1.2 Applications

The earliest applications for thermodiffusion are methods that use its sensitivity to molecular properties such as size, mass, charge and interactions with the solute for separation. In 1939, Clusius and Dickel reported isotope separation of chlorine gas in a thermogravitational column [6]. Later, this technique was used to enrich uranium in the Manhattan project [59]. Due to high energy costs and the technical difficulties in maintaining the almost 15 m high columns, other methods for isotope separation have replaced thermogravitational columns in technical applications.

A method to fractionate and characterize polymers, the Thermal Field Flow Fractionation (th-FFF), was reported to fractionate polystyrene in 1967 [60] and has been successfully applied on colloids and synthetic polymers since [61].

In nature, vertical and horizontal temperature gradients in hydrocarbon reservoirs lead to a separation of components and modelling of the spacial composition of these deposits demands that thermodiffusion is considered when planning the extraction [62]. It has also been shown that thermodiffusion has to be considered in combustion processes [63].

Two fields of application that are important for aqueous systems are presented in more detail below.

1.2.1 Thermogravitational columns

Thermogravitational columns were the first application for thermodiffusion, as well as one of the first methods used to measure the effect. Given that the thermophobic component also has the higher density, they enhance the de-mixing effect of thermodiffusion by combining it with a convection flow (see Fig. 1.2). This is realised by applying a temperature gradient horizontally across the column, leading to a convectional stream that flows down on the cold side and up on the warm side of the column. The thermophobic component of the mixture is driven towards the cold side, and carried down by convection, the thermophilic component is carried up on the warm side. Depending on flow speed and geometry of the column, a steady state is reached after some time with a concentration gradient in vertical direction along the column that is proportional to the Soret coefficient.

Thermodiffusion has also been proposed as a mechanism that might have played a role in the origin of life [64]. One central issue in this context is the so-called concentration problem, posing that any organic matter in the primordial ocean would have been present only in very small concentrations, making hydrolysis the dominating reaction pathway and effective polymerisation virtually impossible. This same problem is expressed from another viewpoint when thermodynamics are considered: the formation of structure that is necessary for even the simplest building blocks of life is connected with a reduction of entropy and will therefore only occur spontaneously under non-equilibrium conditions. The solution to this problem is an accumulation process of some kind, driven by a non-equilibrium setting. In a mechanism similar to thermal columns, mineral pores near hydrothermal vents could have trapped and enriched prebiotic molecules, making polymerisation reactions possible [65]. The key difference is that these pores are not closed, like the thermal columns, but are open
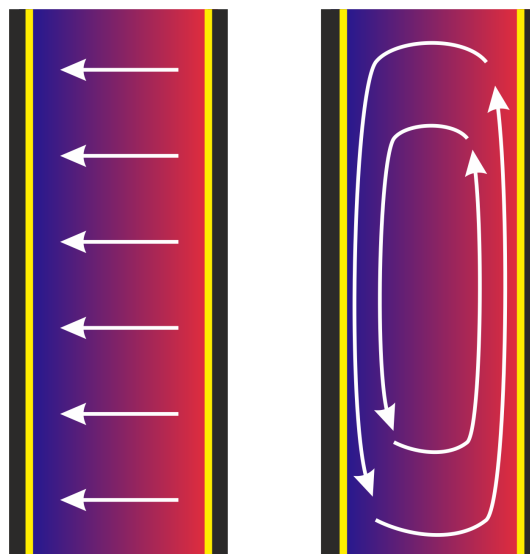
Figure 1.2: *Schematic of a thermal column. Thermodiffusion of the thermophobic component towards the cold side (left) is combined with circular convective flow (right), resulting in a concentration gradient vertically along the column.*

on the top to a reservoir with a fixed concentration of the thermophobic component, so that this component can diffuse into the pore and be accumulated at the bottom.

Similar geometries are also promising for microfluidic applications. One example is the optimisation of polymerase chain reaction (PCR) in a capillary [66]. A slightly more complex convectional pattern allows the accumulation of replicated DNA fragments in the capillary, while the reaction is continually fed and the necessary temperature cycling provided by the convectional stream that circulates the molecules through hot and cold regions.

1.2.2 Microscale thermophoresis

A relatively new application for thermophoresis in the field of life sciences, but already one of the most important ones, is the so-called Microscale thermophoresis (MST) [44–47]. In this experiment a laser is pointed into a solution with fluorescently labelled biomolecules, *e.g.* a protein. The solution heats up in the lit area, thermodiffusion sets in, and a fluorescence intensity I_{fluor} is measured that is proportional to S_{T} of the protein. When a ligand is titrated into the solution, the thermodiffusion behaviour of the protein changes and with it the fluorescence intensity. This can be observed in a titration curve and equilibration constants
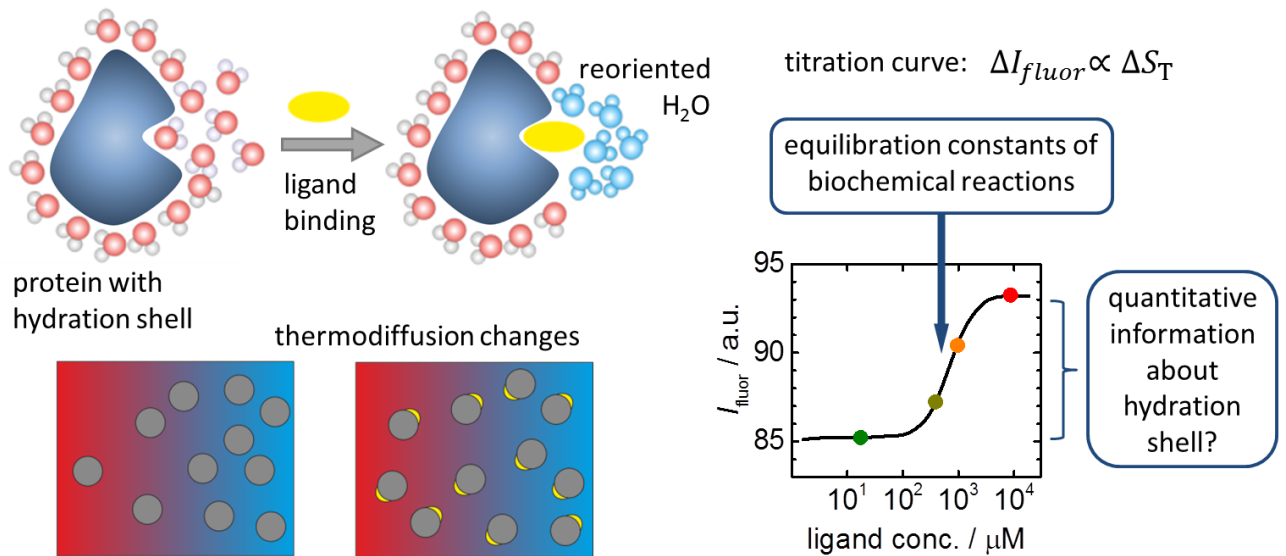
Figure 1.3: *Schematic principle of Microscale thermophoresis (MST): ligand binding causes changes in the hydration shell of the fluorescently labelled protein. The resulting change in thermodiffusion can be observed as a titration curve. The concentration dependence of the change is used to calculate the association constant K_a, the amplitude, which is usually not examined further, should allow inferences about changes in hydration.*

of the binding reaction can be calculated from the dependence on ligand concentration (see Fig. 1.3).

The strong change of S_T upon ligand binding is due to its sensitivity towards changes in the hydration shell, since ligand molecules are usually very small compared to proteins, so that the mass change can be neglected. What exactly occurs in the hydration shell when a ligand binds is not known and might be specific to the protein. It should, however, be possible to extract information about the changes occurring in the hydration shell from the amplitude of the S_T-change. This inspired the investigation of complexes presented in this work.

1.3 Experimental Details

1.3.1 Description of the IR-TDFRS setup

The thermodiffusion data presented in this work was measured by infra-red Thermal Diffusion Forced Rayleigh Scattering (IR-TDFRS) [67, 68]. This technique is especially useful to measure thermodiffusion behaviour in binary aqueous solutions [69]. An infra-red laser beam ($\lambda = 980$ nm) is split and the two beams are crossed at a small angle inside the sample, so that the interference leads to a periodic intensity grating in the solution. The infra-red light is absorbed by the water [70] and results in a periodic heat grating approximately 100 μs after the laser is switched on. The mass transport due to thermodiffusion is well separated on the time scale, setting in after well over 1 ms, even for very small solutes. The thermo-diffusion of the solute leads to a concentration grating overlaying the temperature grating. Both, the temperature and the concentration differences in the grating, result in a refractive index contrast. A read-out beam ($\lambda = 633$ nm) that crosses the grating is scattered and the heterodyne intensity of the refracted beam is measured against time. This intensity is proportional to the refractive index contrast of the grating and described by

$$\xi_{het}(t) = 1 - exp\left(-\frac{t}{\tau_{th}}\right) - \frac{A}{\tau - \tau_{th}}\left\{\tau\left[1 - exp\left(-\frac{t}{\tau}\right)\right] - \tau_{th}\left[1 - exp\left(-\frac{t}{\tau_{th}}\right)\right]\right\}, \qquad (1.9)$$

where $\tau_{th} = 1/(q^2 D_{th})$ and $\tau = 1/(q^2 D)$ are the heat and mass diffusion times, respectively, with the thermal diffusion coefficient D_{th}, the mass diffusion coefficient D, and the grating vector q.

Equation 1.9 is fitted to the measured intensity curve (see Fig. 1.4) and the Soret coefficient S_T can be calculated from the amplitude of the concentration signal A with

$$S_T = \frac{A}{c(1-c)}\frac{(\partial n/\partial T)_{p,c}}{(\partial n/\partial c)_{p,T}}, \qquad (1.10)$$

given that the contrast factors $(\partial n/\partial T)_{p,c}$ and $(\partial n/\partial c)_{p,T}$, that is the change of refractive index with temperature and concentration, are known.

Figure 1.5 shows the IR- TDFRS setup in some detail. The writing and the read-out lasers
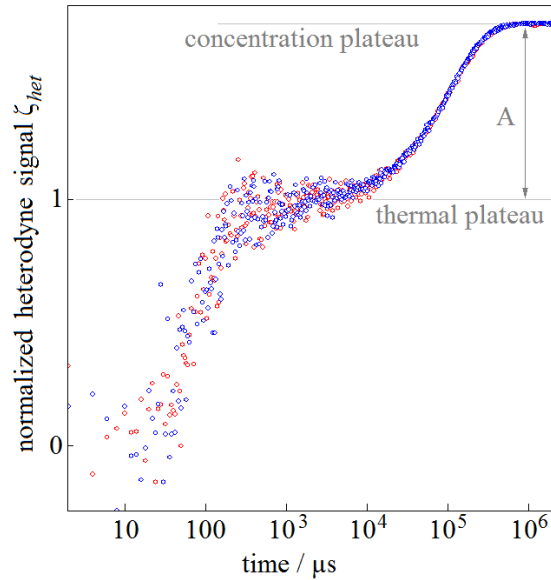
Figure 1.4: *Normalized intensity signal averaged over several measurements. The Soret coefficient S_T can be calculated from the amplitude of the concentration signal A.*
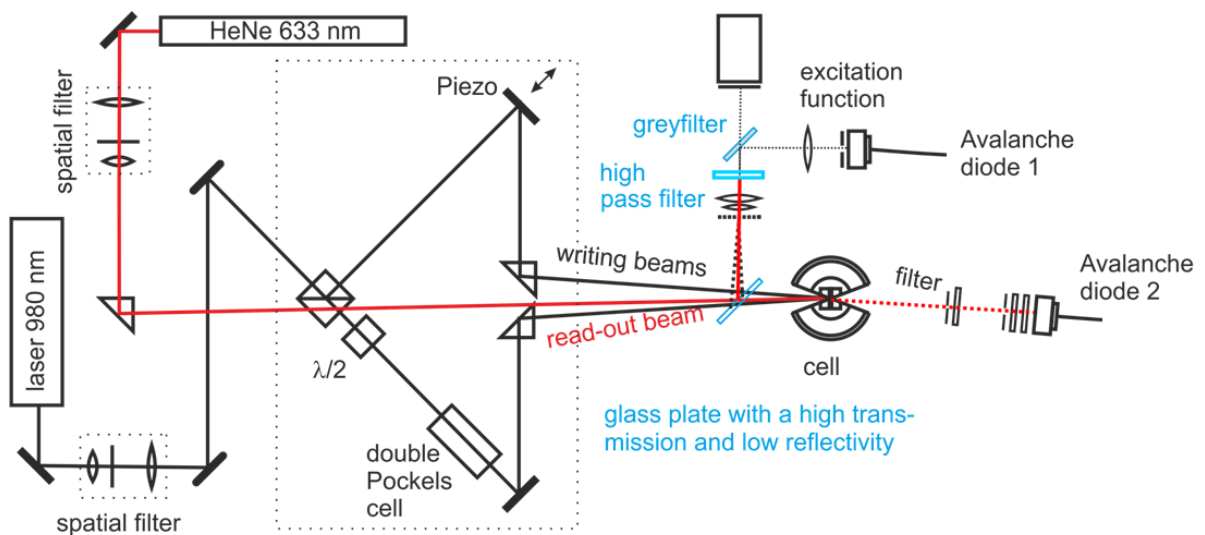

Figure 1.5: *Sketch of the IR-TDFRS setup. The writing beams (980 nm) are black, the read-out beam (633 nm) is red.*

run continuously, phase shift of the grating by 180° is achieved with the Pockels cell. A small fraction of the writing beam is reflected before the sample into a CCD camera and an avalanche diode to measure the grating period and monitor the excitation function, respectively. The refracted read-out beam passes through an IR-filter and its intensity is detected by the second avalanche diode and recorded over time. An example for the normalized signal is shown in Fig. 1.4.

1.3.2 Contrast factors

The contrast factors $(\partial n/\partial T)_{p,c}$ and $(\partial n/\partial c)_{p,T}$ are necessary to quantify the concentration gradient that sets in through thermodiffusion. They are measured separately for each sample investigated by IR-TDFRS. The refractive index change with temperature at constant pressure and concentration $(\partial n/\partial T)_{p,c}$ is measured interferometrically [71] in a temperature range of 5K around several temperatures, always including the highest and lowest temperature investigated with TDFRS. The measurement was done twice to make sure that no change occured in the sample. The data was fitted with a 2nd-order polynomial to average the data sets and interpolate for temperatures not measured directly. The refractive index change with concentration at constant pressure and temperature $(\partial n/\partial c)_{p,T}$ was measured on dilution series with an Abbe refractometer (Anton Paar ABBEMAT RXA 158) with an accuracy of 0.00002 nD and a temperature control of ΔT = ±0.03 K. The refractometer uses a wavelength of 589.3 nm (sodium line), shorter than the wavelength of the read-out beam in our IR-TDFRS setup (HeNe-laser, 632.8 nm). This causes a small systematic error in the refractive index increment in the order of 0.5-1% [72,73]. The refractive index n was plotted against concentration and the slope $(\partial n/\partial c)_{p,T}$ determined with a linear fit. This was done for 4-5 temperatures, in-between values were interpolated.

1.3.3 Sample preparation

Formamide/water: Solutions were prepared using formamide ($\geq$ 99.5%, Sigma-Alrich, 89555 Steinheim, Germany) and water (Millipore). 1-10 mL of the sample were prepared each time, components were measured by volume (error < 0.5%). All solutions were filtered (0.2 μm). Possible changes in the solution was monitored by refractive index measurement.

There was no change when the solutions were stored in the fridge, but after long measurements (> 3 d) at higher temperatures a small rise of n could sometimes occur. This might be due to evaporation of water or a hydrolysis of the formamide. In both cases, the result is a slight uncertainty concerning the formamide concentration during TDFRS-measurements. However, repeated measurements of the same samples over time show no systematic change of the measured S_T, which leads to the conclusion that the change is too small to impact TDFRS-results.

Urea/water: Urea ($\geq 99.0\%$, Fluka, Sigma-Aldrich, 89555 Steinheim, Germany) and water (Millipore) were used. 5-10 g of solution were prepared each time by weighting of the components. Accuracy of the scales was $\pm$ 0.2 mg. All solutions were filtered (0.2 μm) to remove dust. Again, a slight rise of the refractive index could be observed for long measurements.

Cyclodextrins: Samples were prepared using α-cyclodextrin ($> 98.0\%$, Tokyo chemical industry), β-cyclodextrin (99.0%, Tokyo chemical industry), γ-cyclodextrin ($> 98.0\%$, Tokyo chemical industry), methyl-β-cyclodextrin (with 55% of hydroxyl groups randomly methylated, no specification was given for the purity by Tokyo chemical industry), water (Millipore), and formamide ($\geq 99.5\%$, Sigma-Alrich, 89555 Steinheim, Germany). All samples were prepared to have a cyclodextrin mass fraction of $(1.0 \pm 0.01)\%$. Approximately 2 mL solution were prepared at a time, stirred for an hour at room temperature and filtered (0.2 μm).

Cyclodextrins-ASA complexes: Additionally to the cyclodextrins described above, heptakis(2,6-di-O-methyl)-β-cyclodextrin (m-β-def, $>98.0\%$, Sigma-Aldrich) and acetylsalicylic acid (ASA, $\geq$ 99.0%, Sigma-Aldrich) were used. As before, samples had a cyclodextrin concentration of (1.0 ± 0.01) wt%, ASA was added in equimolar amounts. For some samples m-β-def and unmethylated β-cyclodextrin were mixed to achieve intermediate degrees of methylation. Samples were stirred for an hour at RT and filtered (0.2 μm). A systematic change of the S_T-measurements due to aspirin hydrolysis could not be observed.

Streptavidin-biotin complexes: Streptavidin Streptomyces Avidinii Recombinant produced in *E. coli* (Prospec, 7670308 Rehovol, Israel) with the amino acid sequence MAE AGITGTWYNQLGSTFIVTAGADGALTGTYESAVGNAESRYVLTGRYDSAPATDGSG TALGWTVAWKNNYRNAHSATTWSGQYVGGAEARINTQWLLTSGTTEANAWKSTL VGHDTFTKVKPSAAS and a molecular weight of 52 kDa (tetramer) was used. It was cleaned with PD-10 comlumns, lyophilized and then kept at -20°C. The biotin was pur-

chased as lyophilized powder with ≥99% purity (Sigma-Aldrich, 89555 Steinheim, Germany). The buffer stock solution has the following composition: 250 mM TrisHCl (Tris[hydroxymethyl]aminomethane, ≥99.9%, Sigma-Aldrich, 89555 Steinheim, Germany; Hydrochlorid acid 37%, Merck, 64271 Darmstadt, Germany), 1.2 M NaCl (≥99.5%, Merck, 64271 Darmstadt, Germany), 50 mM KCl (≥99.5%, Merck, 64271 Darmstadt, Germany), 30 mM $MgCl_2$ (≥99.0%, Merck, 64271 Darmstadt, Germany).

For IR-TDFRS experiments the buffer stock solution was diluted with Milipore water 1:9 by volume and filtered (0.2 μm) to remove dust. The concentrations of SA in buffer solution was 50 mg/mL (SA weight fraction 0.048 ± 0.001). For the samples with biotin, the ligand was dissolved beforehand in a larger amount of buffer, so that it could be weighted with reasonable error, then this solution was added to the SA. The samples were vortexed 5-10 min, then centrifuged at 60 rpm for 20 min. The solution was carefully extracted with a syringe and put into the Quartz cells used for IR-TDFRS. The samples were left standing at RT for ~12 h, because a good contrast for measuring could usually not be found before.

1.4 Outline of the thesis

This thesis focuses on thermophoresis in a biological context, more precisely on the question how changes in the hydration shell of the solute influence the temperature dependence of thermophoresis in aqueous systems. The work consists of two main parts.

The first part investigates the concentration and temperature dependence of thermodiffusion of formamide and urea in water (chapters 2-4). At high concentrations both systems show a deviation from the usual temperature dependence of the Soret coefficient S_T in aqueous solutions. Due to the fact that formamide is considered as a likely educt in the formation of prebiotic molecules on the early earth, the experimental data was used to simulate the accumulation of formamide in a hydrothermal pore. The results show that accumulations of up to 85 wt% formamide are possible, a concentration which would allow the formation of more complex organic molecules. In the following chapter these simulations are revisited to gain a better understanding of the accumulation mechanism in the pore and how it is influenced by pore geometry. Finally, the unusual temperature dependence at high concentrations is addressed for the urea/water system in chapter 3. Non-equilibrium molecular dynamics simulations are used to arrive at a microscopic understanding of changes in the

system caused by a rising urea concentration. The results help to rationalise the correlation between the temperature dependence of S_T and hydrophilicity of the solute (ΔS_T vs. $\log P$), an observation that is explained more closely in chapter 6.

The second part of the thesis (chapters 5-7) is concerned with the question how complex formation leads to changes in thermodiffusion of biomolecules and if inferences can be drawn from these changes about the hydration shell. Chapter 5 introduces cyclodextrins, cyclic oligosaccharides that are interesting for drug delivery, because they form complexes with several small drug molecules. First, the thermodiffusion of different cyclodextrins is investigated in two polar solvents, water and formamide. The following chapter focuses on the complex formation with aspirin, showing a strong connection between thermodiffusion behaviour and hydrophilicity. In Ch. 7 preliminary results are shown investigating the streptavidin-biotin complex as a model for protein-ligand complexes. While the thermodiffusion behaviour and especially its temperature dependence are very sensitive to changes in the hydration shell and allow some qualitative statements, complimentary methods and further investigation are necessary to quantify the results.

The discussion gives an overview of the work, highlighting the insights into accumulation in a hydrothermal pore, influence of hydrophilicity of a solute on its thermodiffusion, and changes of thermodiffusion upon complex formation. The observed correlation between hydrophilicity and the temperature dependence of thermodiffusion is discussed in more detail. Additional experimental results on the thermodiffusion of streptavidin-biotin are presented and discussed.

2 Accumulation of Formamide in Hydrothermal Pores to Form Prebiotic Nucleobases

Accumulation of formamide in hydrothermal pores to form prebiotic nucleobases

Doreen Niether[a], Dzmitry Afanasenkau[a], Jan K. G. Dhont[a,b], and Simone Wiegand[a,c,1]

[a]Institute of Complex Systems-3 Soft Condensed Matter, Forschungszentrum Juelich GmbH, D-52428 Juelich, Germany; [b]Department of Physics, Heinrich-Heine-Universitaet Duesseldorf, D-40225 Dusseldorf, Germany; and [c]Department für Chemie - Physikalische Chemie, Universitaet zu Koeln, 50939 Cologne, Germany

Formamide is one of the important compounds from which prebiotic molecules can be synthesized, provided that its concentration is sufficiently high. For nucleotides and short DNA strands, it has been shown that a high degree of accumulation in hydrothermal pores occurs, so that temperature gradients might play a role in the origin of life [Baaske P, et al. (2007) *Proc Natl Acad Sci USA* 104(22): 9346−9351]. We show that the same combination of thermophoresis and convection in hydrothermal pores leads to accumulation of formamide up to concentrations where nucleobases are formed. The thermophoretic properties of aqueous formamide solutions are studied by means of Infrared Thermal Diffusion Forced Rayleigh Scattering. These data are used in numerical finite element calculations in hydrothermal pores for various initial concentrations, ambient temperatures, and pore sizes. The high degree of formamide accumulation is due to an unusual temperature and concentration dependence of the thermophoretic behavior of formamide. The accumulation fold in part of the pores increases strongly with increasing aspect ratio of the pores, and saturates to highly concentrated aqueous formamide solutions of ∼85 wt% at large aspect ratios. Time-dependent studies show that these high concentrations are reached after 45–90 d, starting with an initial formamide weight fraction of 10^{-3} wt % that is typical for concentrations in shallow lakes on early Earth.

concentration problem | hydrothermal vents | molecular evolution | origin-of-life problem | thermophoresis

Thermophoresis has been suggested as an active transport mechanism to reach high concentrations of prebiotic molecules to culminate in the formation of RNA (1). A still open question is whether thermophoresis can also be a possible mechanism to form prebiotic nucleobases from simple molecules such as hydrogen cyanide (HCN) and formamide (FA). Already for almost 50 years, FA has been discussed as an important compound from which prebiotic molecules originate (2–7). It has been shown that all known nucleobases can be synthesized from aqueous FA solutions (4). In diluted HCN solutions, polymerization of HCN to form nucleobases becomes favored over hydrolysis of HCN at concentrations of 0.03–0.3 wt % (8). To our knowledge, there are no similar studies of diluted FA solutions. Taking into account the faster hydrolysis of FA (3), we estimated that a 100-times-higher concentration between 3 wt % and 33 wt % should be sufficient for the synthesis of prebiotic molecules in aqueous solutions. In the ocean during the early stages of Earth, the natural occurring concentrations at a low temperature (10 °C) and a pH between 6 and 8 are estimated to be only on the order of 10^{-9} wt %, whereas, in shallow lakes (depth 10 m), due to vaporization and FA input from the atmosphere, higher concentrations of about 10^{-3} wt % are possible (3). Still, these natural concentrations are far too small compared with those required for the formation of nucleobases.

In this work, we perform numerical calculations for the spatial and time dependence of the concentration of aqueous FA solutions in hydrothermal pores exposed to a temperature gradient to investigate whether it is possible to reach sufficiently high FA concentrations that are necessary to initiate the synthesis of prebiotic nucleobases. The dependence of the highest FA concentration in part of the pore is analyzed as a function of the initial FA concentration, which is the reservoir concentration within the shallow lake, at various ambient temperatures, initial concentrations, and aspect ratios of the pores. The highest FA concentration within the pore relative to the initial FA concentration defines the so-called accumulation fold. The concentration dependence and temperature dependence of the thermodiffusion and mass diffusion coefficients of FA in aqueous solutions as determined by means of Infrared Thermal Diffusion Forced Rayleigh Scattering (IR-TDFRS), as well as other revelant physical properties of FA solutions, are used as an input to these calculations. In contrast to the previous study (1) for nucleotides and short DNA fragments, we do not find an exponential increase of the accumulation fold with increasing pore aspect ratio. Instead, the accumulation fold increases exponentially only at relatively small aspect ratios, sharply increases at intermediate aspect ratios, and, finally, saturates to highly concentrated FA solutions on the order of 85 wt % at relatively large aspect ratios. The sharp increase of the accumulation fold with increasing pore size is found to be essentially independent of the initial, shallow lake concentration.

Thermophoresis, also known as the Ludwig−Soret effect or thermodiffusion, is the migration of particles or molecules induced by a temperature gradient (9). In a binary fluid mixture, this mass transport is described by a contribution of the form $\sim -D_T \vec{\nabla} T$ to the mass flux $\vec{j}$, where D_T is the thermodiffusion coefficient. When $D_T > 0$, mass transport occurs from high to low temperature. The total mass flux is thus given by

$$\vec{j} = -D\vec{\nabla}c - c(1-c)D_T\vec{\nabla}T, \qquad \textbf{[1]}$$

where the first term describes mass transport due to gradients in the concentration c, with D the Fickian or mass diffusion coefficient. For a time-independent temperature gradient, a steady state is reached when the mass fluxes due the Fickian diffusion and thermodiffusion contributions cancel each other. The ratio of the resulting concentration gradient and the applied temperature gradient is characterized by the Soret coefficient $S_T = D_T/D$. A larger Soret coefficient implies a larger concentration gradient for a given temperature gradient.

Several theoretical approaches exist to describe thermodiffusion (or thermophoresis) of liquid mixtures, polymer solutions, and colloidal suspensions (10–13). The review by Würger (11) introduces theoretical concepts for colloids, and the book chapter by Wiegand (13) gives an overview on the basic physics of the effect. An excellent agreement between experimental results and theoretical models has been found for charged spherical and rod-like colloids (14, 15), whereas the interfacial effects as they occur in microemulsions are still not fully understood (16, 17). So far, no microscopic particle-based theory exists to describe thermophoresis on a microscopic level for liquid mixtures, such as aqueous FA solutions. Simulations have been performed to investigate the effects of attractive and repulsive interactions between uncharged and charged colloidal particles (18, 19), to study the influence of chain length and stiffness of polymers (20), or to study specific interactions as they occur in aqueous mixtures (21). Due to their importance in biotechnology, many aqueous systems have been studied experimentally. Although the charge contributions to the thermophoretic movement of the solute molecules are well understood, the contributions of the hydration layer, although of high importance (for example in protein−ligand interactions), are not yet understood. It is known that the Soret coefficient S_T of the solute molecules increases when hydrogen bonds break. There are two mechanisms that can lead to a breaking of hydrogen bonds between solute and water molecules. One possibility is to add an ingredient with a strong affinity to water, so that the bonds open (22). Alternatively, an increase of the temperature disrupts the hydrogen bonds between water and the solute. This leads, for aqueous solutions of biological and synthetic molecules, to a similar temperature dependence of S_T (23), which can be described by an empirical equation proposed by Iacopini et al. (24).

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

with fitting parameters S_T^∞, T^*, and T_0. Recently, it has been shown that the number of hydrogen bond sites of solute molecules plays a key role for describing the temperature dependence of S_T and the thermodiffusion coefficient D_T. It turns out that S_T depends linearly on the difference of donor and acceptor sites of the solute molecule belonging to a homologous series (25). Hydrogen bonding certainly plays an important role also in aqueous FA solutions.

Experimental Results

The Soret coefficient of FA/water mixtures was measured by means of IR-TDFRS in the temperature range from 10 °C to 70 °C and in the FA weight fraction range from $\omega = 0.02$ to $\omega = 0.9$. Such measurements require the refractive index of FA solutions as a function of temperature and concentration (*SI Appendix, Refractive Index Contrast Measurements*). Fig. 1 shows the measured Soret coefficients as a function of temperature for various concentrations. S_T is always positive, which indicates that FA is thermophobic and enriches in the cold regions. Specific to the FA/water system
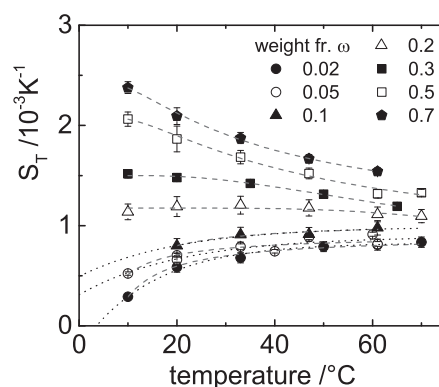

Fig. 1. The Soret coefficient as a function of temperature for various FA concentrations (ω is the weight fraction of FA). The dotted lines for the three low concentrations are fits according to Eq. **2**, and the dashed lines are fits to Eq. **3**.

is the sign change of the slope of the temperature dependence from positive to negative on increasing the FA concentration. At low concentrations ($\omega < 0.2$), the temperature dependence can be described by Eq. **2** (see the dotted lines in Fig. 1), whereas Eq. **2** is not applicable at higher concentrations, where the Soret coefficient increases with increasing FA concentration. Qualitatively, the often-found temperature dependence of the Soret coefficient as described by Eq. **2** might be explained as follows. At low overall temperatures, the system tries to minimize its local free energy, $F = U - T \cdot S$, by forming hydrogen bonds, thus minimizing the internal energy U with a relatively small entropic contribution, so that the water molecules accumulate at the cold side. At higher temperatures, where the entropic contribution is dominant, the system minimizes its free energy by maximizing the orientational and translational entropy S, which leads to an enrichment of water molecules on the warm side (26). At higher FA concentrations, Eq. **2** can no longer be used, and it turns out that the temperature dependence can be described empirically by a simple exponential decay,

$$S_T(T) = S_T^\infty + S_T^0 \cdot \exp(-T/T_0), \qquad \textbf{[3]}$$

which corresponds to the dashed curves at high concentrations in Fig. 1. Deviations from Eq. **2** have also been observed for other systems. The most prominent example is the system ethanol/water (27), but, also, dimethyl sulfoxide (DMSO)/water (28) does not follow the general trend at low concentrations. In contrast to FA/water, for high ethanol and DMSO concentrations, an increase of S_T with increasing temperature and a decrease for low concentrations is oberved. Also, Maeda et al. (25) observed a decrease for various types of crown ethers of S_T with increasing concentration.

As a consequence of the temperature and concentration dependence shown in Fig. 1, FA accumulates in colder regions for all concentrations, whereas, with increasing concentration, the driving force for accumulation in these regions increases. One could expect that this leads to a self-amplifying mechanism leading to significant FA accumulation. The FA molecule $HCONH_2$ is a weak base, as it can bind protons to its negatively charged oxygen and the amino group. The $pK_a \approx 20$ value of FA is very large, so that the fraction of molecules carrying a positive elementary charge is typically as small as 10^{-13}. There is thus no measurable charge contribution to the thermodiffusive motion of FA. The molecular dynamic simulations in ref. 29 offer the possibility of understanding the temperature and concentration dependence of aqueous FA solutions in more detail. In Fig. 2, we replot the average number of water−water (W−W), FA−water (FA−W),

21

and FA−FA hydrogen bonds as a function of the FA weight fraction, ω. As expected, the number of hydrogen bonds between water molecules decreases with increasing FA concentration. Accordingly, the number of FA−FA bonds increases. Note that, for pure water, the number of W−W bonds is slightly larger than the number of FA−FA bonds for pure FA, but both solvents show a strong tendency to form hydrogen bonds. With increasing FA concentration, the number of FA−W hydrogen bonds decreases. Around a weight fraction of 0.13, the number of FA−W bonds is equal to the total number of H bonds that FA forms. This is precisely the concentration range where the slope of the temperature dependence of S_T in Fig. 1 changes from positive to negative. This indicates that the temperature dependence as given in Eq. 2 is valid as long as FA molecules are mostly surrounded by water molecules. As soon as FA−FA hydrogen bonds between FA molecules are formed, the temperature dependence of S_T changes from increasing to decreasing with increasing temperature. A possible explanation is that FA at higher concentrations migrates in temperature gradients as entire FA clusters. With increasing concentration, larger and heavier clusters are formed, which have a larger Soret coefficient, whereas, for increasing temperature, the clusters become smaller due to thermal motion, which leads to a decrease of the Soret coefficient.

As mentioned above, the systems ethanol/water and DMSO/water also do not follow Eq. 2. Compared with FA, both ethanol and DMSO have a much lower hydrogen bond capability, which is only roughly two bonds per molecules, whereas the water hydrogen bond capability lies between 3.5 and 4 per water molecule (30, 31). Both aqueous mixtures show microheterogeneous structures at low concentration (31, 32), which are not formed in the case of FA/water due to their almost equal ability to form hydrogen bonds. As in the case of the crown ethers, the solute molecules are not well interlinked with water molecules. According to computer simulations (33) and near-infrared spectroscopy (34), the water molecules form a clathrate-like structure around the crown ether. Only two water molecules are doubly hydrogen-bonded (bridging) to the crown ether oxygen atoms. In contrast, proteins are often linked to water by 50–100 hydrogen bonds (35), so that the hydrogen bonds of the solute molecules at low concentrations are not influenced by other solute molecules. In conclusion, we can state that Eq. 2 only holds if the solute molecules are well connected to water by hydrogen bonds and no microheterogeneous structures or cages are formed.
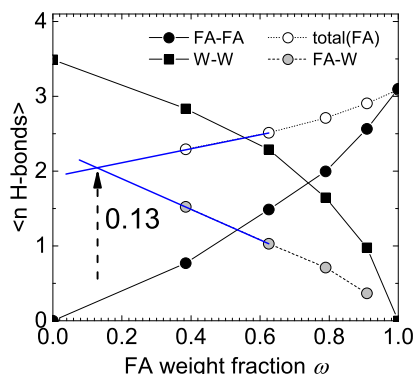
Fig. 3. Soret coefficient as a function of the FA weight fraction, ω, for various temperatures. The solid lines correspond to a fit according to Eq. 4.

Fig. 3 shows the Soret coefficient of FA as a function of the FA weight fraction ω. The concentration dependence can be described by an empirical equation, which has the form of the so-called Hill equation (36, 37),

$$S_T(\omega) = \frac{\omega^a}{K + \omega^a} + S_T^0,$$ [4]

where a, K, and S_T^0 are fitting parameters (the dashed lines in Fig. 3). In contrast to many other systems, S_T for FA increases with concentration. The change in the slope of curves at different temperatures leads to a common intersection at a weight fraction of 0.13. Such an intersection point is often found for associated mixtures (28).

Accumulation in Hydrothermal Pores

The stronger accumulation of FA for larger concentrations and lower temperatures raises the question whether it is possible to accumulate sufficient FA by thermophoresis and convection in hydrothermal pores, such that chemical reactions can be initiated to form nucleobases as prebiotic molecules from FA. Using commercial finite element software (COMSOL Multiphysics Modeling Software), we solved the coupled Navier−Stokes, diffusion, and heat transfer equations in two dimensions, and determined the accumulation of FA in similar hydrothermal pores as in ref. 1. The diffusion equation includes both convection and thermophoresis. Numerical calculations use as an input the experimentally determined concentration and temperature dependence of the thermal diffusion and mass diffusion coefficients as obtained from the IR-TDFRS measurements, as well as the viscosity, the specific mass density, the heat conductivity, and the heat capacity of FA/water mixtures (*SI Appendix, Temperature and Concentration Dependence*). Details on the mesh sizes that were used in the calculations can be found in *SI Appendix, Numerical Calculations*.

Fig. 4 shows a contour plot of the concentration profile in a pore with aspect ratio 10 in the stationary state. At the top of the pore, the FA concentration is constant, reminiscent of the concentration in a shallow lake. This is also the initial concentration within the pore, before the temperature gradient is switched on. The right side of the pore is warmer compared with the left side, with a temperature difference of 30 K for all calculations. The maximum concentration in the stationary state within the dark red-colored region at the bottom corner of the pore defines the accumulation fold that is of interest here. This is the region where possible formation of nucleobases from FA will take place.

Fig. 5A shows the accumulation fold as a function of the height to width aspect ratio r. For comparison with literature results, we first performed calculations for an aqueous nucleotide solution

Fig. 2. Average number of H bonds in an FA−W mixture as a function of FA weight fraction ω taken from ref. 29. The lines connect the points. The black symbols mark the average number of H bonds between W−W (black squares) and FA−FA (black circles). The total number of FA hydrogen bonds (white circles) is the sum of FA−FA and FA−W bonds. It becomes equal to the average number of H bonds between FA and W (gray circles) in dilute solution around $\omega = 0.13$.
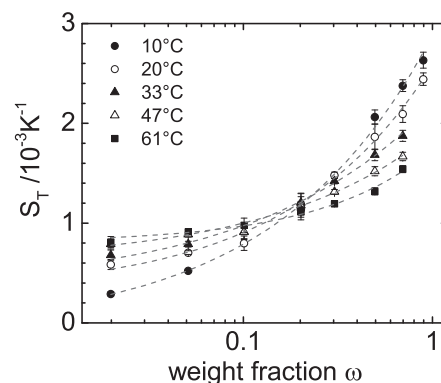
Fig. 4. Contour plot of the concentration profile in a pore with aspect ratio 10 connected to a reservoir in the stationary state. The vertical and horizontal arrows mark the convective and thermodiffusive flow, respectively.

(see Fig. 5, the green solid dots), with an initial concentration of $\omega_0 = 10^{-5}$. In this calculation for the nucleotide, literature values for $S_T = 0.015$ K^{-1}, as well as for $D = 400$ μm^2/s were used, whereas the physicochemical properties of pure water were used for the solvent properties, as done in ref. 1. For low aspect ratios, we find good agreement with the exponential function (see Fig. 5, solid line), as proposed in ref. 1. to describe the aspect ratio dependence of the accumulation fold. At high accumulation folds, however, we find the expected deviations from an exponential accumulation as the nucleotide concentration approached 100% saturation. The concentration and temperature dependence of the Soret coefficient and the mass diffusion coefficient of nucleotide solutions for somewhat elevated concentrations is not known. Additionally the solubility of nucleotides is limited, therefore the last data point in Fig. 5A, corresponding to a concentration of 35 wt %, is marked as an open circle to indicate the uncertainty of this data point.

In contrast to nucleotides, FA and water are miscible at any fraction, so that the entire concentration range is accessible. We studied three different mean temperatures, $T_{mean} = 25$ °C, 45 °C, and 75 °C, for the optimal pore widths of 180 μm, 160 μm, and 100 μm, respectively, to achieve an efficient accumulation (*SI Appendix, Numerical Calculations*). The initial concentration was varied between $\omega_0 = 10^{-9}$ and $\omega_0 = 10^{-5}$, corresponding to FA concentrations as predicted for oceans and shallow lakes under early Earth conditions (3). After an initial exponential increase of the accumulation fold with the pore aspect ratio, we observe, for all studies, a steep increase followed by a plateau when the accumulation fold reaches $1/\omega_0$ of a pure FA solution, as can be seen from Fig. 5A. This saturation plateau is approached at lower aspect ratios for larger temperatures, thus favoring an accumulation in wider pores at lower ambient temperatures.

Fig. 5B shows the time dependence of the accumulation fold for three initial concentrations ω_0, for the same temperature of $T_{mean} = 45$ °C and the same pore aspect ratio of 156. For an initial concentration $\omega_0 = 10^{-5}$, the saturation plateau is reached 45 d after switching on the temperature gradient. Reducing the initial concentration to $\omega_0 = 10^{-7}$ prolongs the saturation time to 90 d. These are reasonable time ranges to establish regions of sufficiently high FA concentration to synthesize nucleobases.

Discussion

FA is a naturally occurring substance on the early Earth (2–8), with relatively high concentrations on the order of 10^{-3} wt % in shallow lakes due to the stronger evaporization of water compared with FA, which has a much higher boiling point of $T_{boil} = 210$ °C (3). Numerical finite element calculations for initial FA concentrations that correspond to early Earth shallow lake conditions reveal that FA accumulates at the bottom of hydrothermal pores with aspect ratios between about 100 and 200 and a width in the range of 100–200 μm in about 45 d to high concentrated FA solutions ($\omega \approx 85$ wt %). The conclusion from these findings is that the combination of thermophoretic mass transport and convection

Fig. 5. (A) Accumulation fold of FA as a function of the aspect ratio r for various initial weight fractions, ω_0, and temperatures as indicated in comparison with the accumulation fold for a single nucleotide. The solid line has been calculated using equation 1 in ref. 1, and the dots refer to COMSOL simulations using the physical chemistry properties of water. The accumulation fold of FA at 25 °C, 45 °C, and 75 °C has been determined at a optimal width of 180 μm, 160 μm, and 100 μm, respectively (*SI Appendix, Numerical Calculations*). All curves show an initial exponential increase, which levels of if the accumulation becomes so strong that it is close to the pure component. (B) Time-dependent study of the accumulation as a function of time for various initial concentrations, ω_0, at a width of 160 μm and a height of 25 mm. (*Inset*) Time to reach the concentration plateau, $\tau_{plateau}$, as a function of the dependence of the accumulation for different initial concentrations ω_0.

may have been at the origin of the synthesis of prebiotic nucleobases in porous rocks in contact with shallow lakes.

The numerical calculations use as an input the experimentally determined concentration and temperature dependence of the Soret coefficient and the mass diffusion coefficient, as obtained by means of IR-TDFRS, of the viscosity, the specific mass density, and the heat conductivity of FA/water mixtures. The positive value of the Soret coefficient for all concentrations leads to an initial accumulation, which is self-enhanced due to the increase of the Soret coefficient with increasing concentration.

Compared with the previously discussed hydrothermal pore accumulation of nucleotides and DNA fragments (1) by thermophoresis and convection, the accumulation of FA is slower and occurs at larger pore aspect ratios. In contrast to FA, the solubility of nucleotides is quite limited, so that the accumulation of nucleotides will be restricted. Besides shallow lakes, mineral surfaces are also proposed to play a role in the origin-of-life concept. These surfaces can act as catalysts in chemical reactions (4) and promote polymerization to form RNA (38–40). Whether such adsorption processes also play a role for the much smaller molecules considered in the present study is an open question. These conditions could affect the accumulation times.

Materials and Methods

Sample Preparation and IR-TDFRS Measurements. Solutions were prepared using FA ($\geq$99.5%; Sigma-Aldrich), without further purification, and water (Millipore). All solutions were either used immediately (measurement of contrast factors and density) or kept in a fridge for the duration of the experiment as stock solutions (for the IR-TDFRS measurements). To ascertain the stability of the mixtures against hydrolysis or other chemical reactions, the purity of the stock solution was validated by an accurate determination of the refractive index before each measurement. During the maximum storage time of 8 wk, no significant change was observed.

For the IR-TDFRS measurement, the solutions have been filtered (0.22 μm) directly into the Hellma quartz cells with an optical pass length of 0.2 mm. We used IR-TDFRS, a holographic transient grating technique, to determine the thermophoretic properties. A detailed description of the setup can be found in ref. 41. For analysis of the IR-TDFRS measurement signal, the refractive index contrast factors need to be determined. We measured the refractive index contrast factor with temperature $(\partial n/\partial T)_{p,\omega}$ at constant pressure p and the FA

weight fraction ω interferometrically (42) and found, as expected, negative values in the investigated temperature and concentration range. Its magnitude increases with higher FA concentration and decreases with increasing temperature. Measuring the refractive index for various concentrations, we determined $(\partial n/\partial\omega)_{p,T}$. It increases at higher FA concentrations and decreases with rising temperature (see also *SI Appendix, Refractive Index Contrast Measurements*).

Finite Element Calculations. To calculate the accumulation of FA in a hydrothermal pore, we solved a combination of Navier–Stokes, heat transfer, and thermodiffusion equations using commercially available finite element software (COMSOL Multiphysics 5.1). The model was built in accordance with the model of Baaske et al. (1). All calculations were done in two dimensions.

In the model, the pore was represented as a rectangle. For the heat transfer equation boundary conditions, the temperatures at the vertical walls ($T_{left} = T_{mean} - \Delta T$, $T_{right} = T_{mean} + \Delta T$) was fixed, whereas the top and the bottom of the column were considered thermally isolated. The temperature difference, ΔT, was kept at 30 K for all simulations. For the Navier–Stokes, we used the nonslip boundary conditions for all walls. For the diffusion equation, we fixed the normal flux to zero at the bottom and at the vertical walls, while, at the top of the pore, the concentration was fixed to ω_0 as in the simulations by Baaske et al. (1). Fixing the concentration at the top does not contradict the nonslip boundary conditions for the Navier–Stokes equation (top closed for the liquid flow), as the top surface can be, e.g., a porous membrane connected to an external reservoir with concentration ω_0. We took into account the temperature and concentration dependence of the thermodiffusion and mass diffusion coefficient, the viscosity, the specific mass density, the heat conductivity, and the heat capacity, part of which is taken from literature (43–49) (*SI Appendix, Temperature and Concentration Dependence*).

The calculations were done for various aspect ratios of the pore at the optimal pore width. The latter is the width of the column at which the maximum accumulation occurs (1). It was found by running a series of simulations for various widths at a specific height. The optimal width varied with temperature but was constant for all heights. We used three different mean temperatures, $T_{mean} = $ 25 °C, 45 °C, and 75 °C. The initial concentration ω_0 was taken equal to 10^{-5}, 10^{-7}, or 10^{-9} in accordance with the estimations for the FA concentration at the early Earth conditions (3).

1. Baaske P, et al. (2007) Extreme accumulation of nucleotides in simulated hydrothermal pore systems. *Proc Natl Acad Sci USA* 104(22):9346–9351.
2. Harada K (1967) Formation of amino-acids by thermal decomposition of formamide–oligomerization of hydrogen cyanide. *Nature* 214(5087):479–480.
3. Miyakawa S, Cleaves HJ, Miller SL (2002) The cold origin of life: A. Implications based on the hydrolytic stabilities of hydrogen cyanide and formamide. *Orig Life Evol Biosph* 32(3):195–208.
4. Saladino R, Crestini C, Pino S, Costanzo G, Di Mauro E (2012) Formamide and the origin of life. *Phys Life Rev* 9(1):84–104.
5. Mulkidjanian AY, Bychkov AY, Dibrova DV, Galperin MY, Koonin EV (2012) Origin of first cells at terrestrial, anoxic geothermal fields. *Proc Natl Acad Sci USA* 109(14):E821–E830.
6. Ferus M, et al. (2015) High-energy chemistry of formamide: A unified mechanism of nucleobase formation. *Proc Natl Acad Sci USA* 112(3):657–662.
7. Pino S, Sponer JE, Costanzo G, Saladino R, Mauro ED (2015) From formamide to RNA, the path is tenuous but continuous. *Life (Basel)* 5(1):372–384.
8. Sanchez RA, Ferris JP, Orgel LE (1967) Studies in prebiotic synthesis. II. Synthesis of purine precursors and amino acids from aqueous hydrogen cyanide. *J Mol Biol* 30(2):223–253.
9. de Groot S, Mazur P (1984) *Non-Equilibrium Thermodynamics* (Dover, New York).
10. Morozov KI, Köhler W (2014) Thermophoresis of polymers: Nondraining vs draining coil. *Langmuir* 30(22):6571–6576.
11. Würger A (2010) Thermal non-equilibrium transport in colloids. *Rep Prog Phys* 73(12): 126601.
12. Dhont JKG, Briels WJ (2008) Single-particle thermal diffusion of charged colloids: Double-layer theory in a temperature gradient. *Eur Phys J E Soft Matter* 25(1):61–76.
13. Wiegand S (2015) Introduction to Thermal Gradient Related Effects, eds Dhont J, et al. (Forschungszentrum Jülich, Juelich, Germany), pp F4.1–F4.24.
14. Ning H, Dhont JKG, Wiegand S (2008) Thermal-diffusive behavior of a dilute solution of charged colloids. *Langmuir* 24(6):2426–2432.
15. Wang Z, Kriegs H, Buitenhuis J, Dhont JKG, Wiegand S (2013) Thermophoresis of charged colloidal rods. *Soft Matter* 9(36):8697–8704.
16. Naumann P, et al. (2014) Isothermal behavior of the Soret effect in nonionic microemulsions: Size variation by using different n-alkanes. *J Phys Chem B* 118(12):3451–3460.
17. Parola A, Piazza R (2005) A microscopic approach to thermophoresis in colloidal suspensions. *J Phys. Condens Matter* 17(45):S3639–S3643.
18. Yang MC, Ripoll M (2012) Driving forces and polymer hydrodynamics in the soret effect. *J Phys Condens Matter* 24(19), 195101.
19. Galliero G, Volz S (2008) Thermodiffusion in model nanofluids by molecular dynamics simulations. *J Chem Phys* 128(6):064505.
20. Zhang M, Müller-Plathe F (2006) The Soret effect in dilute polymer solutions: Influence of chain length, chain stiffness, and solvent quality. *J Chem Phys* 125(12):124903.
21. Rousseau B, Nieto-Draghi C, Avalos JB (2004) The role of molecular interactions in the change of sign of the soret coefficient. *Europhys Lett* 67(6):976–982.
22. Sugaya R, Wolf BA, Kita R (2006) Thermal diffusion of dextran in aqueous solutions in the absence and the presence of urea. *Biomacromolecules* 7(2):435–440.
23. Kishikawa Y, Wiegand S, Kita R (2010) Temperature dependence of soret coefficient in aqueous and nonaqueous solutions of pullulan. *Biomacromolecules* 11(3):740–747.
24. Iacopini S, Rusconi R, Piazza R (2006) The "macromolecular tourist": Universal temperature dependence of thermal diffusion in aqueous colloidal suspensions. *Eur Phys J E Soft Matter* 19(1):59–67.
25. Maeda K, Shinyashiki N, Yagihara S, Wiegand S, Kita R (2015) Ludwig-Soret effect of aqueous solutions of ethylene glycol oligomers, crown ethers, and glycerol: Temperature, molecular weight, and hydrogen bond effect. *J Chem Phys* 143(12):124504.
26. Wang Z, Kriegs H, Wiegand S (2012) Thermal diffusion of nucleotides. *J Phys Chem B* 116(25):7463–7469.
27. Kolodner P, Williams H, Moe C (1988) Optical measurement of the soret coefficient of ethanol water solutions. *J Chem Phys* 88(10):6512–6524.
28. Polyakov P, Wiegand S (2008) Systematic study of the thermal diffusion in associated mixtures. *J Chem Phys* 128(3):034505.
29. Elola MD, Ladanyi BM (2006) Computational study of structural and dynamical properties of formamide-water mixtures. *J Chem Phys* 125(18):184506.
30. Gereben O, Pusztai L (2015) Investigation of the structure of ethanol-water mixtures by molecular dynamics simulation I: Analyses concerning the hydrogen-bonded pairs. *J Phys Chem B* 119(7):3070–3084.
31. Perera A, Mazighi R (2015) On the nature of the molecular ordering of water in aqueous DMSO mixtures. *J Chem Phys* 143(15):154502.
32. Asenbaum A, et al. (2012) Structural changes in ethanol-water mixtures: Ultrasonics, brillouin scattering and molecular dynamics studies. *Vib Spectrosc* 60:102–106.

Niether et al.

33. Kowall T, Geiger A (1994) Molecular dynamics simulation study of 18-crown-6 in aqueous solution. 1. Structure and dynamics of the hydration shell. *J Phys Chem* 98(24):6216–6224.

34. Patil K, Pawar R (1999) Near-infrared spectral studies for investigating the hydration of 18-crown-6 in aqueous solutions. *J Phys Chem B* 103(12):2256–2261.

35. Smolin N, Winter R (2004) Molecular dynamics simulations of staphylococcal nuclease: Properties of water at the protein surface. *J Phys Chem B* 108(40):15928–15937.

36. Hill A (1910) The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *J Physiol* 40:iv–vii.

37. Goutelle S, et al. (2008) The Hill equation: A review of its capabilities in pharmacological modelling. *Fundam Clin Pharmacol* 22(6):633–648.

38. Orgel LE (1998) Polymerization on the rocks: Theoretical introduction. *Orig Life Evol Biosph* 28(3):227–234.

39. Ferris JP, Hill AR, Jr, Liu R, Orgel LE (1996) Synthesis of long prebiotic oligomers on mineral surfaces. *Nature* 381(6577):59–61.

40. Franchi M, Gallori E (2005) A surface-mediated origin of the RNA world: Biogenic activities of clay-adsorbed RNA molecules. *Gene* 346:205–214.

41. Wiegand S, Ning H, Kriegs H (2007) Thermal diffusion forced Rayleigh scattering setup optimized for aqueous mixtures. *J Phys Chem B* 111(51):14169–14174.

42. Wittko G, Köhler W (2003) Precise determination of the soret, thermal diffusion and mass diffusion coefficients of binary mixtures of dodecane, isobutylbenzene and 1,2,3,4-tetrahydronaphthalene by a holographic grating technique. *Philos Mag* 83(17-18):1973–1987.

43. Egan EP, Luff BB (1966) Heat of solution heat capacity and density of aqueous formamide solutions at 25° C. *J Chem Eng Data* 11(2):194–196.

44. Akhtar S, Faruk ANMO, Saleh MA (2001) Viscosity of aqueous solutions of formamide, n-methylformamide and n,n-dimethylformamide. *Phys Chem Liq* 39(3): 383–399.

45. Wohlfarth C (2008) Viscosity of the mixture (1) water; (2) formamide. *Data Extract from Landolt-Börnstein IV/25: Viscosity of Pure Organic Liquids and Binary Liquid Mixtures*, ed Lechner MD (SpringerMaterials, Berlin), Vol 25.

46. Tobitani A, Tanaka T (1987) Predicting thermal-conductivity of binary-liquid mixtures on basis of coordination-number. *Can J Chem Eng* 65(2):321–328.

47. Checoni RF, Volpe PLO (2010) Measurements of the molar heat capacities and excess molar heat capacities for water plus organic solvents mixtures at 288.15 K to 303.15 K and atmospheric pressure. *J Solution Chem* 39(2):259–276.

48. Ganiev YA, Rastorguev YL (1968) Thermal conductivity of organic liquids. *J Eng Phys* 15(3):880–886.

49. Young HD (1992) *University Physics* (Addison Wesley, New York), 7th Ed.

3 Heuristic Approach to Understanding the Accumulation Process in Hydrothermal Pores

Heuristic Approach to Understanding the Accumulation Process in Hydrothermal Pores

Doreen Niether [1],[†] **and Simone Wiegand** [1],[2],[*],[†]

1   ICS-3 Soft Condensed Matter, Forschungszentrum Jülich GmbH, D-52428 Jülich, Germany; d.niether@fz-juelich.de

2   Department für Chemie-Physikalische Chemie, Universität zu Köln, 50939 Cologne, Germany

*   Correspondence: s.wiegand@fz-juelich.de; Tel.: +49-2461-61-6654

†   These authors contributed equally to this work.

Abstract: One of the central questions of humankind is: which chemical and physical conditions are necessary to make life possible? In this "origin-of-life" context, formamide plays an important role, because it has been demonstrated that prebiotic molecules can be synthesized from concentrated formamide solutions. Recently, it could be shown, using finite-element calculations combining thermophoresis and convection processes in hydrothermal pores, that sufficiently high formamide concentrations could be accumulated to form prebiotic molecules (Niether et al. (2016)). Depending on the initial formamide concentration, the aspect ratio of the pores, and the ambient temperature, formamide concentrations up to 85 wt % could be reached. The stationary calculations show an effective accumulation, only if the aspect ratio is above a certain threshold, and the corresponding transient studies display a sudden increase of the accumulation after a certain time. Neither of the observations were explained. In this work, we derive a simple heuristic model, which explains both phenomena. The physical idea of the approach is a comparison of the time to reach the top of the pore with the time to cross from the convective upstream towards the convective downstream. If the time to reach the top of the pore is shorter than the crossing time, the formamide molecules are flushed out of the pore. If the time is long enough, the formamide molecules can reach the downstream and accumulate at the bottom of the pore. Analysing the optimal aspect ratio as function of concentration, we find that, at a weight fraction of $w = 0.5$, a minimal pore height is required for effective accumulation. At the same concentration, the transient calculations show a maximum of the accumulation rate.

Keywords: concentration threshold; hydrothermal vents; origin of life conundrum

1. Introduction

One of the main issues about the origin of life is the question how organic material could have accumulated in the primordial ocean to reach high enough concentrations so that reactions towards larger, more complex molecules outweigh hydrolysis. For the formation of ribonucleic acid (RNA) from nucleotides, an active transport mechanism in a temperature gradient has been suggested to reach high concentrations of these prebiotic molecules [1]. The investigated accumulation process results from a combination of convection and thermophoresis inside a pore with an asymmetrical temperature profile. A probable setting for such systems in great number are porous minerals which are heated by hydrothermal vents from one side and cooled by the ocean from the other.

In our earlier paper [2], we posed the question whether the reaction from simple, anorganically formed molecules, such as hydrogen cyanide and formamide (FA), into the building blocks of RNA could be promoted by the same accumulation mechanism. Formamide has been discussed as an educt

for the formation of prebiotic molecules for almost 50 years [3–8]. Saladino et al. synthesized all nucleobases from concentrated aqueous FA solutions [5]. In our previous work, we measured the thermophoretic properties of FA in water as a function of temperature and concentration and used these data to conduct finite element calculations. We investigated how the distribution of FA in water develops as function of time in hydrothermal pores, which underlie a temperature gradient. In pores with sufficiently large aspect ratios, we found a very high FA concentration in the order of 85 wt %, which can be reached even with very small initial concentrations, as low as 10^{-7} wt %, if the time for accumulation is long enough. In a previous study, restricted to the dilute regime, Baaske et al. [1] found an exponential rise of the accumulation as function of the aspect ratio of the pore. In their study, they used the approximation $\omega(1 - \omega) \approx \omega$. Using the full expression, we identified three regimes of the accumulation fold: a weak exponential growth at low aspect ratios, a sharp rise in the intermediate range, and finally, a saturation of the accumulation fold at large aspect ratios. Independent of the initial concentration, these three regimes could always be identified, if it was possible to reach the high aspect ratios. Due to numerical instabilities this was not always the case, if the diffusion was too fast or the Soret coefficient too low. The focus of this work is to expand on our previous paper by additional time-dependent simulations and a heuristic model that explains the strong dependence of the accumulation on pore geometry as well as its progression with time.

Thermophoresis, also known as thermodiffusion or Ludwig-Soret effect, is the mass diffusion of particles or molecules induced by a temperature gradient [9]. Several theoretical approaches exist to describe thermodiffusion of polymer solutions, colloidal suspensions, and other liquid mixtures [10–13]. A good overview on the physics of the effect is given by the recent reviews by Würger [11] and by Köhler and Morozov [13], highlighting theoretical and experimental aspects of the phenomena for colloids and non-polar liquid mixtures, respectively. Further, simulations have been performed to investigate attractive and repulsive interactions between charged and uncharged colloidal particles [14,15] or to study the influence of chain length and stiffness of polymers [16]. For aqueous low molecular weight mixtures, specific interactions and the addition of salt have been investigated [17,18]. The influence of interfacial effects on the thermophoresis have been studied systematically using microemulsions [19], but it turned out that it was not possible to describe the experimental results by using existing theories [19,20]. The best agreement between experiment and theoretical concepts are found for charged spherical and rod-like colloids [21–23], but when interfacial effects like the coverage by surfactants play a role, existing theoretical concepts fail [24]. For polar liquid mixtures, such as aqueous FA solutions, there is so far no microscopic theory to describe thermophoresis. In a binary fluid mixture the mass flux in a temperature gradient can be expressed as

$$\vec{j} = -D\vec{\nabla}w - w\,(1-w)\,D_{\mathrm{T}}\vec{\nabla}T \, , \tag{1}$$

with contributions from the thermodiffusion along the temperature gradient $\sim -D_{\mathrm{T}}\,\vec{\nabla}T$ and from the Fickian diffusion along the resulting concentration gradient $D\vec{\nabla}w$. For a stationary temperature gradient, the Fickian diffusions balances after some time the thermodiffusion and a steady state is reached. This defines the Soret coefficient $S_{\mathrm{T}} = D_{\mathrm{T}}/D$, which is also a measure for the resulting concentration gradient, if a certain temperature gradient is applied. Generally, the magnitude of the Soret coefficient becomes larger, if the diffusion slows down. This implies that for slow diffusing molecules or particles smaller gradients are required to obtain the same concentration difference. Aqueous systems are of special interest due to their relevance in biotechnology. While charge contributions to the thermophoresis of solute molecules are well understood, the influence of contributions by the hydration layer are still unclear. It is known that the breaking of hydrogen bonds due to the surrounding solvent increases the Soret coefficient S_{T} of the solute molecules. To induce a breakage of hydrogen bonds, one can add an ingredient with a strong affinity to water [25] or the bonds can be disrupted by increasing the temperature [26]. This leads to a temperature dependence of

S_T that is alike for a great number of biological and synthetic molecules in water [27]. Iacopini and Piazza [28] descibed the dependence with the empirical equation

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

where S_T^∞, T^* and T_0 are fitting parameters. Recently, it became clear that the number of hydrogen-bond sites in the solute molecule is an important parameter when describing the temperature dependence of S_T and the thermodiffusion coefficient D_T. It turns out that, for solutes belonging to a homologous series, there is a linear dependence of S_T to the difference of donor and acceptor sites [29]. Although the reason for this linearity is not yet clear, it can be safely assumed that hydrogen bonding is relevant to the FA/water system as well.

2. Accumulation in Hydrothermal Pores

Figure 1 shows the investigated system, a 2D pore with width L_x and height L_y. The aspect ratio is $r = L_y/L_x$. It has been shown that finite element calculations of 3D systems give qualitatively the same results as calculations of 2D systems [1]. The reservoir with an initial concentration ω_0 is implemented by holding the concentration at the upper wall constant at ω_0. Finite element calculations with different initial concentrations were performed. The naturally occurring concentration of FA in the primordial ocean at a pH between 6–8 and an average temperature of 10 °C is estimated to be only of the order 10^{-7} wt %. In shallow lakes, where the large surface leads to faster vaporization of water and a more effective diffusion of FA from the atmosphere into the lake, higher concentrations (about 10^{-3} wt %) are estimated [4]. The left and the right walls are set to different temperatures (with a temperature difference of 30 K), which results in thermodiffusion of FA (grey arrows in Figure 1) and convection (white arrows). We solved the coupled Navier-Stokes-, diffusion-, and heat-transfer equations, using commercial finite element software (COMSOL), and determined the accumulation of FA. For the FA/water system, the whole concentration range from 0 (pure water) to 1 (pure formamide) is accessible, because they are miscible in any ratio. We measured the temperature and concentration dependence of S_T and D_T (thermal- and mass diffusion coefficient) and used the results as well as literature data of specific mass density, viscosity, heat capacity, and heat conductivity of FA/water mixtures as input for the numerical calculations [2].
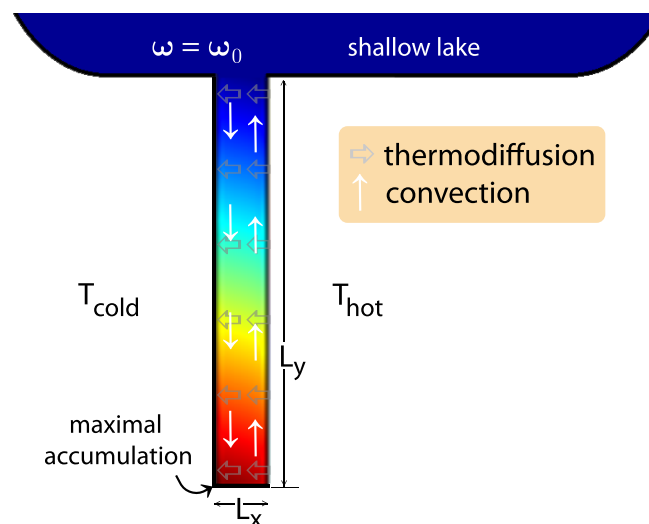

Figure 1. Contour plot of the concentration profile in a pore with aspect ratio 10 connected to a reservoir in the stationary state. The vertical and horizontal arrows mark the convective and thermodiffusive flow, respectively.

The combination of convection and thermodiffusion inside the pore results in accumulation of FA in the cold bottom corner of the pore. From an initially constant concentration ω_0 within the whole pore, a concentration profile arises as illustrated by the contour plot in Figure 1. We see the highest concentrations within the dark red region at the cold bottom corner; here, a possible formation of nucleobases from FA would take place. The concentration in this corner defines the accumulation-fold that is of interest in the following work. We analysed the dependence of FA accumulation on initial FA concentration, ambient temperature, and aspect ratio of the pore.

3. Experimental Results

We used a transient grating technique to measure the Soret coefficient S_{T} of FA in water [2]. Figure 2 displays S_{T} as a function of the weight fraction in the temperature range between 10 °C and 70 °C. Note that S_{T} is positive in the entire invesigated temperature and concentration range indicating that FA is thermophobic and accumulates in cold regions. In order to describe the concentration dependence, we use an empirical equation,

$$S_{\mathrm{T}}(\omega) = \frac{\omega^a}{K + \omega^a} + S_{\mathrm{T}}^0 \,, \tag{3}$$

with the fitting parameters a, K and S_{T}^0. A least square fit according to Equation (3), often denoted as Hill equation [30,31], is depicted as dashed lines in Figure 2. With increasing temperature, the concentration dependent slope decreases, so that we observe an intersection around a weight fraction of 0.2. This is typical and has also been found for other associated mixtures [32]. Peculiar is the fact that S_{T} of FA increases with concentration, which can support a self-enhanced accumulation process.
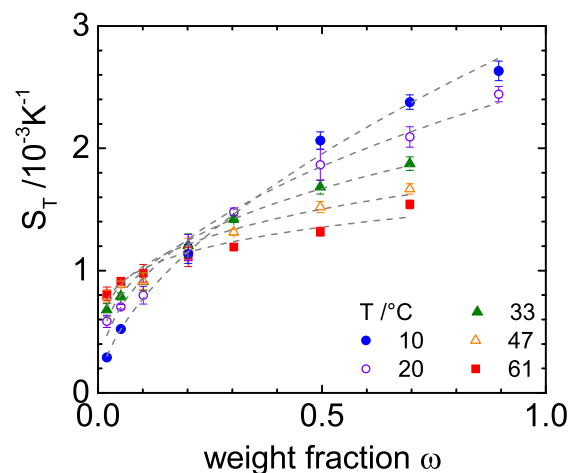
Figure 2. Soret coefficient as function of the formamide weight fraction, ω, for various temperatures. The solid lines correspond to a fit according to Equation (3).

We studied for three different mean temperatures T_{mean} = 25, 45 and 75 °C the accumulation-fold in a hydrothermal pore as function of the aspect ratio height over width. At each ambient temperature, the width of the pore was fixed to its optimal value, which decreased with increasing ambient mean temperature from 180 to 100 μm [2]. Note that an optimum is only found for very low concentrations, while at high concentrations the accumulation-fold remains on the same level for low aspect ratios. At high aspect ratios, we observe always a decrease, because in this case the convection process becomes very fast and prevents an accumulation inside the pore. In Figure 3a the accumulation-fold is displayed as a function of the aspect ratio r. We varied the initial concentration between $\omega_0 = 10^{-9}$–10^{-5}. For all studies, we observe for the accumulation-fold as function of the aspect ratio r an initial weak exponential growth, followed by a steep rise at a specific aspect ratio r^* (an expression for r^* is derived

in the heuristic model (see Equation (7))) and finally the accumulation-fold passes on to a plateau. In this regime, the accumulation-fold is approaching a value of $1/\omega_0$, so that we have almost pure FA in the lower left corner of the hydrothermal pore. It turns out that the saturation plateau is reached at lower aspect ratios for larger temperatures. At lower ambient temperatures, an accumulation in longer pores is favoured. In conclusion, we can say that the steep increase of the accumulation fold strongly depends on the magnitude of the Soret coefficient at low FA concentrations.

The time dependence of the accumulation-fold is shown in Figure 3b. We fixed the ambient temperature and pore aspect ratio to T_{mean} = 45 °C and r = 156, respectively. For the highest studied initial concentration $\omega_0 = 10^{-3}$, the saturation plateau is reached after 11 days, while for the lowest concentration of $\omega_0 = 10^{-7}$ the saturation time is reached after 90 days. For the more likley low initial concentrations between $\omega_0 = 10^{-5}$ and $\omega_0 = 10^{-7}$, the accumlations takes several months, which seems plausible. As in the case of Figure 3a, we observe at high aspect ratios ($r > r^*$) that the accumulation-fold as function of time increases steeply before it saturates. A similar behaviour can be observed for all initial concentrations. In contrast, Figure 3c shows that, for a lower aspect ratio of $r = 137.5$ ($r < r^*$), we do not observe a steep rise of the accumulation fold as function of time. This signifies that there is no influx of FA molecules into the pore, but only a slight increase of concentration in the cold bottom corner due to relocation of the FA already present inside the pore.
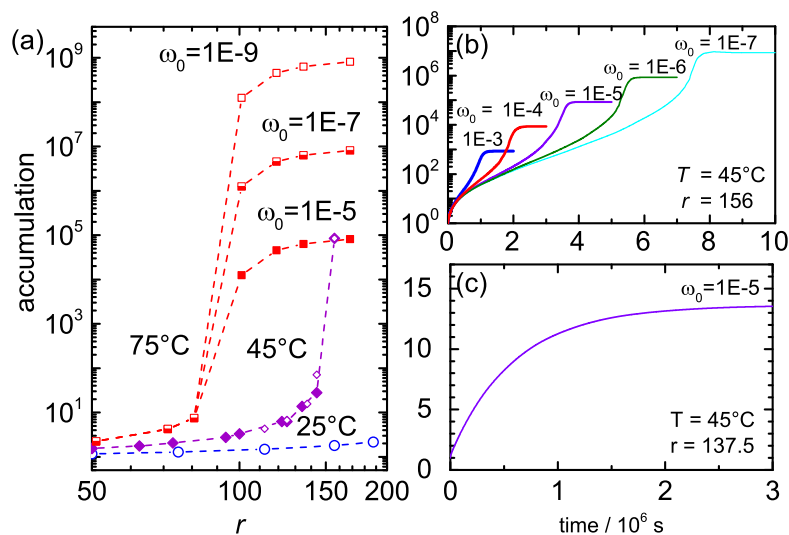
Figure 3. (**a**) Accumulation-fold of formamide as function of the aspect ratio r for various initial weight fractions, ω_0, and temperatures as indicated. The accumulation-fold of formamide at 25, 45 and 75 °C has been determined at an optimal width of 180, 160 and 100 μm, respectively [2]. All curves show an initial exponential growth and then a steep increase, which levels off when the accumulation becomes so strong that it is close to the pure component; (**b**) Time-dependent study of the accumulation as function of time at 45 °C for various initial concentrations at a high aspect ratio of $r = 156$; (**c**) Time dependence of the accumulation fold for a low aspect ratio of $r = 137.5$.

In Figure 4a the accumulation rate is plotted against time (derivative of the accumulation displayed in Figure 3b). For all concentrations, there is a clear maximum before the accumulation rate drops. Figure 4b shows the maximum concentration in the pore at the time of maximum accumulation rates. Independently of initial concentration (*x*-axis), the accumulation slows down once the concentration reaches 50 wt %. Figure 4c shows a contour plot of the flow speed and illustrates the three types of mass transport occurring inside the pore: convection, Fickian diffusion, and thermodiffusion.
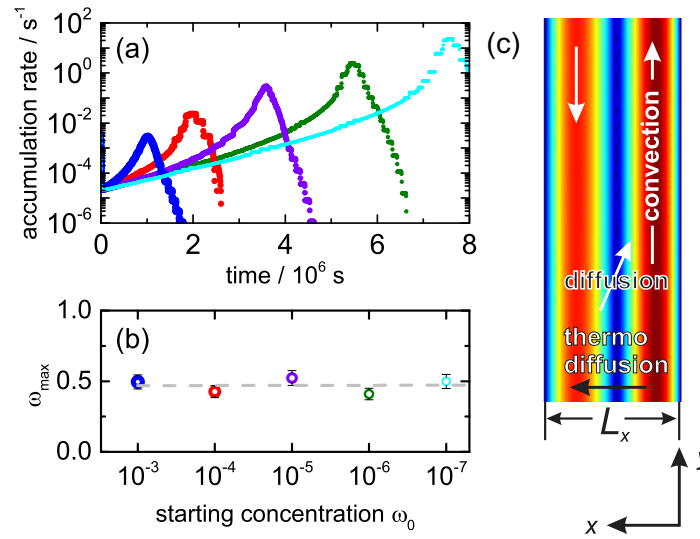
Figure 4. (**a**) Accumulation rate as function of time for the same five different initial concentrations as in Figure 3b; (**b**) Concentration at the maximum accumulation rate as function of the initial concentration, ω_0; (**c**) Illustration of the three processes, diffusion, thermal diffusion, and convection, in the hydrothermal pore with the width L_x. The colour scale illustrates flow speed.

4. Discussion

In order to answer the question, why a certain height to width ratio, r^*, is required to achieve high accumulation-folds of the order of 85 wt %, we looked at the time derivative of the accumulation fold. The determined accumulation rate versus time is shown in Figure 4a. For all simulation studies, we find a maximum of the accumulation rate at a certain time. We find for all initial concentrations that the time at which the accumulation speed shows a maximum corresponds to the same accumulated concentration of approximately 50 wt % (see Figure 4b).

To get an estimate for the height at which the accumulation becomes effective, we suggest the following *Gedankenexperiment*. The time to reach the top of the pore, which scales with its height, needs to be long enough so that the molecule can cross the middle of the pore in *x*-direction to reach the down stream on the cold side. This leads to a depletion of FA in the upper region of the pore. Once the concentration there is lower than ω_0, a diffusion of FA from the reservoir into the pore occurs. Then, an efficient accumulation in the pore becomes possible. If the pore is too short, the molecule will be flushed out into the shallow lake reservoir.

Figure 4 illustrates the different processes occurring in a 2D-hydrothermal pore. Due to convection, there is an upstream on the hot side and a down stream on the cold side in *y*-direction. The thermodiffusion is across the pore in *x*-direction and it gets stronger, if the temperature difference or the thermal diffusion coefficient become larger. The diffusion has a component in both directions. In an heuristic approach, the velocity in *x* and *y*-direction might be estimated by,

$$
\begin{aligned}
v_x &= -D \cdot \frac{\Delta_x w}{L_x} + w(1-w) \cdot D_\mathrm{T} \frac{\Delta T}{L_x} \\
v_y &= v_\mathrm{convection} + D \frac{\Delta_y w}{L_y}
\end{aligned}
\tag{4}
$$

with the convection velocity $v_\mathrm{convection}$ and the temperature difference $\Delta T = T_\mathrm{cold} - T_\mathrm{warm}$ between cold and warm side. In this simplified approach, we assume a linear change of the concentrations between top and bottom and between cold and warm side with $\Delta_x w = w_\mathrm{cold} - w_\mathrm{warm}$ and $\Delta_y w = w_\mathrm{bottom} - w_\mathrm{top}$, respectively. Note that in a full description of the problem the concentration w depends on x or y. For a symmetrical velocity profile, we assume that the accumulation becomes effective, when the particles, which are located on average in the middle of the upstream are moving at

least a distance of $L_x/4$ to reach the downstream before they reach the end of the column, L_y. Therefore, the time to leave the hydrothermal pore L_y/v_y needs to be larger than the time $(L_x/4)/v_x$ to cross a quarter of the pore width,

$$\frac{L_y}{v_y} > \frac{(L_x/4)}{v_x} \tag{5}$$

so follows the inequality for the height ratio

$$\frac{L_y}{L_x} > \frac{v_y}{4 \cdot v_x}. \tag{6}$$

Using the expressions for the velocities (Equation (4)) and assuming that the concentration difference in x direction is small compared to the concentration difference in y-direction, we then find

$$(r^*)^2 = \left(\frac{L_y}{L_x}\right)^2 > \frac{L_y \cdot v_{\text{convection}} + D \cdot \Delta_y w}{4 \cdot w \left(1 - w\right) D_{\text{T}} \Delta T}. \tag{7}$$

Solving Equation (7) for $\Delta_y w$ and differentiating with respect to the concentration w, we find a maximum at $w = 0.5$. This result is in agreement with our transient simulations. It indicates the strongest accumulation around a concentration of 50 wt % (see Figure 4a,b). Further, Equation (7) shows that the height ratio has to increase, if the temperature difference ΔT or the thermal diffusion coefficient D_{T} decreases. Additionally, an increasing diffusion coefficient requires a longer column, because due to the faster diffusion the molecules can leave the pore to the reservoir before reaching the down stream. This simple heuristic model gives an intuitive understanding, why the numerical simulations show for some height ratios a steep increase in the accumulation rate.

5. Conclusions

On the early earth, the estimated formamide concentrations were fairly low of the order of 10^{-7}–10^{-3} wt % [3–8,33]. The highest concentrations were probably reached in lakes with depths of up to 10 m. In these regions, the formamide concentration increased because of preferential vaporization of water, with a roughly two times lower boiling point compared to formamide [4]. Recently, numerical finite-element calculations for initial FA concentrations between 10^{-7}–10^{-3} wt % have been performed [2], which showed that FA could be concentrated at the bottom of hydrothermal pores in about 1–3 months. The possible FA concentrations were around $\omega \approx 85$ wt %. As an input for the numerical calculations, we used experimentally determined thermal and mass diffusion coefficients, measured over a wide concentration and temperature range, as well as literature data of thermophysical properties of the FA/water mixture. Note that in the entire experimental range, the Soret coefficient S_{T} of FA is positive and it is increasing with concentration, leading to a self-enhanced accumulation. The heights and widths of the simulated pores were in the range of 5–35 mm and 100–200 μm, respectively, resulting in aspect ratios between 50–200. The open question of the previous work was, why the aspect ratio had to be above a certain threshold to reach an effective accumulation of FA in the pores. At these aspect ratios, the transient calculations also showed a steep increase of the accumulation at a certain time, which was not observed at the lower aspect ratios.

Compared to the previous study [2] we examined more closely how the accumulation process works and how it is influenced by the pore geometry. Using an heuristic model we could show that the accumulation becomes effective if the aspect ratio is so large that FA molecules diffuse into the convective down stream on the cold side of the pore before the convective upstream flushes FA into the reservoir outside of the pore. Analysing the optimal aspect ratio as function of concentration, we find a minimal aspect ratio at a weight fraction of $w = 0.5$. The result agrees with the analysis of our transient simulations, which show a maximal accumulation at $w = 0.5$.

The here proposed times for accumulation of formamide in hydrothermal pores should only be considered as an order of magnitude. The accumulation times will be influenced by many other

unknown factors such as pore dimensions, porosity and temperature fluctuations. Beside shallow lakes, also mineral surfaces are proposed for the "origin-of-life" conundrum. They can work as catalyst in the synthesis of small biological molecules [5] and initiate polymerization [34,35]. Adsorption leads to enrichment on the surface and it has been shown that adsorption slows down the degradation of RNA [36]. Mechanisms proposed in this context might also be applicable for other chemical reactions discussed. In principle, the thermophoretic accumulation is combinable with the concept of surface adsorption since the hydrothermal pores would offer a large mineral surface area that might promote reactions or retain relatively high concentrations of material when the temperature gradient that drives the accumulation decays. For future work, it would be interesting to realize this scenario experimentally.

6. Materials and Methods

In the following two subsection the experimental and simulation details are given.

6.1. Sample Preparation and IR-TDFRS Measurements

We used formamide with a purity of better than 99.5% (Sigma-Aldrich, Steinheim am Albuch, Germany) and water from a Millipore filtration system. Part of the prepared solutions were stored as stock solutions in a fridge. To ensure the stability of the mixtures, we validated the refractive index before each measurement. We did not observe significant changes over the maximum storage time of 8 weeks.

For the Infra-Red Thermal Diffusion Forced Rayleigh Scattering (IR-TDFRS) measurements, the solutions had to be free of dust. By filtering the solutions with a membrane filter with a diameter of 0.22 μm directly into the Hellma quartz cells, we avoided contamination of the solutions with dust particles in the air. The optical path length was only 0.2 mm to minimize convection effects. Additionally, we performed measurements at different laser intensities to ensure that the measurements were not influenced by convection. Further details about the set-up can be found elsewhere [37] and details about the measurements of the refractive index contrast factors with temperature and concentration can be found in the supporting information of Reference [2].

6.2. Finite Element Calculations

Using the commercially available finite element software (COMSOL Multiphysiscs 5.1) we solved a combination of thermodiffusion-, heat transfer-, and Navier-Stokes-equations to calculate the spatial- and time distribution of FA in a hydrothermal pore. We compared our model with Baaske et al. [1] calculating the accumulation of diluted nucleotide solutions. In the diluted regime we obtained the same results, but at higher concentrations deviations were observed. It turned out that the model by Baaske et al. was limited to very diluted solutions, because they used the approximation $w(1 - w) \approx w$ in the thermophoretic flux equation. In our model, we used the full expression and all calculations were done in 2D.

The temperature difference across the rectangular pore ΔT was kept to 30 K for all simulations. The temperatures at the cold side of the pore was fixed to $T_{\text{left}} = T_{\text{mean}} - \Delta T$ and at the warm side to $T_{\text{right}} = T_{\text{mean}} + \Delta T$. At top and bottom of the pore, we assumed thermal isolation so that the temperatures were given by the solution of the heat equation. For all walls, non-slip boundary conditions were used. Further, the normal flux at the bottom and at the side walls was fixed to zero. To simulate the connection to the reservoir we fixed the concentration at the top of the pore to ω_0. This approach is in accordance with the simulations by Baaske et al. [1]. Due to this assumptions, it is possible to suck FA molecules into the pore. Note that a fixed concentration at the top is in agreement with the non-slip boundary conditions we use for the Navier-Stokes equations, because it is possible to connect the top of the pore by using for instance a porous membrane with a reservoir with concentration ω_0. Due to the use of a suitable membrane the flow velocity can approach zero by simultaneously allowing a material exchange with the reservoir. The thermo- and mass diffusion

coefficient had been determined by IR-TDFRS [2], while the other thermophysical properties have been taken from literature [38–44]. We accounted for the temperature and concentration dependence of all properties and the temperature and concentration dependent equations can be found in the supporting information of Reference [2].

We carefully checked the stability of the stationary and transient calculations for different mesh sizes. It turned out the numerical calculations were stable, when the mesh size was in the order of 3 μm between junctions. As was shown in previous studies [1], there is an optimal width of the pore for maximum accumulation, independent from height. We ran calculation at a fixed height to find the optimal width and observed that it varies with temperature. Then, we ran height and time dependent calculations with the respective optimal widths at three different mean temperatures, $T_{mean} = 25\,°C$, $45\,°C$ and $75\,°C$. Different initial weight fractions of formamide ω_0 in the range from 10^{-9} to 10^{-3} were used, taking into account estimations of concentrations during the early history of the earth [4].

Author Contributions: Doreen Niether and Simone Wiegand conceived and designed the experiments; Doreen Niether performed the experiments; Doreen Niether and Simone Wiegand analyzed the data; Doreen Niether and Simone Wiegand wrote the paper. Both authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

IR-TDFRS   infrared thermal diffusion forced Rayleigh scattering
FA   formamide
RNA   ribonucleic acid
DNA   deoxyribonucleic acid

References

1. Baaske, P.; Weinert, F.M.; Duhr, S.; Lemke, K.H.; Russell, M.J.; Braun, D. Extreme Accumulation of Nucleotides in Simulated Hydrothermal Pore Systems. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 9346–9351.
2. Niether, D.; Afanasenkau, D.; Dhont, J.K.G.; Wiegand, S. Accumulation of Formamide in Hydrothermal Pores to Form Prebiotic Nucleobases. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 4272–4277.
3. Harada, K. Formation of Amino-Acids by Thermal Decomposition of Formamide—Oligomerization of Hydrogen Cyanide. *Nature* **1967**, *214*, 479–480.
4. Miyakawa, S.; Cleaves, H.J.; Miller, S.L. The Cold Origin of Life: A. Implications Based on the Hydrolytic Stabilities of Hydrogen Cyanide and Formamide. *Orig. Life Evol. Biosph.* **2002**, *32*, 195–208.
5. Saladino, R.; Crestini, C.; Pino, S.; Costanzo, G.; di Mauro, E. Formamide and the Origin of Life. *Phys. Life Rev.* **2012**, *9*, 84–104.
6. Mulkidjanian, A.Y.; Bychkov, A.Y.; Dibrova, D.V.; Galperin, M.Y.; Koonin, E.V. Origin of First Cells at Terrestrial, Anoxic Geothermal Fields. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E821–E830.
7. Ferus, M.; Nesvorny, D.; Sponer, J.; Kubelik, P.; Michalcikova, R.; Shestivska, V.; Sponer, J.E.; Civis, S. High-Energy Chemistry of Formamide: A Unified Mechanism of Nucleobase Formation. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 657–662.
8. Pino, S.; Sponer, J.E.; Costanzo, G.; Saladino, R.; di Mauro, E. From Formamide to RNA, the Path Is Tenuous but Continuous. *Life* **2015**, *5*, 372–384.
9. De Groot, S.; Mazur, P. *Non-Equilibrium Thermodynamics*; Dover: New York, NY, USA, 1984.
10. Morozov, K.I.; Köhler, W. Thermophoresis of Polymers: Nondraining vs. Draining Coil. *Langmuir* **2014**, *30*, 6571–6576.

11. Würger, A. Thermal Non-Equilibrium Transport in Colloids. *Rep. Prog. Phys.* **2010**, *73*, 126601.

12. Dhont, J.K.G.; Briels, W.J. Single-Particle Thermal Diffusion of Charged Colloids: Double-Layer Theory in a Temperature Gradient. *Eur. Phys. J. E* **2008**, *25*, 61–76.

13. Köhler, W.; Morozov, K.I. The Soret Effect in Liquid Mixtures—A Review. *J. Non-Equilb. Thermodyn.* **2016**, *41*, 151–197.

14. Yang, M.C.; Ripoll, M. Driving Forces and Polymer Hydrodynamics in the Soret Effect. *J. Phys. Condens. Matter* **2012**, *24*, 195101.

15. Galliero, G.; Volz, S. Thermodiffusion in Model Nanofluids by Molecular Dynamics Simulations. *J. Chem. Phys.* **2008**, *128*, 064505.

16. Zhang, M.; Müller-Plathe, F. The Soret Effect in Dilute Polymer Solutions: Influence of Chain Length, Chain Stiffness, and Solvent Quality. *J. Chem. Phys.* **2006**, *125*, 124903.

17. Rousseau, B.; Nieto-Draghi, C.; Avalos, J.B. The Role of Molecular Interactions in the Change of Sign of the Soret Coefficient. *Europhys. Lett.* **2004**, *67*, 976–982.

18. Römer, F.; Wang, Z.; Wiegand, S.; Bresme, F. Alkali Halide Solutions under Thermal Gradients: Soret Coefficients and Heat Transfer Mechanisms. *J. Phys. Chem. B* **2013**, *117*, 8209–8222.

19. Naumann, P.; Datta, S.; Sottmann, T.; Arlt, B.; Frielinghaus, H.; Wiegand, S. Isothermal Behavior of the Soret Effect in Nonionic Microemulsions: Size Variation by Using Different *n*-Alkanes. *J . Phys. Chem. B* **2014**, *118*, 3451–3460.

20. Parola, A.; Piazza, R. A Microscopic Approach to Thermophoresis in Colloidal Suspensions. *J. Phys. Condens. Matter* **2005**, *17*, S3639–S3643.

21. Ning, H.; Dhont, J.K.G.; Wiegand, S. Thermal-Diffusive Behavior of a Dilute Solution of Charged Colloids. *Langmuir* **2008**, *24*, 2426–2432.

22. Wang, Z.; Kriegs, H.; Buitenhuis, J.; Dhont, J.K.G.; Wiegand, S. Thermophoresis of Charged Colloidal Rods. *Soft Matter* **2013**, *9*, 8697–8704.

23. Sehnem, A.L.; Neto, A.M.F.; Aquino, R.; Campos, A.F.C.; Tourinho, F.A.; Depeyrot, J. Temperature Dependence of the Soret Coefficient of Ionic Colloids. *Phys. Rev. E* **2015**, *92*, 042311.

24. Syshchyk, O.; Afanasenkau, D.; Wang, Z.; Kriegs, H.; Buitenhuis, J.; Wiegand, S. Influence of Temperature and Charge Effects on Thermophoresis of Polystyrene Beads. *Eur. Phys. J. E* **2016**, *39*, doi:10.1140/epje/i2016-16129-y.

25. Sugaya, R.; Wolf, B.A.; Kita, R. Thermal Diffusion of Dextran in Aqueous Solutions in the Absence and the Presence of Urea. *Biomacromolecules* **2006**, *7*, 435–440.

26. Wang, Z.; Kriegs, H.; Wiegand, S. Thermal Diffusion of Nucleotides. *J. Phys. Chem. B* **2012**, *116*, 7463–7469.

27. Kishikawa, Y.; Wiegand, S.; Kita, R. Temperature Dependence of Soret Coefficient in Aqueous and Nonaqueous Solutions of Pullulan. *Biomacromolecules* **2010**, *11*, 740–747.

28. Iacopini, S.; Rusconi, R.; Piazza, R. The "Macromolecular Tourist": Universal Temperature Dependence of Thermal Diffusion in Aqueous Colloidal Suspensions. *Eur. Phys. J. E* **2006**, *19*, 59–67.

29. Maeda, K.; Shinyashiki, N.; Yagihara, S.; Wiegand, S.; Kita, R. Ludwig-Soret Effect of Aqueous Solutions of Ethylene Glycol Oligomers, Crown Ethers, and Glycerol: Temperature, Molecular Weight, and Hydrogen Bond Effect. *J. Chem. Phys.* **2015**, *143*, 124504.

30. Hill, A. The Possible Effects of the Aggregation of the Molecules of Haemoglobin on Its Dissociation Curves. *J. Physiol.* **1910**, *40*, 4–7.

31. Goutelle, S.; Maurin, M.; Rougier, F.; Barbaut, X.; Bourguignon, L.; Ducher, M.; Maire, P. The Hill Equation: A Review of Its Capabilities in Pharmacological Modelling. *Fundam. Clin. Pharmacol.* **2008**, *22*, 633–648.

32. Polyakov, P.; Wiegand, S. Systematic Study of the Thermal Diffusion in Associated Mixtures. *J. Chem. Phys.* **2008**, *128*, 034505.

33. Sanchez, R.A.; Ferris, J.P.; Orgel, L.E. Studies in Prebiotic Synthesis: 2. Synthesis of Purine Precursors and Amino Acids from Aqueous Hydrogen Cyanide. *J. Mol. Biol.* **1967**, *30*, 223–253.

34. Orgel, L.E. Polymerization on the Rocks: Theoretical Introduction. *Orig. Life Evol. Biosph.* **1998**, *28*, 227–234.

35. Ferris, J.P.; Hill, A.R.; Liu, R.H.; Orgel, L.E. Synthesis of Long Prebiotic Oligomers on Mineral Surfaces. *Nature* **1996**, *381*, 59–61.

36. Franchi, M.; Gallori, E. A Surface-Mediated Origin of the RNA World: Biogenic Activities of Clay-Adsorbed RNA Molecules. *Gene* **2005**, *346*, 205–214.

37. Wiegand, S.; Ning, H.; Kriegs, H. Thermal Diffusion Forced Rayleigh Scattering Setup Optimized for Aqueous Mixtures. *J. Phys. Chem. B* **2007**, *111*, 14169–14174.

38. Egan, E.P.; Luff, B.B. Heat of Solution Heat Capacity and Density of Aqueous Formamide Solutions at 25 °C. *J. Chem. Eng. Data* **1966**, *11*, 194–196.

39. Akhtar, S.; Faruk, A.N.M.O.; Saleh, M.A. Viscosity of aqueous solutions of formamide, *N*-methylformamide and *N*,*N*-dimethylformamide. *Phys. Chem. Liq.* **2001**, *39*, 383–399.

40. Wohlfarth, C. Viscosity of the mixture (1) water; (2) formamide. In *Supplement to IV/18*; Lechner, M.D., Ed.; Landolt-Börnstein—Group IV Physical Chemistry; Springer: Berlin/Heidelberg, Germany, 2009; pp. 709–711.

41. Tobitani, A.; Tanaka, T. Predicting Thermal-Conductivity of Binary-Liquid Mixtures on Basis of Coordination-Number. *Can. J. Chem. Eng.* **1987**, *65*, 321–328.

42. Checoni, R.F.; Volpe, P.L.O. Measurements of the Molar Heat Capacities and Excess Molar Heat Capacities for Water plus Organic Solvents Mixtures at 288.15 K to 303.15 K and Atmospheric Pressure. *J. Solut. Chem.* **2010**, *39*, 259–276.

43. Ganiev, Y.A.; Rastorguev, Y.L. Thermal Conductivity of Organic Liquids. *J. Eng. Phys.* **1968**, *15*, 519–525.

44. Young, H.D. *University Physics*; Addison-Wesley: Reading, MA, USA, 1992.

4 Unravelling the Hydrophobicity of Urea in Water Using Thermodiffusion: Implications for Protein Denaturation

Check for updates

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

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

Urea is widely used as a protein denaturant in aqueous solutions. Experimental 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 make it difficult to connect the denaturation effects to the 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 ~ 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 a protein denaturant, as well as its influence on the structure and dynamics of water have 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 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 a 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

[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, UK. E-mail: f.bresme@imperial.ac.uk

[c] Department für Chemie-Physikalische Chemie, Universität zu Köln, 50939 Cologne, Germany

[d] Department of Chemistry, Norwegian University of Science and Technology, NTNU, Trondheim, Norway

† Electronic supplementary information (ESI) available: Additional experimental and simulation results. See DOI: 10.1039/c7cp05843h

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

expected to slow down the orientational dynamics of water (about six times with respect to bulk).[1] Computer simulations have shown that the orientational dynamics of water depends on the urea concentration,[2,20] and have reported a variety of structural changes in the solution, which become more evident at high concentrations of 4–6 M.[8–12] In contrast, Nuclear Magnetic Resonance experiments of concentrated solutions, 10 M, have been 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 the literature can vary significantly.[22] Furthermore, 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 the 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] ethylene glycol oligomers,[29] or formamide at high concentrations.[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 for systematically studying 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 behaviour. We take full advantage of the 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 behaviour signals a transition from the hydrophilic to the hydrophobic response of the solute.

Results and discussion

Fig. 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 of 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 eqn (2). However, this trend is inverted at high urea concentrations (>5.4 M), and the
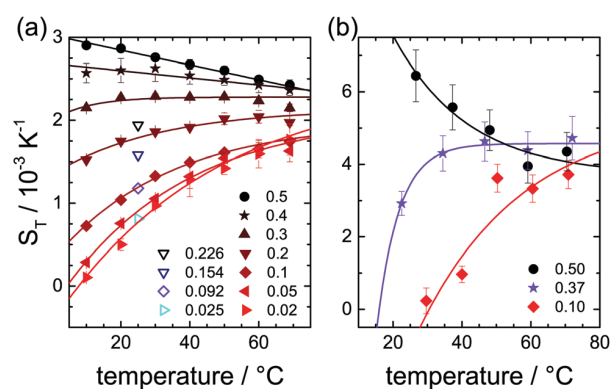

Fig. 1 Soret coefficients of urea–water mixtures obtained from (a) IR-TDFRS experiments and (b) NEMD simulations, as a function of temperature, for different urea concentrations, given in weight fractions (see legend). The open symbols in the left panel were reported by Story and Turner[36] using a different experimental method, thermogravitational columns. The solid lines are fittings of the numerical data (see eqn (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 °C and that at high temperature ($S_T(0) - S_T^\infty$), and T_0 which indicates the temperature dependence of S_T.

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 the slope is observed at a slightly lower concentration, ~ 20 wt% *vs.* ~ 30 wt% for urea. The origin of the inversion behaviour is currently unknown. The concentrations at which we observe the inversion of the thermophoretic response of urea/water and formamide/water mixtures are within the concentration range leading to protein denaturation using a co-solvent[37] (see the ESI† 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 (*cf.* Fig. 1(a) and (b)), even if some quantitative deviations are present. To the 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 a higher urea concentration, 77 wt%, the thermophobic response is stronger and decreases with temperature. We show in Fig. 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 concentrations. 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 concentrations, even if additional layers of atoms start to become evident (see hydrogen contribution for U–U and W–U in Fig. 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 Fig. 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 a 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.
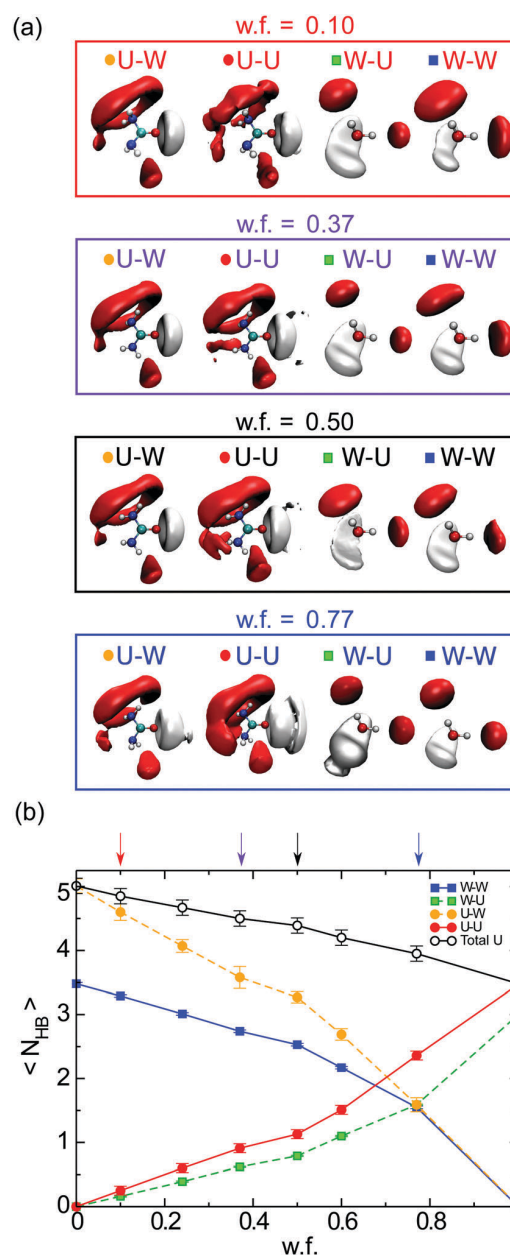

Fig. 2 (a) Isodensity surface enclosing 70% of the hydrogen bonded atoms around the reference molecule, namely urea (U–U and U–W) and water (W–U and W–W). A snapshot of the reference molecule is represented in each case. The red and white clouds indicate the oxygen and hydrogen densities, respectively, 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).

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 Fig. 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 of water, which are very similar at all investigated urea concentrations.

Although the three-dimensional density profile offers insights into the local structure around urea and water molecules, it does not provide information on the overall structuring of the 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 the ESI†). 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 Fig. 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 dependencies on 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 them from the original force-field, U1. We expect that increasing the solute–solvent interactions will increase the magnitude of the chemical potential, and make urea molecules more ''hydrophilic'', and possibly it will reduce the tendency for aggregation. We expect that a higher affinity for 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 in the water–urea interactions (U2) results in a decrease of both size and number of the cluster (see the ESI†). 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. = 0.10 (see Fig. 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 the slope of the Soret coefficient with temperature. We show in Fig. 4(a) the dependence of the Soret coefficient on 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 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. Water molecules are 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 penta-coordination 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 Fig. 4(b) that $dS_T/dT > 0$ is consistently shifted to slightly positive values, similar to those obtained for a solution with a lower concentration that has approximately the same coordination number for urea–water correlations. Hence, we conclude that
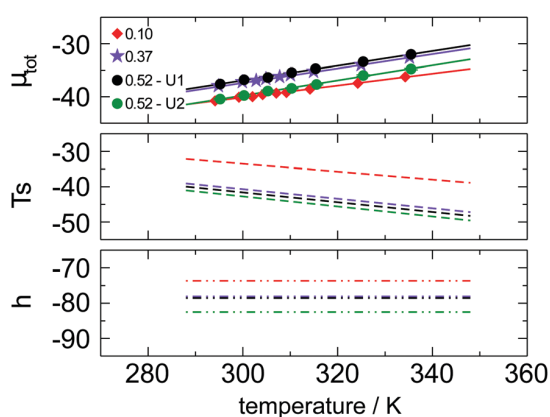

Fig. 3 Solvation free energy (μ_{tot} in kJ mol^{-1}), enthalpy (h in kJ mol^{-1}) and entropy (Ts in kJ mol^{-1} K^{-1}) of urea in water, as a function of temperature, for three different urea weight fractions: 0.10, 0.37 and 0.50. For the highest urea concentration we increased the strength of the urea–water interactions by a factor of two. The colour scheme in the lines in the middle and bottom panel have the same meaning as in the labels in the top panel.
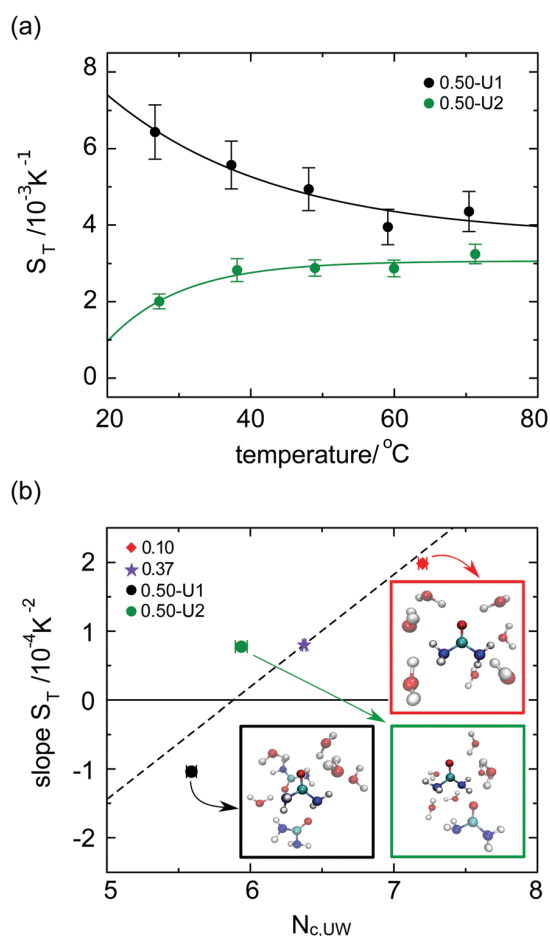
Fig. 4 (a) Dependence of the Soret coefficient of urea on temperature for two different urea–water interaction strengths. U2 is twice as strong as U1. (b) Slope of the Soret coefficient, dS_T/dT, as a function of the number of water molecules in the first solvation shell of the urea molecules. All the results correspond to a temperature of 30 °C.
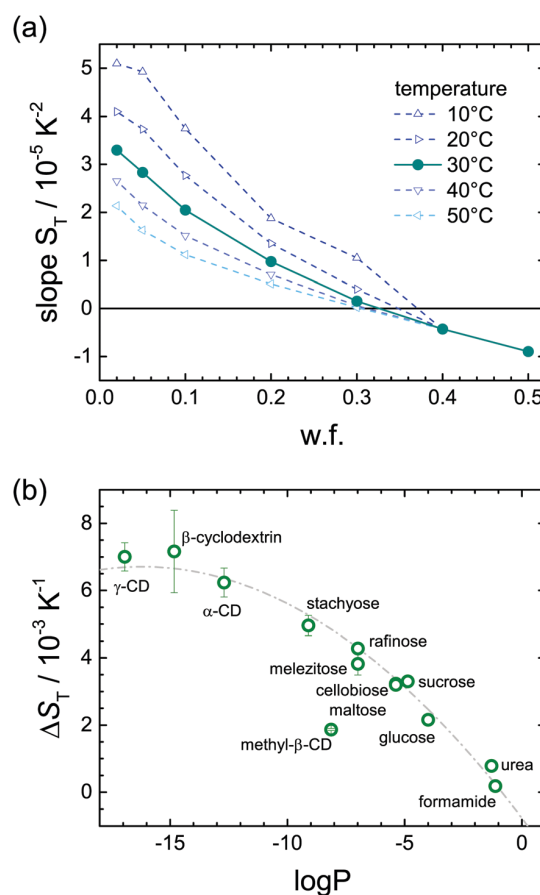

Fig. 5 (a) Slope of the Soret coefficient, dS_T/dT, as a function of the urea weight fraction for different temperatures, as specified in the legend. (b) $\Delta S_T = S_T(50\ °C) - S_T(20\ °C)$ against $\log P$ for cyclodextrins,[40] oligosaccharides,[41,42] formamide[30] and urea at low concentrations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

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, $\Delta S_T = S_T(50\ °C) - S_T(20\ °C)$, which defines the sign of dS_T/dT, and the degree of hydrophilicity, through $\log P$, of different solutes at similar concentrations (see Fig. 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). Fig. 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 Fig. 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 of 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 Fig. 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 concentrations and it is relatively hydrophilic, it will remain well dissolved in water. However, when the concentration increases, 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 identifying the transition from the 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 for testing these ideas and further investigations are needed along these lines to validate this picture.

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 is 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 the 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 the sign, *i.e.* $S_T = 0$. Eqn (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 of the refractive index contrast measurements necessary to evaluate the TDFRS data are provided in the ESI.†

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 ref. 34 and 50. We used a modified version of the code GROMACS v. 4.6.3.[51] A typical snapshot of the system is reported in Fig. S6 in the ESI,† 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 the 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 for 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 a simulation time of 2.7 μs. The trajectories were employed to construct the temperature, density and concentration profiles, by dividing the simulation box into 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,

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

where N the number of molecules of 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 fraction of urea, x_u, and the temperature gradients:

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

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 ESI† representative particle number profiles of 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 of urea solutions for all the concentrations considered in this paper. We found good agreement with both experimental and computed data (see Fig. S4 in the ESI†). 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 with adding a molecule of urea in the aqueous solution, at a specific urea concentration. We follow the procedure used in ref. 34 to calculate the chemical potential of ions. The chemical potential is defined by the sum of 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 $\lambda_{vdW} = 0...1$, where 0 and 1 correspond to zero and full van der Waals interactions, respectively. Subsequently we charged the molecules using 20 coupling parameters, $\lambda_c = 0...1$, where again 1 corresponds 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 barostat, with a time constant of 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 between 22 and 62 °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,

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

To compute the solvation entropy and enthalpy, we fitted our chemical potentials to a linear equation which accurately reproduces the computed data. The fitting parameters are reported in the ESI.†

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 $30°$.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

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. Elber and D. Thirumalai, *Biopolymers*, 2003, **68**, 359–369.

20 P. O. 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: Soft Matter Biol. Phys.*, 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: Soft Matter Biol. Phys.*, 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*, Dover, New York, 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: Thermal Nonequilibrium Phenomena in Fluid Mixtures*, 2002, vol. 584, pp. 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.

5 Thermophoresis of Cyclic Oligosaccharides in Polar Solvents

Regular Article

Thermophoresis of cyclic oligosaccharides in polar solvents[*]

Kazuya Eguchi[1], Doreen Niether[2,a], Simone Wiegand[2,3], and Rio Kita[1,4,b]

[1] School of Science, Tokai University, Hiratsuka, Kanagawa 259-1292, Japan
[2] ICS-3 Soft Condensed Matter, Forschungszentrum Juelich GmbH, D-52428 Juelich, Germany
[3] Chemistry Department – Physical Chemistry, University Cologne, D-50939 Cologne, Germany
[4] Micro/Nano Technology Center, Tokai University, Hiratsuka, Kanagawa 259-1292, Japan

Abstract. Cyclodextrins are cyclic oligosaccharides which are interesting as drug delivery systems, because they can be used as containers for pharmaceutical substances. We studied the Ludwig-Soret effect of α-, β-, γ- and methyl-β-cyclodextrin in water and formamide by infrared thermal diffusion forced Rayleigh scattering (IR-TDFRS). In water the Soret coefficient, S_T, of α-, β- and γ-cyclodextrin increases with increasing temperature and shows a sign change from negative to positive around $T = 35\,°C$, while S_T of methyl-β-cyclodextrin is positive in the entire investigated temperature. In formamide S_T-values of all cyclodextrins coincide and show a slight decrease with temperature. We discuss the obtained results and relate the S_T-values to the different hydrogen bonding capabilities of the cyclodextrins and the used solvents. It turns out that the change of S_T with temperature correlates with the partition coefficient, $\log P$, which indicates that more hydrophilic substances show a more pronounced temperature sensitivity of S_T. Additionally we obtained a surprising result measuring the refractive index contrast factor with temperature, $(\partial n/\partial T)_{c,p}$ of cyclodextrins in formamide, which might be explained by a complex formation between cyclodextrins and formamide.

1 Introduction

Thermodiffusion is the migration of molecules under a temperature gradient. For binary fluid mixtures the flux, $\boldsymbol{J}$, of the solute can be described as [1]

$$\boldsymbol{J} = -\rho D \nabla c - \rho c(1-c) D_T \nabla T, \qquad (1)$$

where ρ is the density. The second term expresses the thermal diffusion along the temperature gradient, ∇T, with the thermal diffusion constant, D_T, and the first term describes the Fickian diffusion along the resulting concentration gradient of the solute, ∇c, with the diffusion constant, D. In a stable temperature gradient a steady state with $\boldsymbol{J} = 0$ is reached. The resulting concentration difference, Δc, at an applied temperature difference, ΔT, is proportional to the ratio of D_T and D, which defines the Soret coefficient of the solute

$$S_T \equiv \frac{D_T}{D} = -\frac{1}{c(1-c)} \frac{\Delta c}{\Delta T}. \qquad (2)$$

[a] e-mail: `d.niether@fz-juelich.de`
[b] e-mail: `rkita@keyaki.cc.u-tokai.ac.jp`

A larger S_T implies a larger concentration gradient for a given temperature gradient. Note also that S_T is larger for slow diffusing compounds such as colloids or small molecules in porous media due to the slower diffusion. There are many practical applications which are influenced by thermodiffusion in different areas such as petroleum reservoirs [2], solar ponds [3], polymer fractionation [4], biopolymers [5] and prebiotic evolution [6]. An extended discussion about the various application areas can be found in the recent review by Köhler and Morozov [7].

Cyclodextrins (CDs) are cyclic oligosaccharides. The glucopyranose units (6, 7 or 8 in α-, β- and γ-cyclodextrin, respectively) are linked by α-$(1,4)$ glycosidic bonds. The molecules arrange in a torus-like shape with the hydroxyl groups at the outer edges to form a hydrophilic exterior while the cavity is lined with apolar parts of the molecule to form a hydrophobic interior [8]. Due to the formation of this "sub-micron heterogeneous environment" small compounds can enter the cavity and form inclusion complexes with the CD (fig. 1). The cavity diameter varies depending on ring size which leads to differences in the complex formation for the various CDs [9].

The propensity for complex formation is used in pharmaceutical industry to enhance solubility and control speed of uptake of drugs [8]. In this context a better
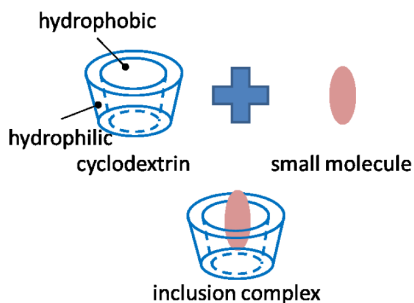
Fig. 1. Schematic drawing of the formation of an inclusion complex formed by cyclodextrins.

Fig. 2. Sketch of the investigated cyclodextrins and chemical structure of un-methylated (left) and methylated (right) glucopyranose units. Red and blue arrows show hydrogen bond donor and acceptor sites, respectively. In methyl-β-CD 55% of the hydroxyl-H are randomly substituted with methyl group. These sites will not form hydrogen bonds (grey arrow). Further explanations are in the text.

understanding of the thermodiffusion behaviour of CDs and their drug-complexes would be advantageous. It is unclear if or how thermodiffusion contributes to the transport processes in the human body. The idea of controlled movement of drugs into certain, *e.g.* inflamed, areas of the bodies is attractive, but maybe not realistic. A more feasible application would be to tune the complexes so that they show no thermophobic behaviour which might lead to a depletion of the drug complexes in areas of inflammation. Another interesting aspect of CDs is the possibility that the interior of the ring and the CD or the attached side groups can be varied independently. This makes it possible to do systematic experiments to differentiate draining and non-draining effects in aqueous systems, which makes it possible to validate, whether concepts derived for non-polar systems are applicable for aqueous mixtures [10].

The theoretical understanding of the thermophoretic behavior of non-polar systems [7] is much more comprehensive than for aqueous mixtures. For non-polar substances systematic studies of isotopic and other mixtures identified three contributions to the Soret coefficient stemming from differences in the molecular masses and moments of inertia and an additional *chemical contribution*. Due to the specific interactions, *e.g.* hydrogen bonds, the *chemical contribution* is the dominant factor for aqueous systems, which changes with temperature and influences the thermodiffusion of aqueous mixtures. To describe this temperature dependence Iacopini and Piazza [11] proposed an empirical equation

$$S_{\mathrm{T}}(T) = S_{\mathrm{T}}^{\infty} \left[1 - \exp\left(\frac{T^* - T}{T_0} \right) \right], \qquad (3)$$

with fitting parameters S_{T}^{∞}, T^* and T_0. In principle T_0 is a measure of the temperature dependence, if S_{T}^{∞} and T^* are fixed. Note that only under these circumstances T_0-values can be compared. At very high temperature the enthalpic *chemical contributions* are negligible and the entropic contribution dominates the behavior, so that we expect that S_{T}^{∞} is determined by differences in physical parameters *e.g.* moment of inertia and mass.

Lately, it has been shown, that this empirical correlation breaks down, if the solute concentration becomes too high [6]. For diluted aqueous solutions the empirical
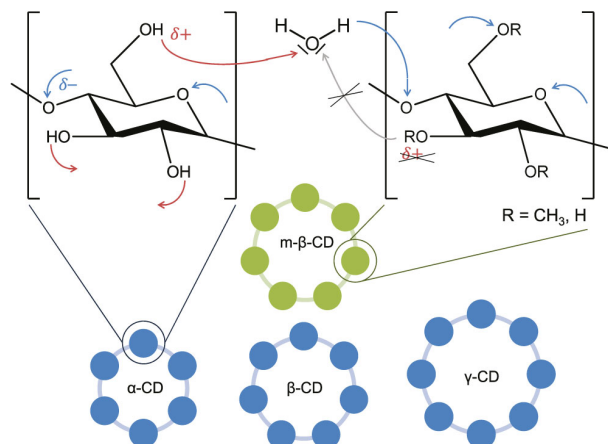
behavior might be understood by a free energy concept as suggested by Wang *et al.* [12]. The basic idea is that at low temperatures the system forms hydrogen bonds to minimize its free energy, so that the water molecules enrich at the cold side, while at higher temperature the entropy is maximized by breaking up the hydrogen bond networks, which leads to an enrichment of water on the warm side. As a consequence the Soret coefficient S_{T} of the solute molecules increases when hydrogen bonds break [13]. Another possibility is to add an ingredient with a strong affinity to water, so that the bonds open [14]. The strength of a hydrogen bond network in a mixture is influenced by the donor and acceptor sites. Recently, it has been shown that S_{T} depends linearly on the difference of donor and acceptor sites of the solute molecule belonging to a homologous series [15]. Hydrogen bonding certainly plays an important role in aqueous CD solutions.

Cyclodextrins can form hydrogen bonds (HB) with water via their hydroxyl groups which act as HB donor sites or as acceptor via the oxygens (see fig. 2, left). The donor and acceptor sites are indicated by red and blue arrows, respectively. There are different ways of counting donor and acceptor sites, we follow in our analysis the work by Maeda *et al.* [15]. So that we count the hydroxyl groups only as a donor cite and ignore that the oxygen atom might serve as an acceptor. The free electron pairs of the ether oxygens can act as HB acceptors, but due to the low polarity the bonding is weak in the order of 20 kJ/mol (corresponding to 5 kcal/mol) [16,17]. We count the oxygen ether only as one acceptor, because due to steric hindrance it is unlikely that it can bind two water molecules. In the used methyl-β-cyclodextrin 55% of the hydroxyl groups are randomly methylated and hence unable to form hydrogen bonds with the solvent (fig. 2, right). The differ-

Table 1. Properties of α-, β-, γ and methyl-β-cyclodextrin. M_W is the molecular mass. The octanol/water partition coefficient $\log P$ given in this table is the average of two incremental calculation methods, Consensus and ChemAxon. Details about the calculation of $N_{don} - N_{acc}$ can be found in the text.

Cyclodextrin	number of glucose units	M_W (g mol^{-1})	$\log P$	$N_{don} - N_{acc}$
α	6	973	−11.65	6
β	7	1135	−13.61	7
γ	8	1297	−15.82	8
methyl-β	7	1303	−6.28	−11.9

ence $N_{don} - N_{acc}$ is listed in table 1. Note that a different way of counting donor and acceptor sites will not change the sequence of the compounds.

In this work we studied the Ludwig-Soret effect of α-, β-, γ-, and methyl-β-cyclodextrin in water and in formamide by infrared thermal diffusion forced Rayleigh scattering (IR-TDFRS). The un-methylated CDs α-, β-, and γ-cyclodextrin are chemically very similar and differ only in the number of their glucopyranose units and accordingly in their molecular weight and partition coefficient $\log P$, which is a measure for the hydrophilicity (see table 1). In methyl-β-CD 55% of the hydroxyl-H are randomly substituted with methyl, which results in a similar molecular weight to γ-CD, but drastically changes the interaction with the surrounding solvent.

2 Experimental section

2.1 Thermal diffusion forced Rayleigh scattering

We used infra-red thermal diffusion forced Rayleigh scattering (IR-TDFRS), a laser-induced transient grating technique [18]. Two laser beams create a holographic grating within the sample. The inherent absorption of water in that range converts the laser light into a temperature grating [19]. Due to thermodiffusion eventually a superimposed concentration grating is created. Both the temperature and the concentration grating result in a refractive index grating that is read out by Bragg diffraction of a third laser beam. The IR-TDFRS has been especially developed for aqueous systems, because the water soluble dyes often change the behavior as a function of pH, so that it is not possible to use a suitable dye as it is done in the classical TDFRS, which has been developed to its present status in the group of Köhler [20]. Also the investigation of surfactant systems is altered by the addition of the dye, which can act as co-surfactant and then leads to changes of the phase behavior [21,22]. A detailed description of the used IR-TDFRS can be found in the paper by Blanco *et al.* [23].

The total heterodyne scattering intensity $\zeta_{het}(t)$, assuming an ideal excitation with a step function, is given by

$$\zeta_{het}(t) = 1 - \exp\left(-\frac{t}{\tau_{th}}\right) - A\left(\tau - \tau_{th}\right)^{-1}$$
$$\times \left\{\tau\left[1 - \exp\left(-\frac{t}{\tau}\right)\right] - \tau_{th}\left[1 - \exp\left(-\frac{t}{\tau_{th}}\right)\right]\right\}, \quad (4)$$

with the steady state amplitude A

$$A = \left(\frac{\partial n}{\partial c}\right)_{p,T}\left(\frac{\partial n}{\partial T}\right)_{p,c}^{-1} S_T c(1-c), \quad (5)$$

where c is the mass concentration of the solute, τ_{th} the heat diffusion time, $(\partial n/\partial c)_{p,T}$ and $(\partial n/\partial T)_{p,c}$ are refractive index contrast factors with respect to mass concentration at constant pressure and temperature, and referring to temperature at constant pressure and mass concentration, respectively. The Soret coefficient, $S_T = D_T/D$, can be expressed as ratio of the thermal diffusion coefficient, D_T, and the collective diffusion coefficient, D. Whereas the diffusion coefficient $D = 1/(q^2\tau)$ can be calculated from the diffusion time, τ (cf. eq. (4)), using the magnitude of the grating vector q, which is given by $q = (4\pi/\lambda_w)\cdot\sin(\theta/2)$. Here θ is the angle between the two writing beams and λ_w is the wavelength of the laser beam. The transport coefficients are determined by fitting eq. (4) to the measured heterodyne signal and deconvoluting the excitation function [24,25].

2.2 Sample preparation

Investigated substance were α-cyclodextrin (α-CD, Tokyo chemical industry, $> 98.0\%$), β-cyclodextrin (β-CD, Tokyo chemical industry, 99.0%), γ-cyclodextrin (γ-CD, Tokyo chemical industry, $> 98.0\%$), and methyl-β-cyclodextrin (methyl-β-CD, Tokyo chemical industry). All of them were dissolved in distilled and deionized water (Millpore) and formamide (Sigma-Aldrich, ≥ 99.5) with a concentration of (1.00 ± 0.01) wt%. To ensure homogeneity of the mixture all CD solutions have been stirred for one hour at room temperature. Approximately 2 mL of the prepared solutions were filtered through a $0.2\,\mu$m filter (Whatman Anotop 10) before filling them into an optical quartz cell (Hellma) with an optical path length of 0.2 mm. At least two measurements with different cells and freshly prepared samples were done for each system.

2.3 Contrast factor measurements

The refractive index increments with the mass concentration $(\partial n/\partial c)_{p,T}$ was measured by an Anton Paar RXA 156 refractometer, with an accuracy of 0.00002 nD and a temperature control of $\Delta T = \pm 0.03$ K. The refractometer uses the sodium line with a wavelength of 589.3 nm, which is roughly 40 nm shorter than the HeNe-laser of 632.8 nm used as read-out beam in the IR-TDFRS. This causes a
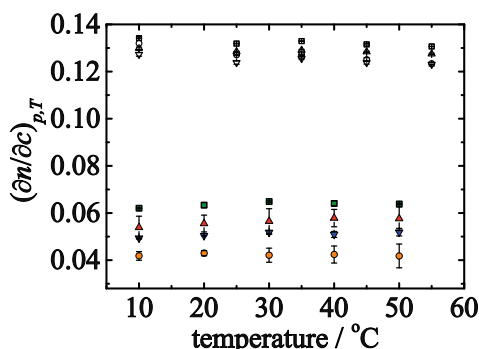
Fig. 3. Contrast factor $(\partial n/\partial c)_{T,p}$ as a function of temperature in water and in formamide. Open and solid symbols refer to the aqueous and formamide solutions, respectively (α-CD (red triangle up), β-CD (blue triangle down), γ-CD (green square) and methyl-β-CD (orange circle)).

small systematic error in the refractive index increment in the order of 0.5–1% [26,27]. For each CD in water and formamide, the refractive index has been measured for at least six concentrations around 1 wt%. The refractive index increments with temperature $(\partial n/\partial T)_{p,c}$ was measured interferometrically [28]. Measurements were either performed around the desired temperature or continuously in the entire temperature range between 10–60 °C. We measured the contrast factors by heating and cooling the solutions with a typical rate of 1 mK/sec.

3 Results

3.1 Contrast factors

In fig. 3 $(\partial n/\partial c)_{p,T}$ is shown, which corresponds to the slope of the linear interpolated refractive index as a function of concentration. For both solvents the refractive index contrast factors with concentration show only a weak temperature dependence. For evaluation of the IR-TDFRS measurements the measured $(\partial n/\partial c)_{p,T}$ and $(\partial n/\partial T)_{p,c}$ values at the respective temperatures were taken or the value was interpolated from the neighboring values. The error bars represent the standard deviation of multiple measurements. The optical contrast for the aqueous systems is two to three times higher than for the formamide solution. The error bars correspond to less than 0.5% and up to 17% for water and formamide solutions, respectively.

For the aqueous CD solutions $(\partial n/\partial T)_{p,c}$ shows a strong linear decrease with temperature, while the variation of the refractive index with temperature in formamide is much weaker (see fig. 4). In the investigated temperature range $(\partial n/\partial T)_{p,c}$ of the aqueous solution changes by a factor four and we find identical values for cooling and heating the sample. In contrast to the aqueous solution we observe an unusual behavior of $(\partial n/\partial T)_{p,c}$ for freshly prepared CD/formamide solutions (see inset of fig. 4). For all four solutions we find a minimum between 16 and 20 °C shifting to higher temperatures with increasing number of glucopyranose units. The absolute variation between the
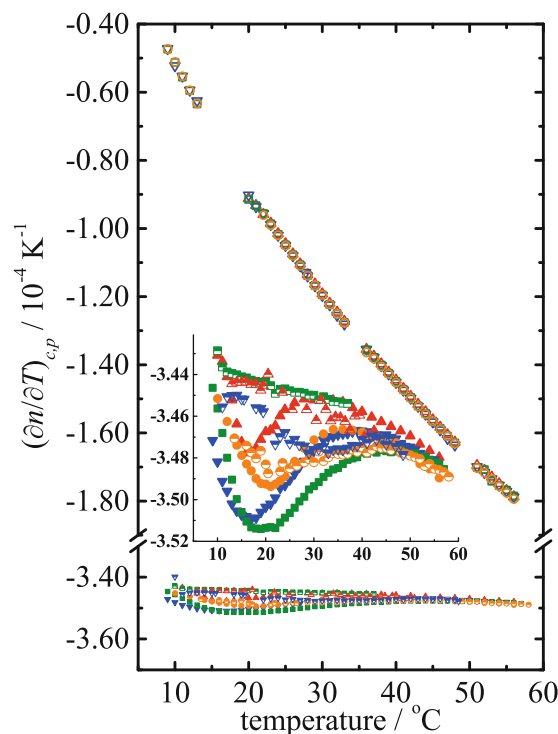

Fig. 4. Contrast factor $(\partial n/\partial T)_{c,p}$ as a function of temperature in water and in formamide. We used the same symbols as in fig. 3. The inset magnifies the measurement results in formamide. The additional half-filled symbols mark the corresponding systems, which have been preheated overnight at 50 °C.

local minimum and maximum of $(\partial n/\partial T)_{c,p}$ lies between $2.7 \cdot 10^{-6}$ K^{-1} and $4.2 \cdot 10^{-6}$ K^{-1} and the relative change of $(\partial n/\partial T)_{c,p}$ is between 0.7 and 2.3%. It turns out that the minimum becomes weaker when the sample ages or is preheated overnight. There are differences between the different CDs. In the case of the γ-CD the minimum disappears completely with time or temperature, while for both β-CDs a shallow minimum remains. The very shallow noisy minima can be observed in β-CD (half-filled blue triangle down) and methyl-β-CD (half-filled orange circle around 30°C and 20°C, respectively.

According to these temperature- and time-dependent changes in the solutions, measurement results for the fresh solutions differ between cooling and heating cycles.

3.2 IR-TDFRS experiments

The thermal diffusion of (1.00 ± 0.01) wt% α-, β-, γ-, and methyl-β-cyclodextrin in water and in formamide was measured in a temperature range from 10 to 55 °C. The results are shown in fig. 5. In water, the Soret coefficients of α-, β-, and γ-cyclodextrin show a temperature dependence as described by eq. (3). There is a sign change around 35 °C and all systems reach the same S_{T}^{∞} value at high temperatures. In comparison, methyl-β-cyclodextrin is much more thermophobic with a predicted sign change
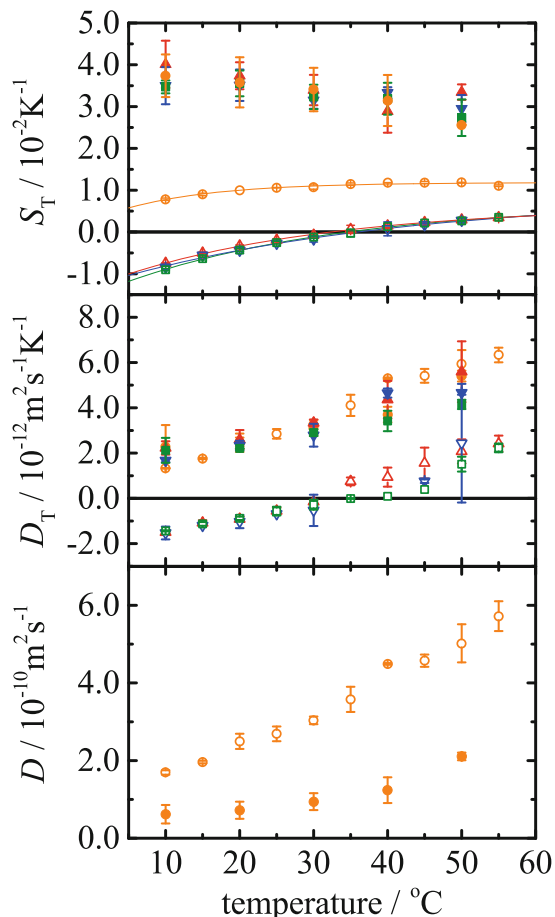
Fig. 5. The Soret coefficient S_T and the thermal diffusion coefficient D_T of α- (red triangle up), β- (blue triangle down), γ- (green square), and methyl-β- (orange circle) cyclodextrin in water and formamide as a function of temperature. For better readability the bottom plot shows only a comparison of the diffusion coefficient of D for methyl-β-cyclodextrin in water and formamide as a function of temperature. Open and solid symbols refer to the aqueous and formamide solution.

Table 2. Fitting parameters and their uncertainties according to eq. (3) shown as solid lines in fig. 5.

Cyclodextrin	S_T^∞ $(10^{-3}\,\mathrm{K}^{-1})$	T^* (°C)	T_0 (K)
α	6.50 ± 0.41	32.4 ± 0.3	29.7 ± 1.7
β	6.42 ± 1.83	35.5 ± 0.9	30.0 ± 6.7
γ	6.95 ± 0.59	34.6 ± 0.3	29.8 ± 2.2
methyl-β	11.8 ± 0.3	-3.3 ± 3.2	12.5 ± 2.8

outside the measured range and a much higher S_T^∞. The fitting parameters are summarized in table 2. In formamide, the behaviour does not follow eq. (3), but shows a decline of the Soret coefficient with increasing temperature. The error bars are much larger than for aqueous solutions. The uncertainties are typically 8% in water and 13% in formamide. The larger uncertainties for formamide

solutions might be understood by the lower optical contrast with concentration and the weaker absorption of the laser light by formamide. Additionally we have a variation of $(\partial n/\partial T)_{p,c}$ in the order of 2.3%.

4 Discussion

4.1 Contrast factor $(\partial n/\partial T)_{p,c}$

While the aqueous solutions show the typical decay of $(\partial n/\partial T)_{p,c}$ with temperature, which is related to the lower density at high temperature, the temperature dependence of $(\partial n/\partial T)_{p,c}$ in formamide solutions is unusual and has to our best knowledge never been observed. Note that the effect is small with a relative amplitude in the order of 2.3%. The inset in fig. 4 shows a decrease of $(\partial n/\partial T)_{p,c}$ in the temperature range between 16 and 20 °C, which would suggest the formation of cavities leading to a more pronounced temperature decrease than caused by the temperature raise. From X-ray diffraction studies it is known that CDs form inclusion complexes with small molecules such as dimethylformamide [29] and formic acid [30], but to our best knowledge the kinetics of such inclusions has not been investigated. From the size of the molecules one would expect very fast kinetics, but our measurements suggest that the process takes several days if the solution is stored at room temperature or it can be accelerated by heating the sample. An explanation might be another, slower, complex formation that blocks the CD cavities from the formamide. A stacking of two or more CDs at temperatures around 20 °C might lead to blocked empty CD cavities that can only be filled up when the stacks break at higher temperatures or over time. To clarify the involved processes further quantitative studies are necessary, which are beyond the scope of this work. Due to the small relative error (2.3%), which is smaller than the uncertainty of the steady state amplitude A (ca. 14%), it will not influence the evaluation of our IR-TDFRS results. Nevertheless it would be interesting to understand the mechanism, which leads to this unusual behavior of $(\partial n/\partial T)_{p,c}$.

4.2 IR-TDFRS measurements

The upper graph of fig. 5 shows the Soret coefficient of the CDs in water and in formamide. While all aqueous solutions show the typical temperature dependence for systems dominated by hydrogen bonds (see eq. (3)), this is not the case for solutions in formamide.

For aqueous systems the Soret coefficient of methyl-β-CD deviates strongly from that of the un-methylated CDs. This change in behaviour is caused by reduced hydrogen bond interactions between the methyl-β-cyclodextrin and the water due to the partial methylation, which reduces the number of potential hydrogen bonds more than a factor two. The observed increase of S_T is comparable to observations made by Sugaya *et al.* where the addition of urea to an aqueous solution led to a similar increase of the

measured S_T [14]. Also Kishikawa *et al.* observed a similar behavior of S_T for pullulan in water and DMSO[13]. In that experiment the addition of urea (up to 5 M, which corresponds to a 28 wt% urea solution) leads to a replacement of water as solvent and the observed shift in S_T due to the weaker solute solvent interaction of urea.

In the case of cyclodextrins in formamide we also measure higher Soret coefficients than in water and the temperature dependent S_T of methyl-β-CD in formamide is equal to those of the other CDs in formamide. The strength of interactions is already weak in comparison to that with water, so that the methylation does not lead to such a pronounced change. Despite the fact that the number of hydrogen bonds formed by formamide is comparable to water [31], the water hydrogen bonds are expected to be stronger due to the larger polarity of water compared to formamide.

Additionally to the already mentioned uncertainties we also observe instabilities of the measured signal of IR-TDFRS at long times (0.5 sec). These instabilities in the intensity of the refracted beam points towards fluctuations in the concentration grating that overlays the laser-induced temperature grating, but the cause is unclear, but might be related to the minimum observed for $(\partial n/\partial T)_{p,c}$. One explanation might be the formation of complexes mentioned before that leads to this atypical behaviour. The scattering of the measurements in formamide prevents an analysis of subtle changes in the behaviour between different cyclodextrins.

The center graph in fig. 5 shows the temperature dependence of the thermal diffusion coefficient D_T. As in the case of S_T we observe an increase of D_T by using formamide as solvent instead of water. Surprisingly D_T and its temperature dependence for methyl-β-CD is very similar in water and formamide. The difference of the Soret coefficients of methyl-β-CD evident in the upper graph of fig. 5 is due to the Fickian diffusion (bottom of fig. 5) which is slower in formamide than in water and consistent with the higher viscosity of formamide. Compared to β-CD in water the hydrogen bonds interactions in the two other systems β-CD/formamide and methyl-β-CD/water are weakened. It is expected that the two last mentioned systems show the same trend, but at the present stage it is unclear why this leads to identical D_T-values. Note that D_T increases for methyl-β-CD in water and formamide, although S_T of methyl-β-CD in water increases, while S_T of methyl-β-CD decreases slightly in formamide.

4.3 Solute-solvent interactions in water

It is evident that interactions with the solvent have a strong impact on the thermodiffusion behaviour of the solute, especially in water, which can strongly interact due to hydrogen bonds. Unfortunately, solute-solvent interactions in liquids are difficult to predict on a microscopic level. In the following discussion we will look for correlations between the thermodiffusion behaviour measured for the four cyclodextrins and two parameters describing the solute-solvent interaction: the difference between donor
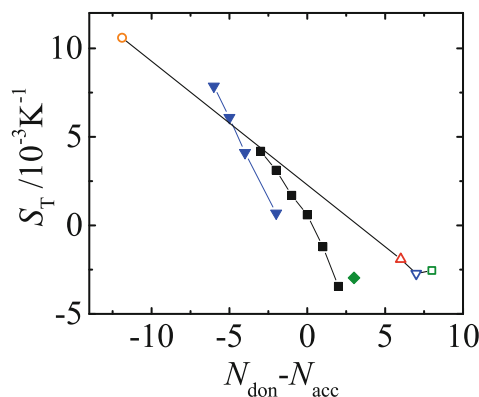

Fig. 6. S_T at $20\,^\circ$C (concentration (1.00 ± 0.01) wt%) *versus* the difference of donor and acceptor sites of the substances. Cyclodextrins (α- (red triangle up), β- (blue triangle down), γ- (green square), and methyl-β- (orange circle) CD) were measured for this work, crown ethers (blue triangles), ethylene glycol oligomers (black squares) and glycerol (green diamond) are reproduced from ref. [15].

and acceptor sites in the solute molecule $N_{\mathrm{don}} - N_{\mathrm{acc}}$ and the octanol/water partition coefficient $\log P$, which are a measure for the hydrophilicity of a compound. Both values are listed in table 1 and will be explained in the following paragraphs.

A direct way to estimate the capability of a substance to interact with water is to count potential donor and acceptor sites in the molecule. A hydroxyl group, for example, would count as a hydrogen bond donor, because a hydrogen bond can be formed by the partially positive hydrogen, an ether oxygen with its free electron pairs and partial negative charge would be counted as hydrogen bond acceptor (see fig. 2). This method does not take into account the polarization strength and hence the relative strengths of the hydrogen bonds. Nevertheless, Maeda *et al.* [15] followed this approach and found for homologous groups a linear correlation between measured Soret coefficient of a solute dissolved in water and the difference between its HB donor and HB acceptor sites. The values for the investigated cyclodextrins fit reasonably well with their findings (see fig. 6). Why the Soret coefficient should dependent on $N_{\mathrm{don}} - N_{\mathrm{acc}}$ is not obvious, because the surrounding water should be equally capable of forming hydrogen bonds with donor as well as acceptor sites of the solute molecule. A possible explanation is the formation of intramolecular hydrogen bonds, so that one donor and one acceptor site block each other against hydrogen bond formation with the surrounding solvent. The difference between donor and acceptor sites would therefore give a better estimate of the potential HB sites open to the solvent than the sum of them. This could be tested by investigation of substances where intramolecular hydrogen bonds are impossible due to sterical hindrance.

The partition coefficient P or, more commonly, its logarithm, $\log P$, are a measure for the relative difference of solubility for a solute in two different solvents. Most commonly used is the octanol/water partition coefficient, be-
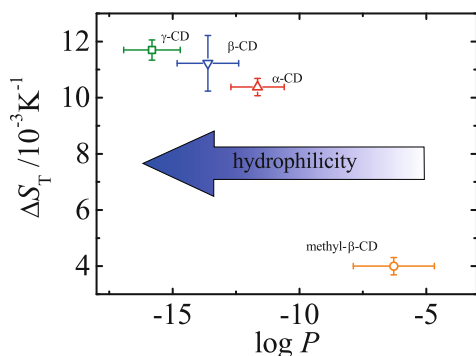
Fig. 7. Correlation of the S_T-difference across the measured range, $\Delta S_T = S_T(50\,^\circ C) - S_T(10\,^\circ C)$, with the partition coefficient $\log P$. Substances with weaker hydrophilicity show smaller S_T-change with temperature.

cause because it is used for modeling physiological and environmental transport processes and an important parameter for drug compounds [32]. In a system where a solute can diffuse freely between two phases, P is the ratio of its equilibrium concentration in octanol over that in water, so a smaller (or negative) $\log P$ signifies stronger hydrophilicity. The hydrophilicity decreases from γ-, β-, α- to methyl-β-CD. In fig. 7 the $\log P$ values of the CDs are plotted *versus* ΔS_T, the difference of the Soret coefficients at 10 and 50 $^\circ$C (cf. fig. 5). We have chosen the S_T-values at these two reference temperatures as a measure for the temperature sensitivity of the systems in the investigated range. ΔS_T increases from methyl-β-, α-, β- to γ-CD in the same way as the hydrophilicity. So the compound with the lowest hydrophilicity shows the weakest temperature dependence.

The isotopic contribution to the Soret coefficient is temperature independent and the cyclodextrins are not charged, the measured temperature dependence of S_T is due to the chemical contribution [33], that is to say the interactions between solute and solvent. In water, these interactions are dominated by hydrogen bonds, which are sensitive to temperature changes. We can assume that at 50 $^\circ$C the hydrogen bonds in the system, those between solute and solvent as well as the HB network of the water, are significantly weakened. This reduces the chemical contribution of S_T leading, in the case of the investigated cyclodextrins, to a rise in the Soret coefficient with rising temperature. So for these substances, as well as all substances for which the temperature dependence of S_T can be described by eq. (3), the chemical contribution to S_T due to the formation of hydrogen bonds is negative. As explained in the introduction this temperature dependence can be understood under the assumption that the formation of hydrogen bonds minimizes the free energy of the system at low temperatures. This should certainly be the case for the formation of hydrogen bonds between solute and solvent. So it seems reasonable that in our investigation the systems which show the highest compatibility with water (lowest $\log P$-value) react most strongly to temperature changes (high ΔS_T).

In conclusion this study confirms that hydrogen bonds play an essential role in the thermophoretic behaviour of aqueous systems. The systematic study of cyclodextrins gives a clear correlation between the change of the Soret coefficient with temperature and the logarithm of the partition coefficient.

References

1. S.R. de Groot, P. Mazur, *Non-equilibrium Thermodynamics* (Dover, New York, 1984).
2. F. Montel, J. Bickert, A. Lagisquet, G. Galliero, J. Pet. Sci. Eng. **58**, 391 (2007).
3. C. Angeli, E. Leonardi, Int. J. Heat Mass Transfer **48**, 4633 (2005).
4. M.E. Schimpf, J.C. Giddings, J. Polym. Sci. Pol. Phys. **27**, 1317 (1989).
5. C.B. Mast, D. Braun, Phys. Rev. Lett. **104**, 188102 (2010).
6. D. Niether, D. Afanasenkau, J.K.G. Dhont, S. Wiegand, Proc. Natl. Acad. Sci. U.S.A. **113**, 4272 (2016).
7. W. Köhler, K.I. Morozov, J. Non-Equilib. Thermodyn. **41**, 151 (2016).
8. E.M. Martin Del Valle, Process Biochem. **39**, 1033 (2004).
9. Sunil S. Jambhekar, Philip Breen, Drug Discov. Today **21**, 356 (2015).
10. K.I. Morozov, W. Köhler, Langmuir **30**, 6571 (2014).
11. S. Iacopini, R. Rusconi, R. Piazza, Eur. Phys. J. E **19**, 59 (2006).
12. Z. Wang, H. Kriegs, S. Wiegand, J. Phys. Chem. B **116**, 7463 (2012).
13. Y. Kishikawa, S. Wiegand, R. Kita, Biomacromolecules **11**, 740 (2010).
14. R. Sugaya, B.A. Wolf, R. Kita, Biomacromolecules **7**, 435 (2006).
15. K. Maeda, N. Shinyashiki, S. Yagihara, S. Wiegand, R. Kita, J. Chem. Phys. **143**, 124504 (2015).
16. P.R. Rablen, J.W. Lockman, W.L. Jorgensen, J. Phys. Chem. A **102**, 3782 (1998).
17. J.P.M. Lommerse, S.L. Price, R. Taylor, J. Comput. Chem. **18**, 757 (1997).
18. Simone Wiegand, Werner Köhler, Thermal Nonequilib. Phenom. Fluid Mixtures **584**, 189 (2002).
19. S. Wiegand, H. Ning, H. Kriegs, J. Phys. Chem. B **111**, 14169 (2007).

20. W. Köhler, P. Rossmanith, J. Phys. Chem. **99**, 5838 (1995).
21. H. Ning, S. Datta, T. Sottmann, S. Wiegand, J. Phys. Chem. B **112**, 10927 (2008).
22. B. Arlt, S. Datta, T. Sottmann, S. Wiegand, J. Phys. Chem. B **114**, 2118 (2010).
23. P. Blanco, H. Kriegs, M.P. Lettinga, P. Holmqvist, S. Wiegand, Biomacromolecules **12**, 1602 (2011).
24. G. Wittko, W. Köhler, Philos. Mag. **83**, 1973 (2003).
25. H. Ning, R. Kita, H. Kriegs, J. Luettmer-Strathmann, S. Wiegand, J. Phys. Chem. B **110**, 10746 (2006).
26. R.D. Camerini-Otero, R.M. Franklin, L.A. Day, Biochemistry **13**, 3763 (1974).
27. V.V. Sechenyh, J. Legros, V. Shevtsova, J. Chem. Thermodyn. **43**, 1700 (2011).
28. A. Becker, W. Köhler, B. Müller, Ber. Bunsen. Phys. Chem. **99**, 600 (1995).
29. K. Harata, Bull. Chem. Soc. Jpn. **52**, 2451 (1979).
30. T. Aree, B. Schulz, G. Reck, J. Inclus. Phenom. Macrocyclic Chem. **47**, 39 (2003).
31. M.D. Elola, B.M. Ladanyi, J. Chem. Phys. **125**, 184506 (2006).
32. C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Adv. Drug Delivery Rev. **64**, 4 (2012).
33. S. Hartmann, W. Köhler, K.I. Morozov, Soft Matter **8**, 1355 (2012).

6 Role of Hydrogen Bonding of Cyclodextrin-Drug Complexes Probed by Thermodiffusion

Role of Hydrogen Bonding of Cyclodextrin−Drug Complexes Probed by Thermodiffusion

Doreen Niether,[†] Tsubasa Kawaguchi,[‡] Jana Hovancová,[§] Kazuya Eguchi,[‡] Jan K. G. Dhont,[†,‖] Rio Kita,*[,‡,⊥] and Simone Wiegand*[,†,#]

[†]ICS-3 Soft Condensed Matter, Forschungszentrum Jülich GmbH, D-52428 Jülich, Germany
[‡]Department of Physics and [⊥]Micro/Nano Technology Center, Tokai University, Hiratsuka, Kanagawa 259-1292, Japan
[§]Chemistry Department, Pavol Jozef Šafárik Univerzity, 041 80 Košice, Slovakia
[‖]Department of Physics, Heinrich-Heine-Universität Düsseldorf, D-40225 Düsseldorf, Germany
[#]Department für Chemie—Physikalische Chemie, Universität zu Köln, D-50939 Cologne, Germany

S *Supporting Information*

ABSTRACT: Temperature gradient-induced migration of biomolecules, known as thermophoresis or thermodiffusion, changes upon ligand binding. In recent years, this effect has been used to determine protein−ligand binding constants. The mechanism through which thermodiffusive properties change when complexes are formed, however, is not understood. An important contribution to thermodiffusive properties originates from the thermal response of hydrogen bonds. Because there is a considerable difference between the degree of solvation of the protein−ligand complex and its isolated components, ligand-binding is accompanied by a significant change in hydration. The aim of the present work is therefore to investigate the role played by hydrogen bonding on the change in thermodiffusive behavior upon ligand-binding. As a model system, we use cyclodextrins (CDs) and acetylsalicylic acid (ASA), where quite a significant change in hydration is expected and where no conformational changes occur when a CD/ASA complex is formed in aqueous solution. Thermophoresis was investigated in the temperature range of 10−50 °C by infrared thermal diffusion forced Rayleigh scattering. Nuclear magnetic resonance measurements were performed at 25 °C to obtain information about the structure of the complexes. All CD/ASA complexes show a stronger affinity toward regions of lower temperature compared to the free CDs. We found that the temperature sensitivity of thermophoresis correlates with the 1-octanol/water partition coefficient. This observation not only establishes the relation between thermodiffusion and degree of hydrogen bonding but also opens the possibility to relate thermodiffusive properties of complexes to their partition coefficient, which cannot be determined otherwise. This concept is especially interesting for protein−ligand complexes where the protein undergoes a conformational change, different from the CD/ASA complexes, giving rise to additional changes in their hydrophilicity.
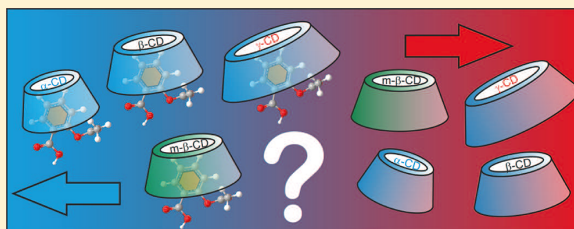
■ INTRODUCTION

A measure to characterize the degree of hydrophilicity of a substance is the 1-octanol/water partition coefficient P (sometimes also denoted as K_{OW} and more commonly given as its logarithm log P). It is used for transport models in several fields, including pharmacological research and environmental science. Although the microscopic meaning of the parameter is not clear, it strongly depends on the formation of a hydration layer, which also influences the transport properties[1] as well as the structure and function[2−4] of water-soluble macromolecules. The degree of hydration of macromolecules, for example, significantly affects their migration induced by spatial gradients in the temperature; this transport mechanism is commonly referred to as thermophoresis or thermodiffusion. Thermophoresis is therefore an effective method to study macromolecular complex formation in cases where the hydration state of the complex is significantly different from that of the non-complexed macromolecule. The change in the degree of hydration of macromolecules and the resulting change in thermodiffusion upon complex formation has been utilized in microscale thermophoresis (MST) to determine the protein−ligand binding constants in dilute solutions as well as the activity of biomolecules.[1] How such binding constants are affected by crowding, such as for protein−ligand complex formation in living cells, is an active area of research.[5−7] Thermophoresis is a promising experimental technique to gain fundamental understanding on the role played by hydrogen bonding with the surrounding water in macromolecular complex formation, both in dilute systems and in crowded environments. It has not yet been employed in this area of

research because a quantitative understanding of how hydrogen bonding affects thermophoretic properties of macromolecules is yet to be developed. A fundamental understanding of the role played by hydrogen bonding in thermophoresis may also be used to tune the thermophoretic properties of drug-delivery complexes to enhance their tendency to migrate toward the warmer regions of inflammation and thereby enhance their effectiveness.

As a first step toward a fundamental understanding of the role played by hydration in complex formation and the utility of thermophoresis as an experimental technique to probe the degree of hydration, we present experiments on a drug-delivery model system consisting of several types of CDs with different hydration properties that bind ASA, also known as aspirin. CDs are ideal model systems for a systematic experimental study to investigate changes in the hydration layer due to complex formation because in contrast to proteins or other biomolecules, CDs are quite rigid so that there are no conformational changes upon complex formation and CDs do not contain polymer-like units that can fold and unfold. Furthermore, in contrast to protein−ligand complexes, the CD/ASA systems have the advantage of being uncharged, are made up of identical units, and are stable in water without the addition of a buffer. The change in thermophoresis properties due to changes in the charge and the influence of the added buffer that would complicate the interpretation of thermophoresis data are therefore absent in the CD/ASA systems.

CDs are cyclic oligosaccharides, which have been developed as drug-delivery systems.[9−12] They have a toroidal shape with a hydrophobic cavity, which serves as a container for guest molecules, enhancing solubility and controlling the speed of uptake of drugs.[10] One of the known guest molecules with an affinity to reside within the cavity of CDs is ASA,[8,13−18] which will be used in the present study. The complex formation by inclusion of this guest molecule in its cavity is mostly enthalpy-driven and is known to be related to hydration effects.[19−21] The complex formation, as sketched in Figure 1, is accompanied by the dehydration of the CD cavity and the guest molecule upon formation of hydrophobic contacts between the guest molecule and hydrophobic sites within the cavity.[22] The hydration of the

complex could reveal whether the remaining contact area with water is more hydrophobic or hydrophilic [sketches (a) and (b) of Figure 1, respectively]. The difference between the two ASA configurations is discussed in more detail in the section on nuclear magnetic resonance (NMR) measurements.

There are several types of CDs with varying degree of hydration. The CDs used in the present study are displayed in Figure 2. The α-, β-, and γ-CDs have 6, 7, and 8 glycopyranose
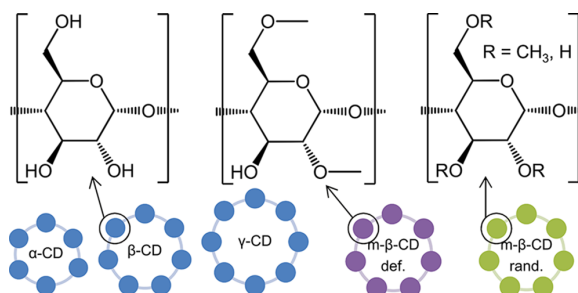
Figure 2. Investigated CDs: α-, β- and γ-CDs are unmethylated (blue); in one methyl-β-CD, 55% of the hydroxyl groups are randomly methylated (green) and in the other, the hydroxyl groups in positions 2 and 6 are methylated (violet).

units, respectively, and are not methylated. The other two CDs are methyl-β-CDs. In one of them, 55% of the hydroxyl groups are methylated randomly (m-β-rand) and in the other, the hydroxyl groups in positions 2 and 6 are methylated resulting in a methylation degree of 67% (m-β-def, where "def" stands for "defined"). We also investigated mixtures of m-β-def and β-CD to vary the degree of methylation. The conformation of ASA within the cavity of these CDs is determined by NMR.

In a binary fluid mixture, thermophoretic mass transport as induced by temperature gradients contributes with $-c(1-c)D_T\vec{\nabla}T$ to the mass flux $\vec{j}$, where D_T is the thermodiffusion coefficient and c is the concentration given as a fraction (mass, mole, or volume fraction, corresponding to the unit connected with the concentration gradient). In this work, concentrations are given in weight fractions, and all experiments are performed in conditions where $c \ll 1$. When $D_T > 0$, mass transport occurs from high to low temperature. In addition, there is a mass flux due to spatial gradients in the concentration c, equal to $-D\vec{\nabla}c$, where D is the Fick's mass diffusion coefficient. The total mass flux is thus given by

$$\vec{j} = -\rho D\,\vec{\nabla}\,c - \rho c(1-c)D_T\,\vec{\nabla}\,T \tag{1}$$

For a time-independent temperature gradient, a steady state is reached when the mass fluxes induced by the Fickian- and thermodiffusion-contributions balance each other. The ratio of the resulting concentration gradient and applied temperature gradient is characterized by the Soret coefficient $S_T = D_T/D$. A larger Soret coefficient implies a larger concentration gradient for a given temperature gradient. For the system of CD and ASA in water, eq 1 cannot describe the flux of every component in the system. It is however not possible to experimentally differentiate between the signals of CD and ASA in solution because of the similar diffusion times of the components. Therefore, we treat the system as quasi-binary mixture, and only one S_T is observed, which contains contributions of ASA, CD, and CD/ASA complex. Knowing the complex fraction and
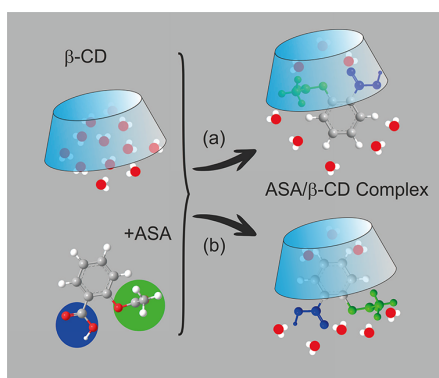

Figure 1. Schematic representation of cyclodextrin (CD) and CD-complexes with acetylsalicylic acid (ASA). Depending on the conformation of ASA in CD, the hydration by the surrounding water changes. Two suggested configurations of the ASA/β-CD complex are shown. These configurations have been obtained by (a) circular dichroism measurements at pH = 2 in H_2O[8] and (b) NMR measurements in D_2O.[9]

the contributions of ASA and CD, S_T of the complex can be calculated (see Supporting Information section 2).

Thermophoresis has been studied theoretically, experimentally, and by computer simulations for many different systems such as low-molecular-weight mixtures, polymer solutions, and colloidal suspensions.[23−40] So far, thermophoresis of low-molecular-weight mixtures is still not understood on a microscopic level. Although for nonpolar systems, the Soret coefficient often shows a linear dependence on physical parameters, for example, mass and moment of inertia of the solute molecules,[37] the understanding of aqueous systems is complicated because of specific interactions. The strong temperature dependence of S_T in aqueous solution can generally be attributed to a chemical contribution[37] caused by hydrogen bond interactions, which decrease with increasing temperature.[41] Many studies have been performed for charged colloids[42−45] and biopolymers;[39] but nonionic aqueous solutions have also been investigated.[46−51] For the latter type of systems, the behavior is dominated by hydrogen bonding, where it has recently been suggested that S_T depends linearly on the difference in the number of donor and acceptor sites of the solute molecule belonging to a homologous series.[30]

Iacopini and Piazza[52] suggested an empirical equation to describe the temperature dependence of the Soret coefficient as

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

with free parameters S_T^∞, T^*, and T_0 that can be adjusted to fit the experimental data. The temperature T_0 is a measure of the sensitivity of the Soret coefficient for changes in the ambient temperature. T^* is the temperature at which the Soret coefficient changes sign. For $T > T^*$, the Soret coefficient is positive, so that migration from high to low temperature occurs, whereas there is a preference for higher temperatures when $T < T^*$. This behavior was explained qualitatively by Wang et al.[53] as minimization of free energy: at low temperatures, the enthalpy contribution dominates, favoring pure water (with a hydrogen bond network undisturbed by the solute) on the colder side; at moderate-to-high temperatures, entropy is increased by maximizing the number of water molecules (small molecules with a high orientational and translational entropy) at the hot side. Later, Vigolo et al.[44] suggested a master curve, which indicates that T^*/T_0 = constant at least for similar systems (see Supporting Information, section 6 for details). It turns out that the empirical working eq 2 describes thermal diffusion of many types of macromolecules in dilute aqueous solutions quite well,[28,29,42] but deviations occur when microstructural heterogeneities are present.[31] There is so far no detailed theory for the contribution of hydrogen bonding to the thermal diffusion coefficient, and therefore a microscopic interpretation of the meaning of the free parameters in eq 2 is in general not available. For charged colloidal particles, thermoelectrophoresis[54,55] and variation of the double-layer energy around the particles[26] give a comprehensive explanation for the temperature dependence of the Soret coefficient.[56] However, at the present stage, it is not possible to derive a microscopic theory of thermophoresis of systems dominated by hydrogen bonds. We will use an approach, which correlates the adjustable parameters with an established empirical parameter describing the hydrophilicity of the system.

■ EXPERIMENTAL SECTION

Sample Preparation. Studied compounds were ASA (Sigma-Aldrich, ≥99.0%), α-CD (Tokyo Chemical Industry, >98.0%), β-CD (Tokyo Chemical Industry, 99.0%), γ-CD (Tokyo Chemical Industry, >98.0%), and two different types of methyl-β-CDs. [one where randomly 55% of the hydroxyl groups were methylated (m-β-rand, Tokyo Chemical Industry) and the other where two out of three hydroxyl groups at defined positions are methylated, heptakis(2,6-di-O-methyl)-β-CD (m-β-def, Sigma-Aldrich, >98.0%)]. Solutions were prepared by using deionized water (Millipore). The concentration of CD was always 1.00 wt %, and they were measured without ASA as well as with equimolar amounts of ASA (see Supporting Information, section 7 for details). The pD/pH value of our solutions without and with ASA is 6 ± 1 and 3.5 ± 0.5, respectively. The hydrolysis half-life of ASA is 644 and 13 h at 10 and 50 °C, respectively.[57] We carefully checked the experimental data by repeating the measurements at low temperatures after the sample had been measured at 50 °C and never noticed a systematic change due to hydrolysis.

Additionally, we adjusted the degree of methylation by mixing β-CD and m-β-def in defined ratios: mix 23% and mix 53%, where 22.5 and 53.4% of the hydroxyl groups are methylated (see Supporting Information, section 7 for details). All solutions were stirred for 40 min at room temperature and then filtered through a membrane filter (0.8/0.2 μm, PALL Acrodisc PF) before the measurement.

NMR. We used NMR spectroscopy to gain structural information about the CD/ASA complex. ^{1}H NMR measurements were performed with Bruker AV500 and AV400WB spectrometers. The samples were prepared as described in the Sample Preparation section; but instead of water, deuterium oxide (D$_2$O, Wako Chemicals, 99.0%) was used as the solvent to diminish the water OH peak. An internal standard such as trimethylsilyl propionate (TSP) was not used to avoid the formation of an inclusion complex with CD. Instead of TSP, the peak of semiheavy water (HDO) was standardized as 4.7 ppm.[58]

The NMR diffusion measurements were performed by a Bruker Diff50 probe with a pulsed-field gradient stimulated echo (PFG-STE) sequence.[59] The attenuation curve was analyzed by the Stejskal−Tanner equation,[60] and the diffusion coefficient D was obtained for all peaks with

$$\frac{I(g)}{I_0} = \exp\left(-(\gamma g \delta)^2 \left(\Delta - \frac{\delta}{3} \right) D \right) \tag{3}$$

where γ is the proton gyromagnetic ratio, g is the field gradient pulse intensity, δ is the gradient pulse duration, Δ is the diffusion time, $I(g)$ is the peak intensity at g, and I_0 is the intensity at the minimum field gradient g of the experiment. To analyze the steric structure of the β-CD/ASA inclusion complex, the phase-sensitive ROESY technique was used.[61] The spinlock mixing was set to 200 ms. Measurement was performed by changing the g value up to 600 G/cm. Δ and δ were fixed at 5 and 1 ms, respectively. All NMR measurements were performed at 25.0 ± 1.0 °C.

TDFRS. The thermal diffusion coefficients were measured in an optical quartz cell (Hellma) with an optical path length of 0.2 mm by infrared thermal diffusion forced Rayleigh scattering (IR-TDFRS), a laser-induced transient grating technique, which has been described in detail before.[62,63] IR-TDFRS was measured in the temperature range of 10−50 °C, in steps of 5 °C. At least two measurements for each sample concentration were done. The error bars represent the standard deviation of the mean. All measurements have been performed in pure water. Because ASA changes the pH of the solution, we performed additional experiments in buffered solution, which are discussed in detail in the Supporting Information (section 4). We found that the thermal diffusion coefficient is not influenced by pH, and only the translation diffusion coefficient shows a small effect because of the higher ionic strength at lower pH. Therefore, we concluded that effects due to pH change can be neglected, and changes of the thermophoretic behavior are solely because of complex formation.

The refractive index increments with mass concentration $(\partial n/\partial c)_{p,T}$ was measured by an Anton Paar RXA 156 refractometer. The

refractive index was measured for four concentrations (1, 0.75, 0.5, and 0.25 wt %) at five different temperatures (10−50 °C). Because the slope $(\partial n/\partial c)_{p,T}$ of the refractive index against concentration was equal within their errors for all temperatures, we used their mean value for the evaluation of our IR-TDFRS data. The resulting value has an error of up to 10% because of the small refractive index changes at these low concentrations. The precision of the refractometer is 0.00002 nD for the refractive index, and $\Delta T = \pm 0.03$ K for the temperature control. Systematic errors in the refractive index increment because of the shorter wavelength used by the refractometer are in the order of 0.5−1%[64,65] and therefore relatively small.

The refractive index increment with temperature $(\partial n/\partial T)_{p,c}$ was measured interferometrically.[66] It was measured in the temperature ranges 7−15, 27−35, and 47−55 °C and interpolated for the values in between. The solution was heated and cooled automatically with a typical rate of 1 mK/s. The $(\partial n/\partial c)_{p,T}$ and $(\partial n/\partial T)_{p,c}$ values for the studied systems are summarized in the Supporting Information.

■ RESULTS

NMR. Figure 3 shows the ^{1}H NMR spectrum of the β-CD/ASA/D$_2$O mixture at 25 °C. The assignments of the proton
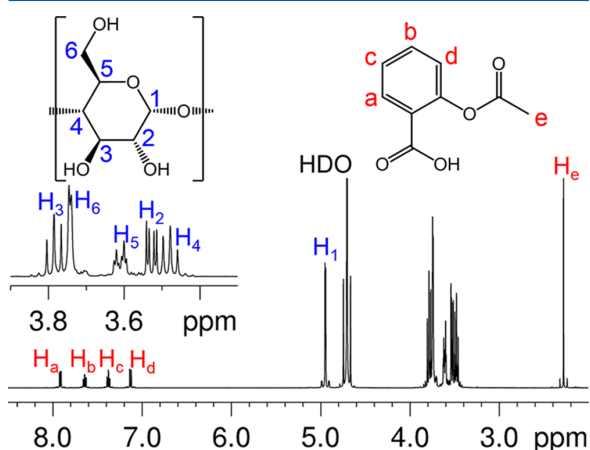

Figure 3. ^{1}H NMR spectrum of an equimolar β-CD/ASA mixture in D$_2$O. H$_a$ to H$_e$ belong to ASA; H$_1$ to H$_6$ belong to CD.

peaks of β-CD and ASA have been obtained from previous reports.[58,67] The protons of ASA are assigned with letters H$_a$ to H$_e$, and the protons of β-CD are assigned with numbers H$_1$ to H$_6$. At room temperature, only H$_3$ and H$_5$ of β-CD in the ternary mixture are shifted by 0.0760 and 0.1335 ppm, respectively, compared to the binary mixture.

Complex formation is confirmed by the ROESY spectrum shown in Figure 4a. Figure 4b shows an enlargement of the spectrum with the cross-peaks. Using the assigned names displayed in Figure 3, we observe the coupling for protons H$_3$(β-CD) and H$_5$(β-CD) pointing to the inner of the ring. The strongest coupling is found for H$_5$(β-CD)−H$_a$(ASA) and H$_5$(β-CD)−H$_d$(ASA). Both ASA protons are located in the phenyl ring. An approximately 40% weaker coupling exists for the same protons of ASA with H$_3$(β-CD). The stronger coupling of H$_5$(β-CD) compared to H$_3$(β-CD) might be related to the slightly smaller diameter of β-CD at that part of the CD. A stronger coupling, which is only 25% weaker than the coupling with H$_5$(β-CD), occurs for H$_3$(β-CD), with H$_e$(ASA) located in the methyl group. This group has some rotational degree of freedom and can probably come closer to the ring.[68]

The NMR diffusion measurements show that the diffusion constants of ASA and β-CD become smaller, but not equal. If the complex was infinitely stable, the diffusion of ASA and β-CD would be slowed down to the same value, that is, the diffusion constant of the heavier complex. In reality, the complex has a limited lifetime, which is shorter than the duration of the NMR measurement of 5 ms. The observed diffusion constant is therefore an average of the diffusion of the individual ASA or β-CD and the slower diffusion of the ASA/β-CD-complex. Similar observations have been made for CDs with surfactants.[69] The self-diffusion coefficients can be used to determine the association constant K_a, which can be expressed for an equimolar mixture as

$$K_a = \frac{p_{com}}{x_0 \cdot (1 - p_{com})^2} \qquad (4)$$

with the initial molar concentration x_0 and the fraction of the formed complexes p_{com} (for further details see Supporting Information, section 1). The fraction p_{com} can be expressed
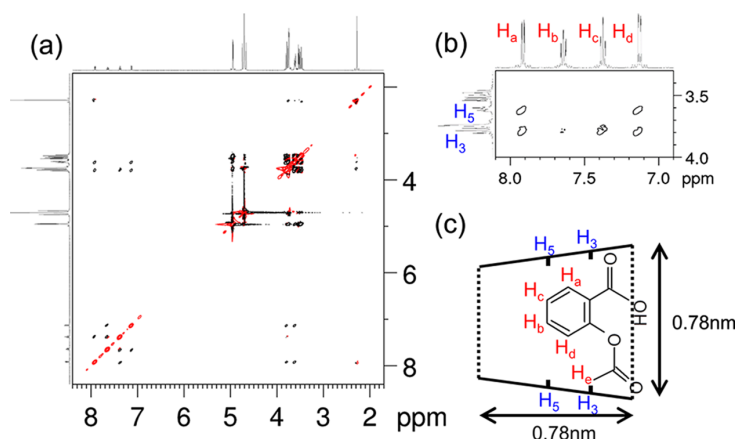

Figure 4. ROESY contour map of an equimolar β-CD/ASA mixture in (a) D$_2$O and (b) its enlarged view, with a mixing time of 200 ms at 25 °C. Black and red lines show positive and negative peaks, respectively. (c) A possible structure of the β-CD/ASA inclusion complex. Ketone group has a rotational degree of freedom; therefore, the double bond faces the outer side of the CD.

with the self-diffusion coefficients D_{ASA} and $D_{\beta\text{-}CD}$ measured in binary solutions of ASA/D$_2$O and β-CD/D$_2$O and the corresponding self-diffusion coefficients $D_{ASA,obs}$ and $G_{\beta\text{-}CD,obs}$ observed in the ternary mixture β-CD/ASA/D$_2$O

$$p_{com} = 1 - \frac{D_{\beta\text{-}CD,mix} - D_{ASA,mix}}{D_{\beta\text{-}CD} - D_{ASA}} \tag{5}$$

Table 1 lists the self-diffusion coefficients found by NMR and compares them with literature results and collective diffusion

Table 1. Diffusion Coefficients of β-CD and ASA in D$_2$O and in the Ternary Mixture β-CD/ASA/D$_2$O, Determined by IR-TDFRS and NMR at 25 ± 0.1 °C

$D/10^{-10}$ m^2 s^{-1}	NMR	TDFRS	lit.
$D_{\beta\text{-}CD}$	2.60 ± 0.03	2.9 ± 0.5	2.91^a
D_{ASA}	6.4 ± 0.3	7 ± 1	8.02^b
			5.0^c
$D_{\beta\text{-}CD,mix}$	2.53 ± 0.05	3.9 ± 0.5^d	
$D_{ASA,mix}$	4.6 ± 0.3		
D_{D_2O}	18.3 ± 0.1		18.72^e
			19.02^f

^{a}Obtained by an optical measurement.[72] ^{b}Obtained with a diffusion cell.[70] ^{c}DOSY measurements.[71] ^{d}Because of the sign change at 25 °C, the concentration signal is very low, so that the value is obtained by interpolating the D values at low and high temperatures. ^{e}Mass extrapolation of a diaphragm cell measurement.[73] ^{f}Tracer diffusion measurement.[73]

coefficients measured by IR-TDFRS. Note that for the investigated dilute solutions, the collective diffusion coefficients agree with the self-diffusion coefficients. Except for D_{ASA}, all values are in excellent agreement with the literature. The diffusion of ASA determined by a diffusion cell is 20% too large,[70] whereas diffusion-ordered spectroscopy (DOSY) experiments lead to a 25% too small value;[71] but the average value of the literature results agrees within the error bars. From our NMR results and the solute concentrations (given in Table S4 in the Supporting Information), we can calculate the association constant using eq 4. The determined complex fraction of $p_{com} = (45 \pm 12)\%$ corresponds to an association constant of $K_a = (174 \pm 135)$ kg·mol^{-1}. It is known that the complex formation depends strongly on pH. Fukahori et al.

measured 10-times higher association of nonionized ASA (pH = 1.7) compared to ionized ASA (pH = 6).[17]

To the best of our knowledge, there are no literature values, which have been measured at pH = 3.5, for a direct comparison of our results. Using isothermal titration calorimetry, Castronuovo and Niccoli[16] found an association constant of $K_a = (210 \pm 30)$ kg·mol^{-1} in a phosphate buffer solution with pH = 9.3, which seems to be rather high if we compare it with $K_a = 51$ kg·mol^{-1} at pH = 6 determined by Fukahori et al. using an ultrasonic relaxation method.[17] Interpolating the values determined by Fukahori et al.,[17] we can estimate $K_a = (360 \pm 70)$ kg·mol^{-1} at pH = 3.5, which just agrees within the error bars.

NMR experiments were performed to assess the structural arrangement of ASA within the CD cage. The position of ASA within the CD and the chemical properties of the outward pointing groups will have an influence on the structure and extent of the hydration layer, leading to a change in the thermophoretic response of the complex compared to the free CD. In the literature, different configurations for ASA in CD have been suggested. Whereas circular dichroism measurements at pH = 2 in H$_2$O suggest that the phenyl ring of the ASA points outward of the CD[8] (compare Figure 1a), NMR measurements in D$_2$O favor a configuration with the phenyl inside the CD[9] (compare Figure 1b). In both studies, the stoichiometric ratio was 1:1, but pH in the NMR study was most likely around 3–4. Probably also the configuration of the complex changes as a function of pH because Fukahori et al.[17] found 10-times higher equilibrium constant for ionized ASA (pH $\approx$ 6) compared to nonionized ASA (pH $\approx$ 2) in solutions with β-CD. The authors speculate that the charged group in the ionized ASA can be easily drawn to solvent bulk water, which is supported by the 10-times larger backward rate constant. On the basis of these results, one would expect that ASA is located deep inside the CD cavity at low pH in contrast to the finding of Dahab and El-Hag.[8] Our NMR study clearly supports the outcome of Loftsson and Duchêne,[9] so that we expect that the hydroxyl group and oxygen point outward from the cavity.

IR-TDFRS. Apart from ASA, which shows a linear decrease of S_T with increasing temperature (Figure 5b, magenta stars), all investigated systems can be described with eq 2.

Figure 5a shows the Soret coefficient S_T of the unmethylated CDs—α, β, and γ. These three CDs differ only in the number
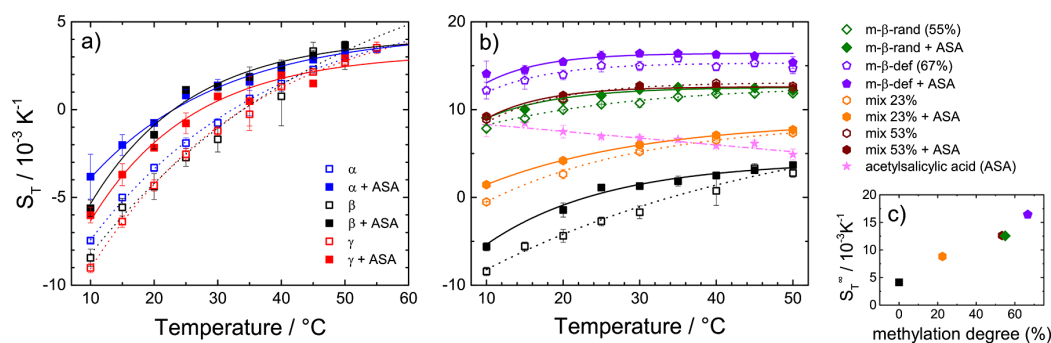

Figure 5. Soret coefficient S_T against temperature for CDs with and without ASA (full and empty symbols, respectively) for (a) unmethylated CDs and (b) methylated CDs, with β-CD (black squares) reproduced for easier comparison. Percentages in the legend give the degree of methylation. Systems called "mix" (orange and dark red hexagons) consist of a mixture of β and m-β-def resulting in the given degree of methylation. Addition of ASA results in more thermophobic behavior (raised S_T). The dotted and solid lines in (a,b) are fits of eq 2 for the CDs and the CD/ASA-complexes, respectively. (c) Soret coefficient at high temperatures S_T^∞ shows a linear dependence on the degree of methylation.

of glycopyranose units and show a very similar behavior. At low temperatures, they are thermophilic ($S_T < 0$); S_T increases with decreasing size of the CD ring. Complex formation with ASA increases the Soret coefficient in all systems by the same amount (0.003 K^{-1} at 10 °C). The inversion temperature $T*$, where the system goes from thermophilic to thermophobic behavior ($S_T = 0$), lies around 35 °C for the CDs alone and is lowered by complex formation with ASA to about 25 °C. At high temperatures, neither the ring size nor the complex formation has a strong impact on the Soret coefficient.

The temperature-dependent Soret coefficient of the methylated systems is shown in Figure 5b. For easier comparison, the results of β-CD are reproduced from Figure 5a. In the investigated systems, 23, 53, 55, and 66% of the hydroxyl groups are methylated, and there is a clear increase of the S_T value with increase in methylation (see Figure 5c). The position of the methyl groups (randomly at any position, defined at two out of three positions or through a mixture of β-CD and m-β-def) does not influence the result. Complex formation with ASA has qualitatively the same result as for the unmethylated CDs, but the increase of S_T is not as pronounced.

Diffusion coefficients for β-CD and ASA were calculated and are listed in Table 1; errors are determined by multiple measurements. In ternary mixture β-CD/ASA/D$_2$O, we observed only one diffusion coefficient of the formed complex β-CD/ASA. It was not possible to differentiate individual diffusion coefficients of β-CD and ASA by IR-TDFRS.

■ DISCUSSION AND CONCLUSIONS

In this work, we use CDs and ASA in water as a model system to study the role played by hydrogen bonding upon complex formation. This is a particularly suitable model system for such a study because the core of CDs does not structurally change when ASA is embedded within the CD ring, so that the major change of the thermodiffusion coefficient upon complex formation is due to changes in the degree of hydrogen bonding. Because the thermophoretic mobility is sensitive to hydrogen bonding, we performed systematic thermophoresis experiments for various types of CDs, noncomplexed and complexed forms, as a function of temperature.

Comparing different CDs, we see a stronger thermophobicity (higher S_T) for higher degrees of methylation (see Figure 5c). For homologous groups, S_T can be predicted with the donor–acceptor concept.[30] It states that for such similar compounds, the Soret coefficient at a fixed temperature depends linearly on the difference in the number of hydrogen bond donors and acceptors in the molecule. Figure 6 shows S_T against $N_{don}-N_{acc}$ for the investigated systems at 20 and 50 °C. Counting the donor and acceptor sites for our CDs, we find that methylation turns a hydroxyl group (hydrogen-bonding donor) into an ether group (hydrogen-bonding acceptor), thus reducing $N_{don}-N_{acc}$ (further information on how the number of donor and acceptor sites is determined is given in the Supporting Information, section 3). In a previous work,[74] we speculated that donor and acceptor sites block each other through intramolecular hydrogen bonds, leading to $N_{don}-N_{acc}$ as the number of hydrogen-bonding sites open to interaction with the solvent. This first idea does not seem likely because with this interpretation, a positive or negative value for $N_{don}-N_{acc}$ would simply denote whether the donor or acceptor sites are left free. Because there is no inherent difference between hydrogen bonds, whether the solute is a donor or acceptor, the sign of $N_{don}-N_{acc}$ should have no influence on S_T. It is clear, however,
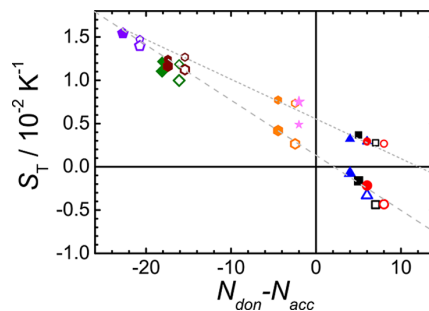

Figure 6. Soret coefficient S_T at 20 °C (larger symbols) and 50 °C (smaller symbols) as a function of the difference between the number of hydrogen bond donors (N_{don}) and the number of hydrogen bond acceptors (N_{acc}). The open symbols refer to the CDs without ASA, and the solid symbols refer to equimolar mixtures of CD and ASA: α-CD (blue triangle), β-CD (black square), γ-CD (red circle), β-CD/m-β-def (23%) (orange hexagon), β-CD/m-β-def (53%) (brown hexagon), m-β-rand (55%) (green diamond), m-β-def (67%) (violet pentagon), and ASA (pink star). The percentages give the degree of methylation.

that hydrogen-bonding acceptors seem to give a positive contribution to S_T, whereas donors give a negative contribution. In the investigated systems, the addition of an acceptor group is always accompanied by addition of hydrophobic groups ($-CH_3$ for CDs and $-CH_2-CH_2-$ for ethylene glycols). A more convincing way to interpret the data is therefore that N_{don} is counting hydrophilic groups, whereas N_{acc} is really counting hydrophobic contributions. Greater hydrophilicity leads to a stronger thermophobic response. This observation holds for 20 °C as well as for 50 °C, although the effect is less pronounced at the higher temperature, indicating the weakening of hydrogen bonds. What is surprising about these findings is that this simple method also predicts the Soret coefficients of the complexes reasonably well, just by adding the donor and acceptor numbers of ASA. This further substantiates our NMR interpretation because the donor and acceptor sites of ASA are accessible to the solvent.

Assuming that the temperature dependence of S_T in aqueous systems is mainly due to the temperature-dependent formation and breaking of hydrogen bonds, we propose a slightly adapted form of eq 2

$$S_T(T) = S_T^\infty - C_H \exp(-A_H T) \tag{6}$$

The interpretation of this empirical formula is as follows. The contribution S_T^∞ to the Soret coefficient stems from the thermal properties of the core material, possible charges, and so forth, without the presence of hydrogen bonds. The second term accounts for the presence of hydrogen bonds, where C_H is a measure of the number of hydrogen bonds. The temperature-dependent factor that multiplies C_H describes the diminishing contribution of hydrogen bonds as they weaken with increasing temperature. The parameter $A_H > 0$ measures the temperature-dependent strength of a hydrogen bond. For larger values of A_H, hydrogen bonds weaken more strongly with increasing temperature. Note that both C_H and A_H have the dimension T^{-1}. The contribution of hydrogen bonds is expected to be the main cause of temperature dependence of the Soret coefficient, so that S_T^∞ is essentially temperature-independent. Fitting the experimental data with eq 6 reveals a significant correlation between parameters C_H and A_H. Figure 7a shows that $\ln(C_H)$ is
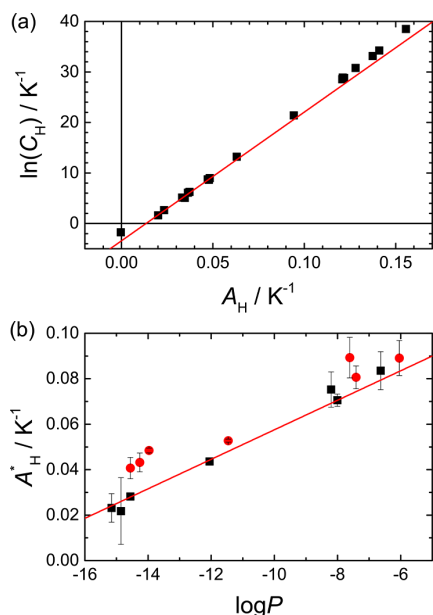
65

Figure 7. (a) Fit parameters C_H and A_H from eq 6 show linear correlation. (b) Inserting the linear dependency of C_H into eq 6, A_H^* becomes the only fit parameter describing temperature dependence in eq 7. It shows a linear correlation with log P. Note that log P values were only available for pure CDs (black squares); the log P values for the ASA complexes (red circles) are estimated (see Supporting Information section 3).

linearly dependent on A_H. Without a more detailed microscopic model, it is impossible to distinguish whether a stronger contribution is due to a stronger interaction or greater number of hydrogen bonds.

Inserting the linear dependence of $\ln(C_H)$ on A_H into eq 6 leads to a function of the form

$$S_T(T) = S_T^\infty - \exp(m + A_H^*(n - T)) \qquad (7)$$

with the constants $m = -3.4 \pm 0.2$ and $n = (255 \pm 5)$ K calculated from the linear fit shown in Figure 7a. The only two fitting parameters left are thus S_T^∞ and A_H^*, where the latter is a measure of the temperature sensitivity of the strength of the hydrogen bonds. Note that m is negative and $n < T$, so that a larger A_H^* indicates a smaller contribution of the temperature-dependent term.

Recently, an empirical correlation between the temperature sensitivity of the Soret coefficient in relation to the partition coefficient log P has been found for different types of molecules.[74] A_H^* is also a measure to characterize the degree of hydrophilicity of a substance. Its values are plotted against log P in Figure 7b. Note that log P values are only available for pure CDs (black squares), whereas log P for complexes (red circles) are not strictly defined and are estimated from the correlation between the difference of donor and acceptor sites, $N_{don}-N_{acc}$, and log P (for further explanation see Supporting Information, section 3). The correlation with hydrophilicity shows that our initial assumption, that is, $S_T(T)$ depends on the interaction strength between the solute and the solvent, is correct. For more hydrophilic compounds, indicated by small partition coefficients, the temperature dependence of S_T is more pronounced (smaller A_H^*) than for more hydrophobic compounds.

Piazza's eq 2 as well as our adapted expression 6 describe the temperature dependence of the Soret coefficient for the investigated CDs and complexes well. Additionally, eq 6 is able to describe the temperature dependence of S_T of more hydrophobic compounds such as ASA. The interpretation of the results, however, is difficult. With an increase in hydrophobicity, A_H and $\ln(C_H)$ become negative, which might be interpreted as a repulsive interaction between the solute and the solvent, such as the formation of clathratelike structures around hydrophobic parts of the molecule. In that case, the homogeneity of the solution on a microscopic scale is questionable, often leading to a decay of S_T as a function of temperature.[31,75,76] The temperature dependence is not due to the strength of hydrogen bonds between the solute and the solvent but due to the increasing flexibility of the hydrogen bond network of water at high temperatures.

Figure 8 shows the temperature sensitivity of the Soret coefficient $\Delta S_T = S_T(50\ °C) - S_T(20\ °C)$ of several linear
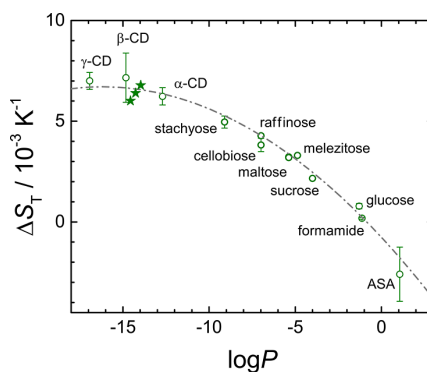

Figure 8. $\Delta S_T = S_T(50\ °C) - S_T(20\ °C)$ against log P for oligosaccharides,[77,78] formamide,[31] CDs, and ASA (open symbols) at low concentrations. Additionally we plotted the values of CD/ASA complexes (green stars) by assuming that S_T observed for CD + ASA is the weighted average of the S_T-values of CD, ASA, and CD/ASA complex (see Supporting Information, section 2 for details). The dashed line is a second-order polynomial fit of all open symbols.

sugars, CDs, and ASA as a function of log P. Taking ΔS_T as a measure of the temperature sensitivity is not principally different from fitting T_0 or A_H^* to the temperature curve, but it necessitates fewer measurements and is therefore easier to obtain. For the complexes, we used the same log P values as for Figure 7. The concentration of solute molecules was quite low with 1 wt % (CDs), 5 wt % (formamide), 10 wt % (oligosaccharides), and 20 wt % (glucose). We find an increase of temperature sensitivity of ΔS_T with increasing hydrophilicity (corresponding to a more negative log P) of the solute molecules. The temperature sensitivity observed for the solutions with CD and ASA ΔS_T^{mix} is much smaller than expected for the complex because of the presence of ASA with its negative ΔS_T^{ASA}. With p_{com} determined for β-CD, the temperature sensitivity of the complex ΔS_T^{com} (green stars in Figure 8) was calculated for the unmethylated CDs (see Supporting Information, section 2). The results agree reasonably well with the trend set by the CDs and the literature compounds (dash dotted line). The methylated CDs without ASA also show a ΔS_T value that is lower than expected, which might be caused by an overestimation of log P by the algorithm used in this work. To test if this interpretation of the

data is correct, further experiments with other drug compounds forming complexes with CDs are required.

Our systematic study of CDs and their complexes with ASA shows that the thermophoretic behavior is strongly connected with hydrophobicity/hydrophilicity. The unambiguous correlation between the temperature sensitivity of the Soret coefficient and log P and the apparent additivity found for the complex compared to the empty CD might be an interesting alternative to define the hydrophilicity of a complex, for which log P cannot be measured independently. Although there are still some questions to be addressed, our investigations show that the change of the temperature sensitivity of S_T upon complex formation could provide structural information about protein−ligand binding. Combining these findings with the rapid MST, it will be possible to quantify not only protein−ligand affinities[1] but also the hydrophilicity of the formed complexes. The temperature sensitivity of thermophoresis also plays a key role to direct drugs into the inflamed areas, and therefore, it would be interesting to perform temperature-dependent studies in crowded media in the future.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.langmuir.7b02313.

> Brief derivation of how the fraction of complexes (see eq 5) can be determined from the measured self-diffusion coefficients, determination of the number of donors and acceptors, assessment of whether the pH-changes due to the addition of ASA has an effect on the results, summary of the measured contrast factors, comparison of our data with the master curve of ref 44, and composition of the investigated solutions (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: rkita@keyaki.cc.u-tokai.ac.jp (R.K.).
*E-mail: s.wiegand@fz-juelich.de (S.W.).

ORCID Ⓘ

Simone Wiegand: 0000-0001-6333-1956

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Jerabek-Willemsen, M.; André, T.; Wanner, R.; Roth, H. M.; Duhr, S.; Baaske, P.; Breitsprecher, D. Microscale Thermophoresis: Interaction Analysis and Beyond. *J. Mol. Struct.* **2014**, *1077*, 101−113.

(2) Stadler, A. M.; Koza, M. M.; Fitter, J. Determination of Conformational Entropy of Fully and Partially Folded Conformations of Holo- and Apomyoglobin. *J. Phys. Chem. B* **2015**, *119*, 72−82.

(3) Smolentsev, N.; Lütgebaucks, C.; Okur, H. I.; de Beer, A. G. F.; Roke, S. Intermolecular Headgroup Interaction and Hydration as Driving Forces for Lipid Transmembrane Asymmetry. *J. Am. Chem. Soc.* **2016**, *138*, 4053−4060.

(4) Liese, S.; Gensler, M.; Krysiak, S.; Schwarzl, R.; Achazi, A.; Paulus, B.; Hugel, T.; Rabe, J. P.; Netz, R. R. Hydration Effects Turn a Highly Stretched Polymer from an Entropic into an Energetic Spring. *ACS Nano* **2017**, *11*, 702−712.

(5) Susumu, K.; Uyeda, H. T.; Medintz, I. L.; Pons, T.; Delehanty, J. B.; Mattoussi, H. Enhancing the Stability and Biological Functionalities of Quantum Dots Via Compact Multifunctional Ligands. *J. Am. Chem. Soc.* **2007**, *129*, 13987−13996.

(6) Mattoussi, H.; Palui, G.; Na, H. B. Luminescent Quantum Dots as Platforms for Probing in Vitro and in Vivo Biological Processes. *Adv. Drug Delivery Rev.* **2012**, *64*, 138−166.

(7) Lohse, M. J.; Nuber, S.; Hoffmann, C. Fluorescence/Bioluminescence Resonance Energy Transfer Techniques to Study G-Protein-Coupled Receptor Activation and Signaling. *Pharmacol. Rev.* **2012**, *64*, 299−336.

(8) Dahab, A. A.; El-Hag, D. Rapid Analysis of NSAIDs Binding to β-Cyclodextrin Using the Simultaneous Measurement of Absorption and Circular Dichroism with a Novel Multi-Cell Low-Volume Device. *Anal. Bioanal. Chem.* **2012**, *404*, 1839−1850.

(9) Loftsson, T.; Duchêne, D. Cyclodextrins and their pharmaceutical applications. *Int. J. Pharm.* **2007**, *329*, 1−11.

(10) Del Valle, E. M. M. Cyclodextrins and their uses: A review. *Process Biochem.* **2004**, *39*, 1033−1046.

(11) Szejtli, J. Cyclodextrins—Properties and Applications. *Drug Invest.* **1990**, *2*, 11−21.

(12) Crini, G. Review: A History of Cyclodextrins. *Chem. Rev.* **2014**, *114*, 10940−10975.

(13) Zhou, H.-Y.; Jiang, L.-J.; Zhang, Y.-P.; Li, J.-B. β-Cyclodextrin Inclusion Complex: Preparation, Characterization, and Its Aspirin Release in Vitro. *Front. Mater. Sci.* **2012**, *6*, 259−267.

(14) Shende, P. K.; Trotta, F.; Gaud, R. S.; Deshmukh, K.; Cavalli, R.; Biasizzo, M. Influence of Different Techniques on Formulation and Comparative Characterization of Inclusion Complexes of ASA with β-Cyclodextrin and Inclusion Complexes of ASA with PMDA Cross-Linked β-Cyclodextrin Nanosponges. *J. Inclusion Phenom. Macrocyclic Chem.* **2012**, *74*, 447−454.

(15) El-Gendy, G. A.; Katsuhide, T.; Keiji, Y.; Yoshinobu, N. Molecular Behavior, Dissolution Characteristics and Chemical Stability of Aspirin in the Ground Mixture and in the Inclusion Complex with di-O-Methyl-β-Cyclodextrin. *Int. J. Pharm.* **1986**, *31*, 25−31.

(16) Castronuovo, G.; Niccoli, M. Thermodynamics of Inclusion Complexes of Natural and Modified Cyclodextrins with Acetylsalicylic Acid and Ibuprofen in Aqueous Solution at 298 K. *Thermochim. Acta* **2013**, *557*, 44−49.

(17) Fukahori, T.; Kondo, M.; Nishikawa, S. Dynamic Study of Interaction between β-Cyclodextrin and Aspirin by the Ultrasonic Relaxation Method. *J. Phys. Chem. B* **2006**, *110*, 4487−4491.

(18) Loftsson, T.; Ólafsdóttir, B. J.; Friðriksdóttir, H.; Jónsdóttir, S. Cyclodextrin Complexation of NSAIDss: Physicochemical Characteristics. *Eur. J. Pharm. Sci.* **1993**, *1*, 95−101.

(19) Inoue, Y.; Hakushi, T.; Liu, Y.; Tong, L.; Shen, B.; Jin, D. Thermodynamics of Molecular Recognition by Cyclodextrins. 1. Calorimetric Titration of Inclusion Complexation of Naphthalenesulfonates with α-, β-, and γ-Cyclodextrins: Enthalpy−Entropy Compensation. *J. Am. Chem. Soc.* **1993**, *115*, 475−481.

(20) Inoue, Y.; Liu, Y.; Tong, L. H.; Shen, B. J.; Jin, D. S. Calorimetric titration of inclusion complexation with modified .beta.-cyclodextrins. Enthalpy-entropy compensation in host-guest complexation: from ionophore to cyclodextrin and cyclophane. *J. Am. Chem. Soc.* **1993**, *115*, 10637−10644.

(21) Cooper, A.; Johnson, C. M.; Lakey, J. H.; Nöllmann, M. Heat Does Not Come in Different Colours: Entropy−Enthalpy Compensation, Free Energy Windows, Quantum Confinement, Pressure Perturbation Calorimetry, Solvation and the Multiple Causes of

67

Heat Capacity Effects in Biomolecular Interactions. *Biophys. Chem.* **2001**, *93*, 215−230.

(22) Olvera, A.; Pérez-Casas, S.; Costas, M. Heat Capacity Contributions to the Formation of Inclusion Complexes. *J. Phys. Chem. B* **2007**, *111*, 11497−11505.

(23) Artola, P.-A.; Rousseau, B.; Galliéro, G. A new model for thermal diffusion: Kinetic approach. *J. Am. Chem. Soc.* **2008**, *130*, 10963−10969.

(24) Bringuier, E.; Bourdon, A. Kinetic Theory of Colloid Thermodiffusion. *Phys. A* **2007**, *385*, 9−24.

(25) Braibanti, M.; Vigolo, D.; Piazza, R. Does Thermophoretic Mobility Depend on Particle Size? *Phys. Rev. Lett.* **2008**, *100*, 108303.

(26) Dhont, J. K. G.; Briels, W. J. Single-Particle Thermal Diffusion of Charged Colloids: Double-Layer Theory in a Temperature Gradient. *Eur. Phys. J. E* **2008**, *25*, 61−76.

(27) Hartmann, S.; Wittko, G.; Köhler, W.; Morozov, K. I.; Albers, K.; Sadowski, G. Thermophobicity of Liquids: Heats of Transport in Mixtures as Pure Component Properties. *Phys. Rev. Lett.* **2012**, *109*, 065901.

(28) Iacopini, S.; Rusconi, R.; Piazza, R. The "Macromolecular Tourist": Universal Temperature Dependence of Thermal Diffusion in Aqueous Colloidal Suspensions. *Eur. Phys. J. E* **2006**, *19*, 59−67.

(29) Kishikawa, Y.; Wiegand, S.; Kita, R. Temperature Dependence of Soret Coefficient in Aqueous and Nonaqueous Solutions of Pullulan. *Biomacromolecules* **2010**, *11*, 740−747.

(30) Maeda, K.; Shinyashiki, N.; Yagihara, S.; Wiegand, S.; Kita, R. Ludwig-Soret Effect of Aqueous Solutions of Ethylene Glycol Oligomers, Crown Ethers, and Glycerol: Temperature, Molecular Weight, and Hydrogen Bond Effect. *J. Chem. Phys.* **2015**, *143*, 124504.

(31) Niether, D.; Afanasenkau, D.; Dhont, J. K. G.; Wiegand, S. Accumulation of Formamide in Hydrothermal Pores to Form Prebiotic Nucleobases. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 4272−4277.

(32) Piazza, R.; Iacopini, S.; Triulzi, B. Thermophoresis as a probe of particle−solvent interactions: The case of protein solutions. *Phys. Chem. Chem. Phys.* **2004**, *6*, 1616−1622.

(33) Rauch, J.; Köhler, W. Diffusion and thermal diffusion of semidilute to concentrated solutions of polystyrene in toluene in the vicinity of the glass transition. *Phys. Rev. Lett.* **2002**, *88*, 185901.

(34) Wiegand, S. Thermal diffusion in liquid mixtures and polymer solutions. *J. Phys.: Condens. Matter* **2004**, *16*, R357−R379.

(35) Schimpf, M. E.; Giddings, J. C. Characterization of Thermal Diffusion in Polymer Solutions by Thermal Field-Flow Fractionation: Dependence on Polymer and Solvent Parameters. *J. Polym. Sci., Part B: Polym. Phys.* **1989**, *27*, 1317−1332.

(36) Würger, A. Thermal non-equilibrium transport in colloids. *Rep. Prog. Phys.* **2010**, *73*, 126601.

(37) Köhler, W.; Morozov, K. I. The Soret Effect in Liquid Mixtures—A Review. *J. Non-Equilib. Thermodyn.* **2016**, *41*, 151−197.

(38) Römer, F.; Wang, Z.; Wiegand, S.; Bresme, F. Alkali Halide Solutions under Thermal Gradients: Soret Coefficients and Heat Transfer Mechanisms. *J. Phys. Chem. B* **2013**, *117*, 8209−8222.

(39) Reichl, M.; Herzog, M.; Greiss, F.; Wolff, M.; Braun, D. Understanding the Similarity in Thermophoresis between Single- and Double-Stranded DNA or RNA. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **2015**, *91*, 062709.

(40) Lüsebrink, D.; Yang, M.; Ripoll, M. Thermophoresis of Colloids by Mesoscale Simulations. *J. Phys.: Condens. Matter* **2012**, *24*, 284132.

(41) Dougherty, R. C. Temperature and Pressure Dependence of Hydrogen Bond Strength: A Perturbation Molecular Orbital Approach. *J. Chem. Phys.* **1998**, *109*, 7372−7378.

(42) Wang, Z.; Kriegs, H.; Buitenhuis, J.; Dhont, J. K. G.; Wiegand, S. Thermophoresis of charged colloidal rods. *Soft Matter* **2013**, *9*, 8697−8704.

(43) Ning, H.; Dhont, J. K. G.; Wiegand, S. Thermal-diffusive behavior of a dilute solution of charged colloids. *Langmuir* **2008**, *24*, 2426−2432.

(44) Vigolo, D.; Buzzaccaro, S.; Piazza, R. Thermophoresis and Thermoelectricity in Surfactant Solutions. *Langmuir* **2010**, *26*, 7792−7801.

(45) Fayolle, S.; Bickel, T.; Würger, A. Thermophoresis of Charged Colloidal Particles. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **2008**, *77*, 041404.

(46) Kita, R.; Kircher, G.; Wiegand, S. Thermally induced sign change of Soret coefficient for dilute and semidilute solutions of poly(N-isopropylacrylamide) in ethanol. *J. Chem. Phys.* **2004**, *121*, 9140−9146.

(47) Kita, R.; Wiegand, S.; Luettmer-Strathmann, J. Sign change of the Soret coefficient of poly(ethylene oxide) in water/ethanol mixtures observed by thermal diffusion forced Rayleigh scattering. *J. Chem. Phys.* **2004**, *121*, 3874−3885.

(48) Kita, R.; Wiegand, S. Soret coefficient of poly(N-isopropylacrylamide)/water in the vicinity of coil−globule transition temperature. *Macromolecules* **2005**, *38*, 4554−4556.

(49) Königer, A.; Plack, N.; Köhler, W.; Siebenbürger, M.; Ballauff, M. Thermophoresis of Thermoresponsive Polystyrene−Poly(N-Isopropylacrylamide) Core−Shell Particles. *Soft Matter* **2013**, *9*, 1418−1421.

(50) Wongsuwarn, S.; Vigolo, D.; Cerbino, R.; Howe, A. M.; Vailati, A.; Piazza, R.; Cicuta, P. Giant Thermophoresis of Poly(N-Isopropylacrylamide) Microgel Particles. *Soft Matter* **2012**, *8*, 5857−5863.

(51) Naumann, P.; Datta, S.; Sottmann, T.; Arlt, B.; Frielinghaus, H.; Wiegand, S. Isothermal Behavior of the Soret Effect in Nonionic Microemulsions: Size Variation by Using Different n-Alkanes. *J. Phys. Chem. B* **2014**, *118*, 3451−3460.

(52) Iacopini, S.; Piazza, R. Thermophoresis in protein solutions. *Europhys. Lett.* **2003**, *63*, 247−253.

(53) Wang, Z.; Kriegs, H.; Wiegand, S. Thermal Diffusion of Nucleotides. *J. Phys. Chem. B* **2012**, *116*, 7463−7469.

(54) Würger, A. Transport in Charged Colloids Driven by Thermoelectricity. *Phys. Rev. Lett.* **2008**, *101*, 108302.

(55) Majee, A.; Würger, A. Collective Thermoelectrophoresis of Charged Colloids. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **2011**, *83*, 061403.

(56) Sehnem, A. L.; Neto, A. M. F.; Aquino, R.; Campos, A. F. C.; Tourinho, F. A.; Depeyrot, J. Temperature Dependence of the Soret Coefficient of Ionic Colloids. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **2015**, *92*, 042311.

(57) Edwards, L. J.; Gore, D. N.; Rapson, H. D. C.; Taylor, M. P. The Hydrolysis of Aspirin in Pharmaceutical Preparations. A Limit Test for Free Salicylic Acid. *J. Pharm. Pharmacol.* **1955**, *7*, 892−904.

(58) Li, S.; Purdy, W. C. Circular dichroism, ultraviolet, and proton nuclear magnetic resonance spectroscopic studies of the chiral recognition mechanism of β-cyclodextrin. *Anal. Chem.* **1992**, *64*, 1405−1412.

(59) Cotts, R. M.; Hoch, M. J. R.; Sun, T.; Markert, J. T. Pulsed Field Gradient Stimulated Echo Methods for Improved NMR Diffusion Measurements in Heterogeneous Systems. *J. Magn. Reson.* **1989**, *83*, 252−266.

(60) Stejskal, E. O.; Tanner, J. E. Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. *J. Chem. Phys.* **1965**, *42*, 288−292.

(61) Claridge, T. D. W. *High-Resolution NMR Techniques in Organic Chemistry*, 3rd ed.; Elsevier: Boston, 2016; pp 315−380.

(62) Wiegand, S.; Köhler, W. Measurement of Transport Coefficients by an Optical Grating Technique. *Thermal Nonequilibrium Phenomena in Fluid Mixtures*; Lecture Notes in Physics; Springer, 2002; Vol. *584*, pp 189−210.

(63) Blanco, P.; Kriegs, H.; Lettinga, M. P.; Holmqvist, P.; Wiegand, S. Thermal Diffusion of a Stiff Rod-Like Mutant Y21Mfd-Virus. *Biomacromolecules* **2011**, *12*, 1602−1609.

(64) Camerini-Otero, R. D.; Franklin, R. M.; Day, L. A. Molecular weights, dispersion of refractive index increments, and dimensions from transmittance spectrophotometry. Bacteriophages R17, T7, and PM2, and tobacco mosaic virus. *Biochemistry* **1974**, *13*, 3763−3773.

68

(65) Sechenyh, V. V.; Legros, J.-C.; Shevtsova, V. Experimental and predicted refractive index properties in ternary mixtures of associated liquids. *J. Chem. Thermodyn.* **2011**, *43*, 1700−1707.

(66) Wittko, G.; Köhler, W. Precise determination of the Soret, thermal diffusion and mass diffusion coefficients of binary mixtures of dodecane, isobutylbenzene and 1,2,3,4-tetrahydronaphthalene by a holographic grating technique. *Philos. Mag.* **2003**, *83*, 1973−1987.

(67) *SDBSWeb*. http://sdbs.db.aist.go.jp; National Institute of Advanced Industrial Science and Technology, 2016.

(68) Kitchin, S. J.; Halstead, T. K. Solid-State 2H NMR Studies of Methyl Group Dynamics in Aspirin and Aspirin·*β*-Cyclodextrin. *Appl. Magn. Reson.* **1999**, *17*, 283−300.

(69) Cabaleiro-Lago, C.; Nilsson, M.; Söderman, O. Self-Diffusion NMR Studies of the Host−Guest Interaction between *β*-Cyclodextrin and Alkyltrimethylammonium Bromide Surfactants. *Langmuir* **2005**, *21*, 11637−11644.

(70) Edwards, L. J. The Dissolution and Diffusion of Aspirin in Aqueous Media. *Trans. Faraday Soc.* **1951**, *47*, 1191−1210.

(71) Rogerson, A. K. New Techniques in Diffusion-Ordered NMR Spectroscopy. Ph.D. Thesis, University of Manchester, Faculty of Engineering and Physical Sciences, 2013.

(72) Pavlov, G. M.; Korneeva, E. V.; Smolina, N. A.; Schubert, U. S. Hydrodynamic Properties of Cyclodextrin Molecules in Dilute Solutions. *Eur. Biophys. J.* **2010**, *39*, 371−379.

(73) Weingärtner, H. Diffusion in Liquid Mixtures of Light and Heavy Water. *Ber. Bunsenges. Phys. Chem.* **1984**, *88*, 47−50.

(74) Eguchi, K.; Niether, D.; Wiegand, S.; Kita, R. Thermophoresis of cyclic oligosaccharides in polar solvents. *Eur. Phys. J. E* **2016**, *39*, 86.

(75) Asenbaum, A.; Pruner, C.; Wilhelm, E.; Mijakovic, M.; Zoranic, L.; Sokolic, F.; Kezic, B.; Perera, A. Structural Changes in Ethanol−Water Mixtures: Ultrasonics, Brillouin Scattering and Molecular Dynamics Studies. *Vib. Spectrosc.* **2012**, *60*, 102−106.

(76) Perera, A.; Mazighi, R. On the Nature of the Molecular Ordering of Water in Aqueous DMSO Mixtures. *J. Chem. Phys.* **2015**, *143*, 154502.

(77) Blanco, P.; Wiegand, S. Study of the Soret Effect in Monosaccharide Solutions. *J. Phys. Chem. B* **2010**, *114*, 2807−2813.

(78) Blanco, P.; Kriegs, H.; Arlt, B.; Wiegand, S. Thermal Diffusion of Oligosaccharide Solutions: The Role of Chain Length and Structure. *J. Phys. Chem. B* **2010**, *114*, 10740−10747.

69

7 Thermodiffusion as a Probe of Protein Hydration for Streptavidin and the Streptavidin-Biotin Complex

Thermodiffusion as a Probe of Protein Hydration for Streptavidin and the Streptavidin-Biotin Complex

Doreen Niether[1], Mona Sarter[2,6], Bernd König[3], Michaela Zamponi[4], Jörg Fitter[5,6], Andreas Stadler[2,a)] and Simone Wiegand[1,7,b)]

[1]*ICS-3 Soft Condensed Matter, Forschungszentrum Jülich GmbH, D-52428 Jülich, Germany*
[2]*JCNS-1 & ICS-1 Neutron Scattering, Forschungszentrum Jülich GmbH, D-52428 Jülich, Germany*
[3]*ICS-6 Structural Biochemistry, Forschungszentrum Jülich GmbH, D-52428 Jülich, Germany*
[4]*Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ), Forschungszentrum Jülich GmbH, Lichtenbergstrasse 1, 85748 Garching, Germany*
[5]*ICS-5 Molecular Biophysics, Forschungszentrum Jülich GmbH, D-52428 Jülich, Germany*
[6]*I. Physikalisches Institut (IA), AG Biophysik, RWTH Aachen, Sommerfeldstrasse 14, 52074 Aachen, Germany*
[7]*Department für Chemie - Physikalische Chemie, Universität zu Köln, 50939 Cologne, Germany*

a)Corresponding author: a.stadler@fz-juelich.de
b)Corresponding author: s.wiegand@fz-juelich.de

Abstract. Molecular recognition via protein–ligand interactions is of fundamental importance to numerous processes in living organisms. Microscale thermophoresis (MST) uses the sensitivity of the thermophoretic response upon ligand binding to access information on the reaction kinetics. Additionally, thermophoresis is promising as a tool to gain information on the hydration layer, as the temperature dependence of the thermodiffusion behaviour is sensitive to solute-solvent interactions. To quantify the influence of structural fluctuations and conformational motion of the protein on the entropy change of its hydration layer upon ligand binding, we combine quasi-elastic incoherent neutron scattering (QENS) and isothermal titration calorimetry (ITC) data from literature. However, preliminary results show that replacing water with deuterated water leads to changes of the thermophoretic measurements, which are similar to the changes observed upon binding by biotin. In order to gain a better understanding of the hydration layer all measurements need to be performed in heavy water. This will open a route to develop a microscopic understanding of the correlation between the strength and number of hydrogen bonds and the thermophoretic behaviour.

INTRODUCTION

Thermodiffusion or thermophoresis describes the mass transport in a temperature gradient [1]. Nowadays one of the most important applications is the so-called microscale thermophoresis (MST), which monitors protein-ligand binding interactions and is especially used to determined equilibration constants of biochemical reactions [2]. As sketched in Fig.1 the tendency of a protein to accumulate in the cold regions often changes substantially once a protein binds to a small ligand molecule. The complex has typically only a slightly higher molecular mass compared to the free protein, but during the binding process the hydration layer changes. This can for instance be caused by a conformational change of the protein or due to a different hydrophilicity of the bound ligand molecule compared to the protein in the region of binding. The hypothesis is that changes in the hydration layer influence thermophoretic behaviour upon binding. In order to test this hypothesis on a well-known system and gain a better understanding of hydrogen bonding we investigated the thermophoresis of streptavidin (SA) and compare it with the streptavidin-biotin (SA-B) complex.

Thermodiffusion is characterized by the Soret coefficient S_T, which is equal to the ratio of the thermal diffusion coefficient D_T and the diffusion coefficient D. Interpretation and prediction of thermodiffusion behaviour is difficult due to its sensitivity to the properties of solute (mass, size, charge, moment of inertia) and solvent (ionic strength, chemical interactions). A striking difference between aqueous and unpolar solutions is the strong temperature dependence of the Soret coefficient in water. A sign change from thermophilic (negative S_T) to thermophobic (positive
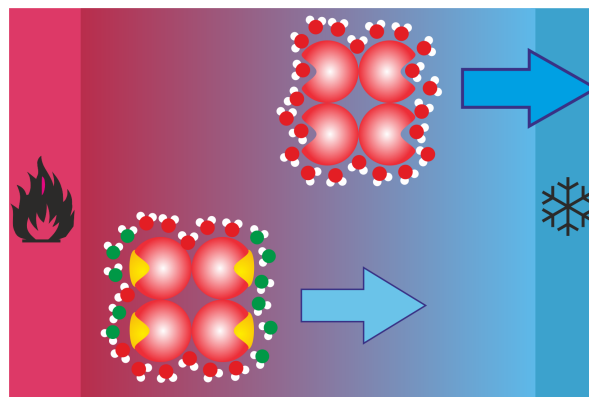
020001-1

FIGURE 1. In this schematic illustration the free streptavidin tetramer has a stronger tendency to accumulate in the cold compared to the streptavidin-biotin complex. It is suspected that the change of the hydration upon binding (indicated by the water molecules with green oxygen atoms) influences the thermophoretic behaviour.

S_T) behaviour at a transition temperature T^* can often be observed [3]. For many biologically relevant systems the temperature dependence of S_T can be described by the following empirical equation

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

where S_T^∞ is the plateau value of S_T that is reached at high temperature [4]. The sensitivity to temperature is due to the contribution of solute-solvent interactions, a contribution that is close to zero in unpolar solvents and decreasing with rising temperature in water due to the breaking of hydrogen bonds at high temperature. While the contribution of these solute-solvent interactions is not strong enough to determine the absolute value of S_T, it dominates the temperature dependent part of the coefficient. The difference of S_T at two temperatures ΔS_T, which is proportional to the chemical contribution to S_T, was shown to correlate with $\log P$ [5].

The partition coefficient $\log P$ is a measure for hydrophilicity/hydrophobicity of a solute and can be used to model the transport of a compound in the environment or to screen for potential pharmaceutical compounds (Lipinski's 'Rule of Five' [6]). It describes the distribution of a solute A between aqueous and oil phase (most commonly 1-octanol) in equilibrium, with P being the ratio of solute concentration in oil and water $P = [A_{oil}]/[A_{water}]$. Due to the fact that measurement of $\log P$ can be quite costly, a number of algorithms have been developed that allow estimation of $\log P$ for any given compound, based on an experimental data base and incremental contributions by functional groups of the molecule. These methods give a reliable $\log P$ value only for room temperature, low concentrations and small (unfolded) molecules that have all contributing groups in contact with the surrounding water.

While the microscopic picture of solvation in water is not yet clear, the tendency of a solute A to accumulate in the oil or aqueous phase can be expressed as a difference in the solvation free energy upon transition between the two phases [7]

$$\Delta G_t(A) = -RT \ln([A_{oil}]/[A_{water}]) = -2.303RT \log P. \tag{2}$$

In an analogous view, the chemical contribution to thermodiffusion, that is the contribution that is due to solute-solvent interactions, can be understood as distribution of the solute between two 'phases' of different temperature, were the hotter water acts more oil-like due to a weakening of the hydrogen bond network.

This gives us only qualitative understanding, however. In order to quantify the entropic change of the hydration layer when biotin binds to SA, we used quasi-elastic incoherent neutron scattering (QENS) to access the conformational entropy of the protein structure, which does not include the entropic contribution of the hydration layer, and compared our results to ligand-displacement isothermal titration calorimetry (LDITC) measurements from the literature [8].

Quasi-elastic incoherent neutron scattering (QENS) observes inelastic scattering of neutrons, where the energy transfer is small compared to the energy of the incident neutrons. Therefore, a broadening of the elastic scattering peak is observed, this gives information about the dynamic properties of the samples [9, 10]. When applied to protein

dynamics QENS mainly observes the incoherent scattering of the hydrogen-atoms, while the deuterium atoms have a significantly smaller incoherent scattering cross section. Therefore, their contribution can be considered negligible when compared to the contribution from the hydrogen-atoms [9, 10]. By exchanging the interchangeable H-atoms in a protein for D-atoms and using a D_2O buffer the main scattering contribution of the sample then comes from the non-exchangeable H-atoms in the protein [11]. QENS observes the dynamics of these H-atoms, the movement of which can be related to a fluctuation of the residues to which they are bound. By measuring QENS from proteins we quantify the conformational fluctuations of the protein under different environmental conditions (i.e. with or without ligands) [12]. The loss of the elastic scattering peaks gives the information about the average amplitude of protein motions, which is given by the mean square displacement (MSD) calculated from the elastic incoherent structure factor EISF A_0 [12]

$$A_0(q) = \exp\left(-\text{MSD}\, q^2\right)(1 - p) + p, \tag{3}$$

where q is the scattering vector and p is the fraction of immobile H-atoms. The EISF is determined from the simplified scattering function

$$S(q, \omega) = A_0(q) \cdot L_G(q, \omega) + [1 - A_0(q)] \cdot L_{G+I}(q, \omega), \tag{4}$$

where $\hbar\omega$ is the energy transfer. The total theoretical scattering function plus linear background $B(q, \omega)$ was convoluted with the instrumental resolution function and fitted to the measured spectra. The HWHM (half width at half maximum) of the two Lorentzians

$$L_G(q, \omega) = \frac{1}{\pi} \times \frac{\Gamma_G(q)}{(\hbar\omega)^2 + \Gamma_G(q)^2} \tag{5}$$

and

$$L_{G+I}(q, \omega) = \frac{1}{\pi} \times \frac{\Gamma_G(q) + \Gamma_I(q)}{(\hbar\omega)^2 + [\Gamma_G(q) + \Gamma_I(q)]^2} \tag{6}$$

account for the global and internal protein diffusion, respectively. The change of the MSD between two states allows for the calculation of the change in the conformational entropy between these states [12].

$$\Delta S_{\text{conformational}} = 3R \ln\left(\sqrt{\left(\frac{\text{MSD}_{\text{complex}}}{\text{MSD}_{\text{free}}}\right)}\right) \tag{7}$$

EXPERIMENTAL SECTION

Sample preparation

Streptavidin (SA) used in these experiments is Streptavidin Streptomyces Avidinii Recombinant produced in *E. coli* (Prospec, 7670308 Rehovol, Israel). The molecular weight of the tetramer is given by the manufacturer as 52 kDa and the amino acid sequence is MAEAGITGTWYNQLGSTFIVTAGADGALTGTYESAVGNAESRYVLTGRYDSAPAT DGSGTALGWTVAWKNNYRNAHSATTWSGQYVGGAEARINTQWLLTSGTTEANAWKSTLVGHDTFTKVKP SAAS. SA was cleaned with PD-10 comlumns, lyophilized and then kept at -20°C. The biotin was purchased as lyophilized powder with ≤99% purity (Sigma-Aldrich, 89555 Steinheim, Germany). The buffer stock solution has the following composition: 250 mM TrisHCl (Tris[hydroxymethyl]aminomethane, ≤99.9%, Sigma-Aldrich, 89555 Steinheim, Germany; Hydrochlorid acid 37%, Merck, 64271 Darmstadt, Germany), 1.2 M NaCl (≤99.5%, Merck, 64271 Darmstadt, Germany), 50 mM KCl (≤99.5%, Merck, 64271 Darmstadt, Germany), 30 mM $MgCl_2$ (≤99.0%, Merck, 64271 Darmstadt, Germany).

For IR-TDFRS experiments the buffer stock solution was diluted with Milipore water 1:9. The concentrations of SA in buffer solution was 50 mg/mL (SA weight fraction 0.048 ± 0.001) for both samples. To the second sample we added biotin at a stoichiometry of SA:biotin = 1:4.

For the neutron scattering experiments the lyophilized SA powder was incubated in D_2O for 24h in order to exchange the interchangeable hydrogen atoms by deuterium. Afterwards the SA was again lyophilized and stored at -20°C. The D_2O based buffer has the following composition: 25 mM TrisDCL, 120 mM NaCl, 5 mM KCl, 3 mM $MgCl_2$, pH=7.4.

Thermal diffusion forced Rayleigh scattering

The thermal diffusion coefficients were measured in an optical quartz cell (Hellma) with optical path length of 0.2 mm by Infra-Red Thermal Diffusion Forced Rayleigh Scattering (IR-TDFRS), a laser-induced transient grating technique, which has been described in detail before [13, 14]. IR-TDFRS was measured in a temperature range from 10 to 50 °C, with steps of 5 °C. At least two measurements for each sample concentration were done. The error bars represent the standard deviation of the mean. All measurements have been performed in the buffer solution. Additional measurements of water/deuterated water mixtures with 10, 25 and 50 wt% of deuterated water were performed at 50°C, the temperature with the strongest signal.

Quasi-elastic neutron scattering

QENS experiments were performed on the backscattering spectrometer SPHERES [15, 16] operated by JCNS at the Heinz-Maier-Leibnitz Zentrum (MLZ) in Garching, Germany. The wavelength of the spectrometer is 6.27 Å, the resolution is approximately $0.65\,\mu$eV for the HWHM and the timescale observed is nanoseconds. The data were analyzed for the q-vectors between 0.5 and 1.6 Å^{-1} and the measurements were performed for 25°C. A sample of SA and of SA-B with a concentration of 65 mg/mL were measured. For the SA-B complex biotin was added to streptavidin in a molar ratio of 4 to 1.

RESULTS

Thermophoretic measurements

IR-TDFRS measurements were conducted for the two samples described above, and for buffer solution and biotin-buffer solution without SA. Although the buffered solutions are multi-component systems, the signal of SA and the SA-B complex can be well separated from those of the buffer salts and free biotin due to the large differences in diffusion speed. The signal of the buffer is very small compared to that of the protein and, as expected due to the strong binding between SA and biotin, we could not detect a separate biotin-signal in the mixture, so that all mixtures could be treated as 2-component systems in the evaluation. The diffusion coefficient of SA was measured as $D(20°C) = (6.7 \pm 0.5) \cdot 10^{-7}$ cm^{2}s^{-1}, which agrees reasonably well with the literature value at room temperature of $D = 6.2 \cdot 10^{-7}$ cm^{2}s^{-1} [17].

The Soret coefficient S_T is shown in Fig. 2 as a function of temperature. Both systems show the typical behaviour of aqueous protein solution with a temperature dependence according to Eq. 1. While the thermodiffusion of the SA-B complex shows no deviation from the free SA at 10°C, at higher temperatures S_T is significantly altered. This is due to
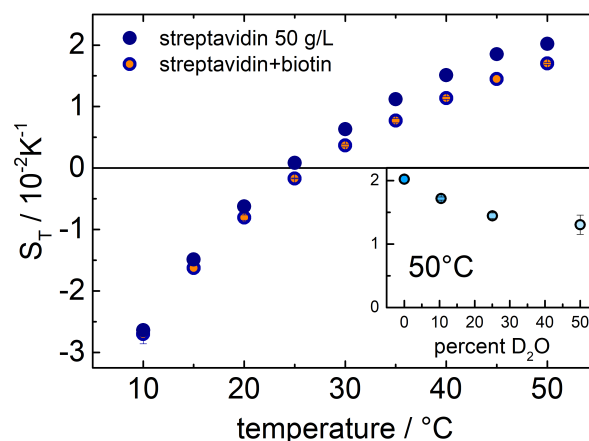

FIGURE 2. Soret coefficient of the unbound SA (blue) and the complex (orange filling) as function of temperature. The inset shows the variation of S_T of the free SA at 50°C with increasing D$_2$O content.
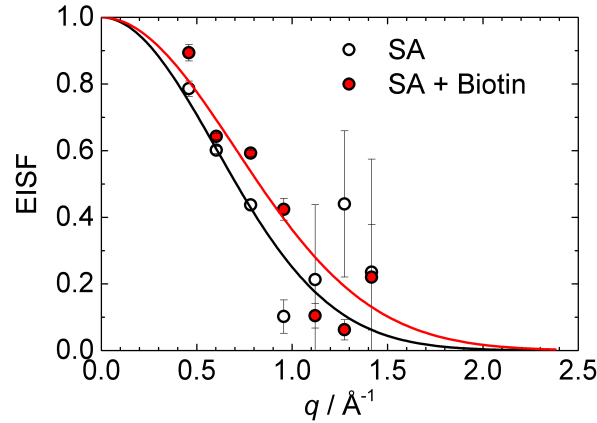
FIGURE 3. Results of QENS measurements performed on SPHERES. Elastic incoherent structure factor of streptavidin (black) and the streptavidin-biotin complex (red) as function of the scattering vector q.

a change in the thermodiffusion coefficient D_T, the diffusion coefficient D is not changed. The temperature sensitivity of S_T is reduced for the complex compared to the free protein. This indicates that the complex is less hydrophilic than the free SA, so that it is likely that the complex forms fewer hydrogen bonds with the surrounding water. Consequently, we expect a higher entropy of the water molecules in the hydration shell.

Due to the absence of an absorption band in D_2O we cannot perform IR-TDFRS measurements in pure heavy water. The inset of Fig. 2 shows that S_T decreases with increasing D_2O content and reaches a plateau at 50wt%. Additional temperature dependent measurements of the free protein and the complex are needed in order to investigate how the temperature sensitivity of S_T is altered.

Neutron scattering experiments

In Fig. 3 the EISF for SA and SA-B complex is shown, the EISF of the SA shows a steeper decline than that of the SA-B complex, this indicates that the MSD of SA-B is reduced when compared to the MSD of SA. This indicates that the SA-B complex is less flexible in its motions than the free SA at the nanosecond range. Calculating the change in conformational entropy for the different MSD yields $\Delta S^{QENS} = -2.0 \pm 0.2 \text{ kJ mol}^{-1} \text{ K}^{-1}$. This indicates a decrease in the conformational order of the protein upon ligand binding. It is expected that the more rigid structure of the complex is stabilised by an increase in the mobility of the water in the hydration layer, thus also compensating the decrease in conformational entropy by increasing in the hydration layer [18].

DISCUSSION

In the neutron scattering experiment only the entropic contribution of the protein is probed, while the isothermal titration probes the protein and the hydration shell. Assuming that the contributions of biotin are small compared to those of SA, the entropic contribution of the hydration shell $\Delta S^{hydration}$ can be calculated from the difference of the two.

$$
\begin{aligned}
\Delta S^{QENS} &= S_{SA-B}^{protein} - S_{SA}^{protein} \\
\Delta S^{ITC} &= S_{SA-B} - S_{SA} = \left(S_{SA-B}^{protein} + S_{SA-B}^{hydration} \right) - \left(S_{SA}^{protein} + S_{SA}^{hydration} \right) \\
\Delta S^{ITC} - \Delta S^{QENS} &= S_{SA-B}^{hydration} - S_{SA}^{hydration} = \Delta S^{hydration}
\end{aligned}
\tag{8}
$$

Ligand-displacement ITC (LDITC) measurements of the binding of SA with biotin in the *same* buffer solution used in this work were carried out by Kuo *et al.* [8], who calculated the entropic contribution at 25°C as $\Delta S = -52.48 \text{ cal mol}^{-1} \text{ K}^{-1}$ with an error of 15% and the biotin/SA ratio of $n = 4.0$. Note that Kuo *et al.* used water instead

of heavy water. The entropic contribution of the whole tetramer is then $\Delta S^{\mathrm{ITC}} = n \cdot \Delta S$.

$$\Delta S^{\mathrm{QENS}} = -2.0 \pm 0.2 \ \mathrm{kJ\,mol^{-1}\,K^{-1}}$$
$$\Delta S^{\mathrm{ITC}} = -0.88 \pm 0.13 \ \mathrm{kJ\,mol^{-1}\,K^{-1}} \tag{9}$$
$$\Delta S^{\mathrm{hydration}} = -0.88 \pm 0.13 \ \mathrm{kJ\,mol^{-1}\,K^{-1}} + 2.0 \pm 0.2 \ \mathrm{kJ\,mol^{-1}\,K^{-1}} = 1.12 \pm 0.33 \ \mathrm{kJ\,mol^{-1}\,K^{-1}}$$

These calculations show that the entropy of the hydration shell increases upon ligand binding, compensating in large parts the entropic loss of the complex. To connect these results with a microscopic picture, a comparison can be made to the work of Liese *et al.* [18], who investigated hydration effects on stretched polyethylene glycol (PEG) chains. They, too, found that hydration has a significant contribution towards the free energy of the observed system: the loss in conformational entropy of the PEG chain due to the increase in stretching energy is compensated by an increase of entropy in the hydration shell. Their molecular dynamics simulations showed that the entropy gain was due to the replacement of double by single hydrogen-bonded hydration water.

Applying these observations to our system, the increasing entropy of the hydration shell upon ligand binding could indicate a reduction of hydrogen bonds between the complex and the hydration water compared to the number of hydrogen bonds between free SA and hydration water. Reasons for this behaviour are not clear, but might include the displacement of hydration water by biotin, the increased rigidity of the protein or a reduction of surface area of the protein due to conformational changes. The finding that the number of hydrogen bonds between SA and the surrounding water is reduced upon ligand binding is in agreement with the IR-TDFRS results indicating less hydrophilic behaviour for the complex.

It has to be noted, however, that the neutron scattering experiments are performed in D_2O, while Kuo *et al.* used normal water for their LDITC experiments. As can be seen in Fig. 2 the thermophoretic behaviour is strongly influenced by the addition of D_2O. Similar results are expected for the SA-B complex. The change of S_T is of the same order of magnitude as the change upon binding with biotin. Since it is known [19, 20, 5] that the thermophoretic measurements are very sensitive to changes of the hydrogen binding, also the entropic changes of the ITC measurements will be influenced, if water is replaced by D_2O. Therefore, it is necessary to perform the ITC measurements in a buffer with deuterated water, in order to determine the correct $\Delta S^{\mathrm{hydration}}$ by comparing ΔS^{ITC} with ΔS^{QENS} determined by the QENS measurements.

ACKNOWLEDGMENTS

REFERENCES

[1] S. Wiegand, "Introduction to thermal gradient related effects," in *Functional Soft Matter*, edited by J. Dhont, G. Gompper, G. Meier, D. Richter, G. Vliegenthart, and R. Zorn (Forschungszentrum Jülich, Jülich, 2015), pp. F4.1–F4.24.

[2] M. Jerabek-Willemsen, T. Andr, W. Wanner, H. Roth, S. Duhr, P. Baaske, and D. Breitsprecher, J. Mol. Struct. **1077**, 101–113 (2014).

[3] Y. Kishikawa, S. Wiegand, and R. Kita, Biomacromolecules **11**, 740–747 (2010).

[4] S. Iacopini, R. Rusconi, and R. Piazza, Eur. Phys. J. E **19**, 59–67 (2006).

[5] D. Niether, T. Kawaguchi, J. Hovancova, K. Eguchi, J. K. G. Dhont, R. Kita, and S. Wiegand, Langmuir **33**, 8483–8492 (2017).

[6] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, Adv. Drug Delivery Rev. **64**, 4–17 (2012).

[7] W. L. Jorgensen, Accounts of Chemical Research **22**, 184–189 (1989).

[8] T. C. Kuo, C. W. Tsai, P. C. Lee, and W. Y. Chen, Journal of Molecular Recognition **28**, 125–128 (2015).

[9] M. Bee, *Quasielastic neutron scattering: Principles and applications in solid state chemistry, biology and materials science* (Hilger, Bristol, 1988).

[10] F. Gabel, D. Bicout, U. Lehnert, M. Tehei, M. Weik, and G. Zaccai, Q. Rev. Biophys. **35**, 327–367 (2002).

[11] J. Fitter, J. Katsaras, and T. Gutberlet, *Neutron Scattering in Biology: [E-Book] : Techniques and Applications* (Springer-Verlag, Berlin, Heidelberg, 2006).

[12] A. Stadler, M. Koza, and J. Fitter, The journal of physical chemistry B **119**, 72–82 (2015).

[13] S. Wiegand and W. Köhler, LNP Vol. 584: Thermal Nonequilibrium Phenomena in Fluid Mixtures **584**, 189–210 (2002).

[14] P. Blanco, H. Kriegs, M. P. Lettinga, P. Holmqvist, and S. Wiegand, Biomacromolecules **12**, 1602–1609 (2011).

[15] J. Wuttke, A. Budwig, M. Drochner, H. Kmmerling, F.-J. Kayser, H. Kleines, V. Ossovyi, L. C. Pardo, M. Prager, D. Richter, G. J. Schneider, H. Schneider, and S. Staringer, Rev. Sci. Instrum. **83**, p. 75109 (2012).

[16] MLZ *et al.*, J. large-scale Res. Facil. **1**, p. A30 (2015).

[17] A. E. Kamholz, E. A. Schilling, and P. Yager, Biophysical Journal **80**, 1967–1972 (2001).

[18] S. Liese, M. Gensler, S. Krysiak, R. Schwarzl, A. Achazi, B. Paulus, T. Hugel, J. P. Rabe, and R. R. Netz, ACS Nano **11**, 702–712 (2017).

[19] R. Sugaya, B. A. Wolf, and R. Kita, Biomacromolecules **7**, 435–440 (2006).

[20] K. Eguchi, D. Niether, S. Wiegand, and R. Kita, The European Physical Journal E **39**, p. 86 (2016).

8 Discussion and Conclusion

8.1 Discussion

There are two main applications for thermal gradients in aqueous systems:

- Trapping and enriching of components in pores through a combination of thermodiffusion and convection.

- Monitoring of protein-ligand binding through the strong changes of thermodiffusion upon complex formation.

The former is relevant in the context of origin-of-life theories, as an accumulation mechanism like this is a possible solution to the concentration problem. It has been investigated for larger biomolecules (nucleotides and RNA-fragments) in dilution [65]. We tested the mechanism for formamide to see if such small molecules could also be accumulated and improved it by including temperature and concentration dependence of the input parameters (Ch. 2). In Ch. 3 an heuristic explanation for the accumulation process was offered.

The latter application is already utilised and commercially available as microscale thermophoresis (MST) [74], which has gained popularity in recent years. This method uses the sensitivity of S_T to changes in the molecule as an indicator of complex formation. The main goal of our investigation is to find out how the difference in S_T between free protein and complex is connected to modifications in their hydration shells and if MST could be used to obtain information about protein hydration in addition to the reaction kinetics. Ligand binding of proteins is of course rather complex, as conformational changes are often involved. Therefore, we first investigate simpler complexes with rigid cyclodextrin molecules (Ch. 5 and 6). To understand the localisation of molecules in the hydration layer and its influence

on thermodiffusion, we investigated urea, a small molecule accessible by non-equilibrium molecular dynamics (NEMD) simulations (Ch. 4). Finally, first results on the protein-ligand system streptavidin-biotin were presented in Ch. 7.

8.1.1 Thermophoretic accumulation

The basic principle of thermophoretic accumulation has been outlined in Sec. 1.2.1. In the context of origin-of-life theories, the accumulation in hydrothermal pores has been investigated before for the formation of RNA: the accumulation of RNA-fragments and nucleotides [65]. In Ch. 2 we showed that this kind of accumulation process is also feasible for small molecules that could have been formed through inorganic pathways, *e.g.* photochemically in the atmosphere, providing a link from inorganic educts to small prebiotic molecules like nucleotides. Since nucleotides and RNA are soluble in water only up to a relatively small concentration, for that case the model worked well within the assumption of a dilute regime, with a fixed S_T-value for the solute and the properties of water for the solution. For formamide, which is miscible with water at any ratio, the concentration and temperature dependence of the mixture properties was taken into account by our numerical model.

In Ch. 3 the accumulation mechanism was investigated in detail. In the case of thermogravitational columns the degree of separation that can be achieved depends on the height of the column. As the simulated hydrothermal pores are principally very similar, the observation that accumulation is more effective in higher pores (at larger aspect ratios) was expected. However, what we found was a rapid increase of accumulation at a certain aspect ratio above which formamide concentrations at the bottom of the pore approached saturation. A closer analysis of the accumulation process shows that while in shorter pores the accumulation rate decreases over time, in higher pores it increases exponentially until a concentration of ~ 50 wt% is reached and then drops fast due to saturation. This behaviour could be explained by a simple heuristic model (see Fig. 8.1): thermodiffusion has to drive the particle that is to be accumulated from the convectional upstream, where it would be washed out of the pore, into the downstream that carries it to the bottom. If the pore is high enough, the concentration at the top of the pores decreases, leading to diffusion of the components from the reservoir with fixed concentration into the pore. This explains the very effective accumulation that is observed once the pore reaches a certain height and the increase in accumulation rate, since thermodiffusion is at a maximum at 50 wt%.
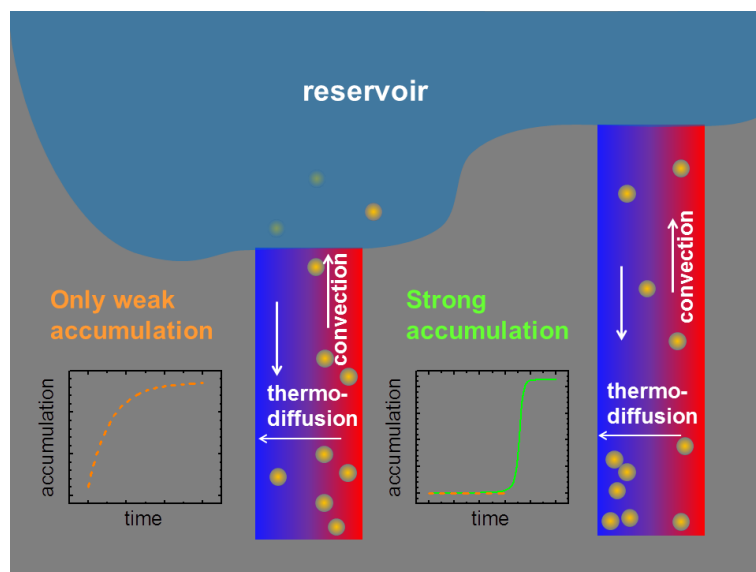
Figure 8.1: *Heuristic model of the accumulation process: (left) in a short pore convection flushes the component out of the pore, (right) the long pore allows thermodiffusion to carry the component into the downstream before it can be flushed out.*

8.1.2 Influence of hydrophilicity

At several points in this work, the temperature dependence of the Soret coefficient (measured either as ΔS_T, where $\Delta S_T = S_T(T_2) - S_T(T_1)$ with the temperatures $T_1 < T_2$, or the slope of S_T) has been correlated with $\log P$, which is a measure for the hydrophilicity of a chemical compound. The rationale behind this connection lies in the additivity of mass-inertia S_T^m and chemical contribution S_T^{chem} to the Soret coefficient (Eq. 1.4). While the mass-inertia contribution is not temperature dependent, the chemical contribution in an aqueous solution is determined by the strength and number of hydrogen bonds and therefore very sensitive to temperature change. The empirical equations describing $S_T(T)$ (Eq. 1.8 and Eq. (6) and (7) in Ch. 6) follow this interpretation as they are a sum of the temperature independent term S_T^∞ (approached at high temperature, where interactions are weakened) and a temperature dependent term describing the chemical contribution. Subtraction of S_T^∞ would be the obvious approach to get the chemical contribution S_T^{chem} from the measured S_T, but the error is large, because S_T^∞ is determined by extrapolation out of the accessible temperature range. For this reason we assume a proportionality between the amplitude of the chemical

contribution and the temperature dependece of S_T observed in a given temperature range

$$S_T^{chem} \propto \Delta S_T. \tag{8.1}$$

We find a correlation of ΔS_T with the partition coefficient $\log P$, which characterises solute-solvent interactions. Partition is observed when a solute can freely diffuse between two immiscible phases, most commonly 1-octanol and water. The partition coefficient is usually only measured under standard conditions and only defined for one single dilute solute. [75] In equilibrium, the ratio of concentration in the two phases is a constant value, the partition coefficient P, characteristic to the solute

$$P = \frac{[A]_{1-octanol}}{[A]_{water}}. \tag{8.2}$$

The hydrophilicity scale resulting from these experiments orders chemical compounds by their tendency to leave or accumulate in aqueous phases. This makes $\log P$ a handy parameter to model transport processes [76, 77]. The partition between the two phases is driven by the difference in solvation free energy of the solute in different phases. The energy of transfer of a solute A from water to the organic phase ΔG_t can be calculated from $\log P$ [78] with

$$\Delta G_t(A) = -RT \ln \frac{[A]_{1-octanol}}{[A]_{water}} = -2.303 \cdot RT \log P. \tag{8.3}$$

An analogy can be drawn between partition and thermodiffusion (compare Fig. 8.2). In the shaking flask experiment used to determine $\log P$ there is a concentration gradient between aqueous and organic phase. In a thermodiffusion experiment the concentration gradient is between two 'phases' of the same solvent, but with a different temperature. The difference in solvation free energy between the two zones should depend on the changed interactions with the surrounding solute and therefore determined by similar microscopic principles in both cases.

To understand how a concentration gradient can arise through diffusion it is important to remember that the thermodynamic force F that causes Fickian diffusion acts along a chemical

potential gradient. With the chemical potential

$$\mu = \mu^{\Theta} + RT \ln a \tag{8.4}$$

force F is

$$F = -\left(\frac{\partial \mu}{\partial x}\right)_{p,T} = -RT\left(\frac{\partial \ln a}{\partial x}\right)_{p,T} = -\frac{RT}{a}\left(\frac{\partial a}{\partial x}\right)_{p,T}. \tag{8.5}$$

The activity a is the concentration a solution should have at ideal behaviour, given by

$$a = \gamma c \tag{8.6}$$

the product of the activity coefficient γ, which corrects the non-ideality, and the actual mole fraction of the solute c. The Fickian law

$$J = -D\frac{\mathrm{d}c}{\mathrm{d}x}, \tag{8.7}$$

where the particle flux $J \propto F$, is based on the assumption of an ideal solution where $a = c$.
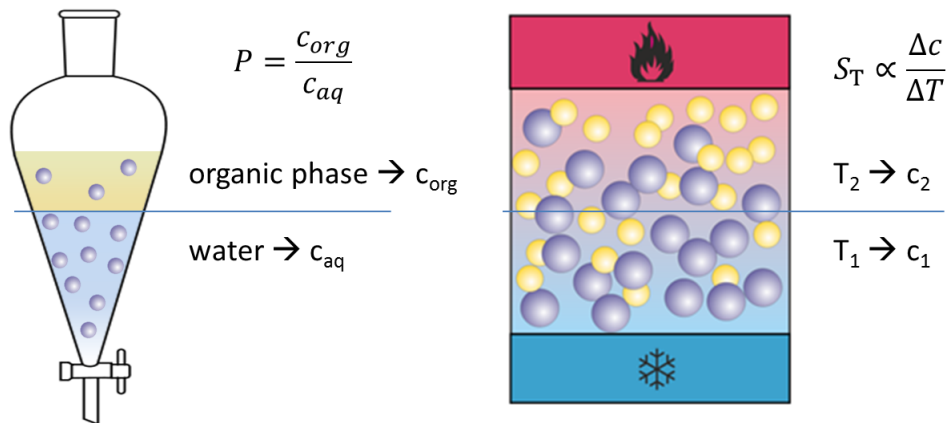

Figure 8.2: *Analogy between partition and thermodiffusion: (left) in a shaking flask with two immiscible phases the ratio of solute concentration in the two phases is the partition coefficient **P**, (right) thermodiffusion leads to a difference in solute concentrations in the two temperature zones.*

As long as the activity only varies with solute concentration, this assumption holds even in non-ideal solutions. In the case of the partition coefficient, where different solvents are involved, it is more useful to write the Fickian law as

$$J = -D_a \frac{\mathrm{d}a}{\mathrm{d}x}. \tag{8.8}$$

Now the particle flux vanishes when the activities of the solute in both solvents (the aqueous and organic phase, denoted by the indices 'aq' and 'org', respectively) are equal:

$$a_{org} = a_{aq} \tag{8.9}$$

$$\gamma_{org} \cdot c_{org} = \gamma_{aq} \cdot c_{aq} \tag{8.10}$$

$$\frac{c_{org}}{c_{aq}} = \frac{\gamma_{aq}}{\gamma_{org}}. \tag{8.11}$$

With $\gamma_{org} \approx 1$ (nearly ideal behaviour in the organic phase), we get

$$\log P \approx \log \gamma_{aq}, \tag{8.12}$$

a relation that has been observed experimentally [79].

In the case of thermodiffusion the particle flux can be described after Eq. 1.1 as

$$J = -D \frac{\mathrm{d}c}{\mathrm{d}x} - c_0(1 - c_0)D_T \frac{\mathrm{d}T}{\mathrm{d}x}, \tag{8.13}$$

where the diffusion coefficient is explicitly not dependent on activity, but on concentration. This is practical, because it leads to a definition of S_T that is proportional to the readily observable gradients in concentration and temperature. By this definition the observed flux J along the temperature gradient is described in relation to the flux that would occur due to Fickian diffusion in an ideal system under isothermal conditions and any non-ideality is necessarily contained in D_T.

To get a picture how the Soret coefficient might be connected to the activity, we suppose now that, analogous to the partition effect, the thermodiffusion is driven by Fickian diffusion
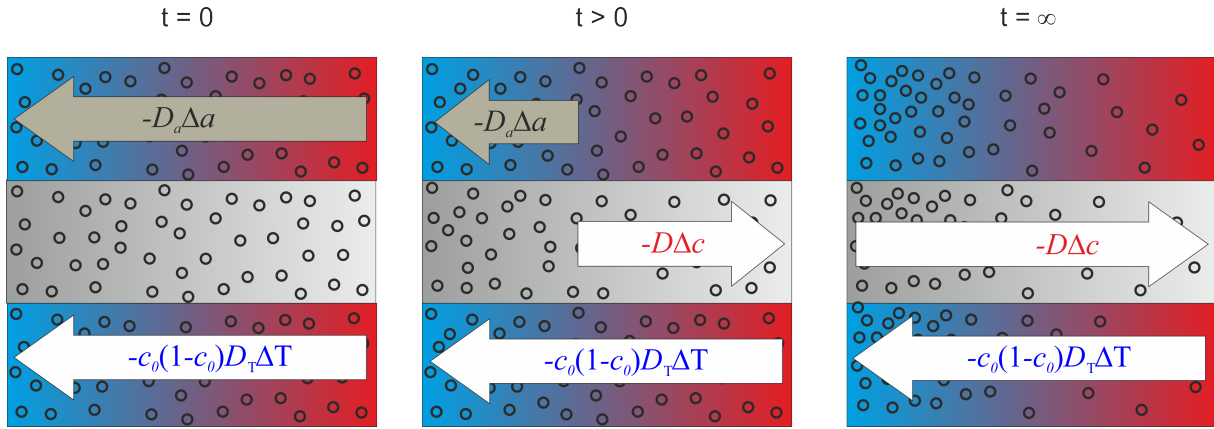
Figure 8.3: *Illustration of the two descriptions of the thermodiffusion flux discussed in the text before thermodiffusion sets in (t = 0), during the process (t > 0), and at the steady state (t = ∞). As Fickian diffusion along a temperature-induced activity gradient according to Eq. 8.8 (grey arrow) and as a sum of thermodiffusion and an opposed ideal Fickian diffusion according to Eq. 8.13 (white arrows).*

along an activity gradient that arises due to a temperature dependence of a. The same flux as in Eq. 8.13 can then be described by Eq. 8.8 and we get

$$D_a \frac{\mathrm{d}a}{\mathrm{d}x} = D \frac{\mathrm{d}c}{\mathrm{d}x} + c_0(1 - c_0) D_\mathrm{T} \frac{\mathrm{d}T}{\mathrm{d}x}. \tag{8.14}$$

Note that any influence of the heat flux is ignored and that Eq. 8.5 on which Eq. 8.8 is based does not account for a temperature dependence of the chemical potential, so that a definition of D_a is vague.

Before the flux sets in $\mathrm{d}c/\mathrm{d}x = 0$ and with constant concentration

$$\frac{\mathrm{d}a}{\mathrm{d}T} = c_0 \frac{\mathrm{d}\gamma}{\mathrm{d}T}, \tag{8.15}$$

so that we can rewrite Eq. 8.14 by

$$\frac{\mathrm{d}\gamma}{\mathrm{d}T} = \frac{D(1 - c_0)}{D_a} S_\mathrm{T}. \tag{8.16}$$

As pointed out before, this solution reflects a very simplified view on the problem as the

thermal non-equilibrium is not considered. However, it makes obvious that a temperature dependence of the activity coefficient (and consequently of the chemical potential), if it exists, can have a strong influence on the Soret coefficient. As the underlying assumption here is that an activity gradient is the *only* driving force of thermodiffsuion, this solution indicates a proportionality of the chemical contribution rather than of the global Soret coefficient:

$$S_T^{chem} \propto \frac{\mathrm{d}\gamma}{\mathrm{d}T}. \tag{8.17}$$

Connecting the observation of a correlation between ΔS_T and $\log P$ to Eq. 8.16, some qualitative remarks can be made. If γ is determined by the number of hydrogen bonds, the energy of which decreases linearly with temperature, then we expect an exponential decay for $\gamma(T)$, with a negative slope that decreases with increasing temperature. This gives us a negative chemical contribution to the Soret coefficient S_T^{chem} that has a large amplitude at low temperatures and vanishes at large temperatures, which fits well with the typical temperature dependence of S_T as described by Eq. 1.8. It also means that $\log \gamma \propto -\mathrm{d}\gamma/\mathrm{d}T$, which fits the linear correlation observed between ΔS_T ($\propto S_T^{chem}$) and $\log P$ ($\approx \log \gamma$) at moderate $\log P$-values.

As mentioned before, the prediction of S_T works well for non-polar mixtures, but fails for polar ones. In her comparison of thermodynamic models [80,81] Gonzalez-Bagnoli finds that while the deviation from experimental data is 10-20% for non-polar systems, for polar ones 200-300% error are not unusual.

Of course, existing thermodynamic models, some of which are presented in Sec. 1.1.2, contain expressions for the chemical potential, but as the activity coefficient is difficult to obtain experimentally, these are usually approximations. For example Eq. 1.7 contains the concentration dependence of the chemical potential $x_1(\partial\tilde{\mu}_1/\partial x_1)_{p,T} = [1 + (\partial \ln \gamma_1/\partial \ln x_1)_{p,T}]$, but does not consider any temperature dependence. And indeed, for non-polar mixtures, such a temperature dependence is not expected. However, it is known that hydrogen bonds are sensitive to temperature. Measurements of the heat of vaporisation H_{vap} have shown that hydrogen bond strength in pure water has a linear temperature dependence between 273-433K [82]. In this range H_{vap} falls with rising temperature by 45 Jmol^{-1}K^{-1}. Any non-ideality of aqueous solutions caused by hydrogen bonds must therefore depend quite strongly on temperature.

Interestingly, there have been numerous attempts to connect S_T to the activity coefficient

or to parameters that should have a close, if often unclear, connection to non-ideality. For example, thermodiffusion has been correlated with the Hildebrand parameter [84] or with the thermal expansion coefficient [16], which might reflect hydrogen bond length and therefore energy. The Hildebrand parameter is used as a measure for the 'likeness' or compatibility of chemical compounds, so that substances with good mutual solubility have similar Hildebrand parameters and the difference increases with decreasing solubility.

However, the microscopic interpretation of hydration (solvation of a solute in water) is not clear. Formation of holes to accommodate the solute disrupts the hydrogen bond (HB) network of water and costs energy, the formation of HBs between solute and the surrounding water reduces solvation free energy. There is also a significant entropic effect involved in hydration: HBs with the solute can disrupt the HB network and raise entropy, but entropy might also be reduced when water molecules are fixed in position by double HB with a hydrophilic solute [83] or when clathrate-like cages are formed around hydrophobic parts of a solute.

A rise of S_T has long been related to a breaking of hydrogen bonds [56]. This work indicates that this is due to a reduced chemical contribution, which has a negative value when there is strong hydrogen bond interaction. It could be shown that, because it is not influenced by the mass-inertia contribution to the Soret coefficient, the temperature sensitivity of thermodiffusion is a better measure for hydrogen bond interaction inside a mixture.

8.1.3 Thermodiffusion of complexes

The second focus of this work is the question how this sensitivity of thermodiffusion to solute-solvent interactions could be used. For some years now, Microscale Thermophoresis (MST) [45] has been one of the most important applications for thermodiffusion. As detailed in Sec. 1.2.2, this method is able to determine binding constants of biomolecular interactions, such as a ligand binding to a protein, by observing the change in their response to a temperature gradient in dependence of ligand concentration.

Chapters 5 and 6 examines cyclodextrins (CD), cyclic oligosaccharides that are known to form complexes with several small drug molecules. Their complexes with acetylsalicylic acid (ASA) are investigated in detail in the latter chapter. The system was chosen as a model for complexes, because, in contrast to proteins, there are no conformational changes

expected from these small and rather rigid molecules, they are not charged, and stable without addition of buffer. A shift towards higher S_T-values and a reduced T-dependence for all complexes formed with ASA is observed compared to the free CDs. ASA, which is hydrophobic ($\log P \approx 1$), has a negative $S_T(T)$-slope. The partition coefficient is not strictly defined for complexes, but could be estimated. The observed temperature dependence ΔS_T is much smaller than expected from these values. NMR-measurements show that the fraction of complex is only $p_{com} = (45 \pm 12)$ % so that free CD and ASA contribute significantly to the TDFRS-signal of the mixture, since diffusion of all components is in the same order of magnitude and can therefore not be separated. Assuming that interactions between the particles are negligible at these low concentrations and the measured signal is an average of the component contributions, the ΔS_T of the complexes can be calculated. It turns out that these values fit the correlation curve of ΔS_T vs. $\log P$, but the change is rather small compared to the free CDs.

As a next step streptavidin (SA) and its complex with biotin (SA-B) have been investigated as a model system for protein-ligand complexes in Ch. 7. In contrast to the cyclodextrins, the TDFRS-signal of SA is well separated from that of biotin by the large difference in diffusion speed. While this simplifies the data analysis, the interpretation is still difficult due to the complexity of the protein. The results show that SA and SA-B show the same thermodiffusion at 10°C, but then S_T of the complex rises slower with temperature. This weaker temperature dependence indicates that the complex is less hydrophilic than the free SA. Fewer hydrogen bonds lead to increased mobility of the surrounding water molecules and higher entropy of the hydration layer. This quantitative interpretation aligns with an increased entropy of the hydration shell upon complex formation calculated from quasi-elastic neutron scattering (QENS) and isothermal titration calorimetry (ITC) experimental results. Obviously, a lot more experimental data on different systems would be required to attempt a quantitative interpretation of IR-TDFRS data.

Figure 8.4 shows the thermodiffusion against temperature for SA and SA-B at different biotin concentrations. Streptavidin naturally occurs as a tetramer in which each unit can bind one biotin molecule. This means that saturation is reached at a molar ratio of 1:4, lower biotin concentrations should lead to unsaturated complexes. The data shows that thermodiffusion of the complex depends only weakly on biotin concentration. This might indicate that the conformational change of the protein that leads to the altered thermodiffusion behaviour of the complex in comparison to the free SA already occurs when the first biotin binds. This is
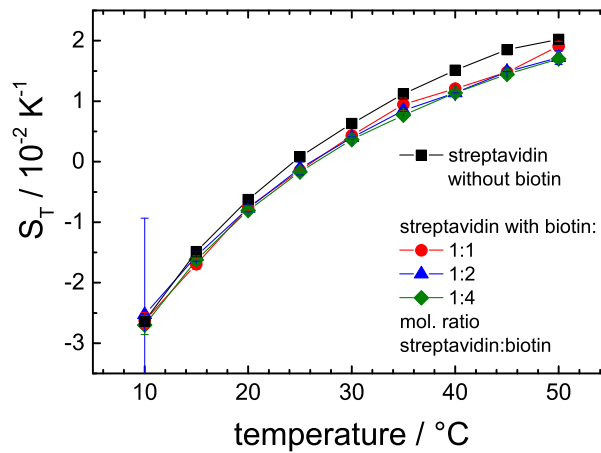
Figure 8.4: *Soret coefficient S_T against temperature for streptavidin (SA, black symbols) and mixtures of SA and biotin (coloured symbols) in buffer. The concentration of SA is 50 g/L, the molar ratio of biotin is specified in the legend.*

corroborated by insensitivity to biotin concentration of the denaturation temperature, which has been observed by circular dichroism (CD) spectroscopy [85]. Such a behaviour might point to cooperative binding, which is common in biomolecules, but controversial in the case of streptavidin-biotin [86–89].

8.2 Conclusion

8.2.1 Accumulation in a hydrothermal pore

A process was investigated by which one component in a mixture is accumulated through a combination of thermodiffusion and convection (Ch. 2 and 3). This has been proposed as a possible mechanism at the origin of life in order to reach high enough concentrations of organic molecules and make polymerisation into more complex structures feasible [64, 65]. Our numerical calculations simulate the accumulation of formamide in hydrothermal pores with a width of 100-200 μm, taking into account the temperature and concentration dependence of S_T and other thermophysical properties of the formamide/water mixture. This is necessary, because the accumulation can become very effective and therefore make the assumption of a dilute regime impractical.

We found that the thermodiffusion behaviour of formamide is ideally suited to the investigated accumulation process, because S_T is positive across the measured concentration and temperature range, and it increases with rising concentration, so that the process is self-enhanced. Apart from the thermodiffusion coefficient D_T, efficiency of accumulation depends on the temperature gradient ΔT, the ambient temperature T, starting concentration of formamide c_0, and pore geometry. We found that accumulation up to a formamide concentration of $c = 85$ wt% at the bottom of the pore from $c_0 = 10^{-3}$ wt% is possible in under three months.

The dependence on pore geometry was investigated in more detail in Ch. 3 and it was found that accumulation becomes extremely effective at a certain aspect ratio $r^* = L_h/L_w$, where L_h is the height of the pore and L_w is its width. This is due to the fact that a certain height allows the formamide to diffuse into the convective downstream $(c(1-c)D_T\Delta T)$ before it can be washed out of the pore by the convective upstream and Fickian diffusion $(v_{conv} + D \cdot \Delta_y c)$. The minimum aspect ratio for this occurrence can be estimated with the following equation:

$$r^* = \sqrt{\frac{L_y \cdot v_{conv} + D \cdot \Delta_y c}{4 \cdot c(1-c)D_T\Delta T}}. \tag{8.18}$$

8.2.2 Hydrophilicity and the temperature dependence of S_T

A correlation between the partition coefficient $\log P$ and the temperature dependence of S_T was observed in Ch. 5 and investigated further in Ch. 4 and 6. While the partition coefficient itself is a phenomenon that is not well understood on a microscopic level, $\log P$ is generally interpreted as a measure for hydrophilicity. It is closely related to the activity coefficient γ and therefore with the strength of solute-solvent interactions in aqueous solutions. We found that a positive slope $(dS_T/dT > 0)$ indicates hydrophilic behaviour (corresponding to the 'typical' behaviour in aqueous solutions according to Eq. 1.8) and a negative slope $(dS_T/dT < 0)$ indicates hydrophobic behaviour, which leads to solutions that are inhomogeneous on a microscopic level due to aggregation of the solute.

The change of temperature dependence observed when concentration is varied can also be explained along those lines. In the case of urea in water, a rise in urea concentration leads to aggregation, which has the same effect as a reduction in urea-water interactions. Further we

argued that the tendency of urea to aggregate at higher concentrations might be the reason for its properties as a protein denaturant: when proteins upon present and the urea concentration is high enough, urea would tend to aggregate along the protein surface, displacing water from the hydration shell ('dry globule' formation) and leading to an unfolding of the protein.

In Ch. 5 S_T of several cyclodextrins at low concentration in water and in formamide was compared. The large change in S_T observed between the two solvents is partly caused by diffusion, which is faster and more temperature dependent in water. But there is also a significant shift towards lower D_T-values in the more hydrophilic cyclodextrins, which does not occur for the methylated one that has weaker interactions with water. In conclusion we can say that the reduction of hydrogen bond interaction between solute and solvent (by methylation of the cyclodextrin or by exchanging water with the less polar formamide) raises D_T and with it S_T, as has been noted before [41, 90]. Apart from this qualitative observation it is difficult to extrapolate from these results to systems with other solvents, because there is no easy way to predict solute-solvent interactions in liquids on a microscopic level.

In Sec. 8.1.2 the temperature sensitivity of the activity coefficient γ was discussed as one of the driving forces for the Soret effect. Since γ is a correction factor to account for non-ideality, it depends on a number of microscopic effects about which little is known. It stands to reason, however, that in aqueous solutions it is dominated by hydrogen bonds and should approach unity at high temperatures when hydrogen bonds are weakened. This might explain the observation of S_T approaching a constant value S_T^∞ at high temperatures.

8.2.3 Change of thermodiffusion behaviour upon complex formation

The sensitivity of thermodiffusion to interactions between the solute and surrounding solvent inspired the idea that changes in the thermodiffusion behaviour upon complex formation could tell us something about what modifications occur in the hydration shell. Two model systems where investigated in order to test this hypothesis: Cyclodextrin complexes with acetylsalicylic acid (ASA) in Ch. 6 and streptavidin-biotin complexes in Ch. 7.

Cyclodextrins are cyclic oligosccharides and the smaller ones that were investigated here are rigid in shape. They form complexes by hydrophobic interaction of the drug molecule with the interior of the ring. We could show that the additivity found for homologous groups by

Maeda *et al.* [57] holds and S_T of such relatively small systems can be estimated by counting donor and acceptor groups. In the investigated cases the acceptor groups are ether bonds and their number is proportional to additional hydrophobic groups, so that what is actually weighted against each other with this method is hydrophilic and hydrophobic contributions in the molecule. In the case of the investigated system this approach even worked for the complexes by including donor- and acceptor groups of the ASA, but this might depend on the degree to which the guest molecule in the complex is still in contact with the solvent. However, it turned out that the relatively large difference in ΔS_T was not due to a strongly modified hydration of the complex, but the free ASA remaining in solution. It appears that the change in hydration upon complex formation is comparably small for the cyclodextrins.

Streptavidin is a protein that binds to biotin with one of the strongest known non-covalent bonds, the K_a is in the order of magnitude of 10^{12} to 10^{14} M^{-1}. There is a significant shift of S_T and ΔS_T when biotin is added. In this system, due to the fact that the diffusion of biotin is by an order of magnitude faster than that of streptavidin, the TDFRS-signals are well separated, so that the change is due to differences in the hydration shells of free protein and complex. The decrease of ΔS_T upon complex formation indicates a decrease in hydrophilicity. Quasi-elastic neutron scattering (QENS) shows that the protein becomes more rigid when the complex is formed. Comparison with data from isothermal titration calorimetry (ITC) shows that this rather strong decrease in entropy of the protein is partly compensated by an increase in entropy of the hydration shell. This fits with a breaking of hydrogen bonds between the complex and surrounding water in a similar mechanism as described by Liese *et al.* [83].

8.3 Outlook

The partition coefficient $\log P$ is used to screen for pharmacological compounds [76] and to model transport processes in environmental science [77]. Apart from the sometimes problematic accuracy of incremental algorithms that are nowadays used to determine $\log P$, the parameter is usually only given at standard conditions and defined for one single solute at high dilution. Depending on the process, transport models might benefit greatly if the dependence of hydrophilicity on concentration, temperature, and co-solvents could be taken into account. The temperature dependence of S_T might offer a suitable hydrophilicity scale.

In the discussion the correlation between ΔS_T and $\log P$ was rationalised by being based in the proportionality $S_T^{chem} \propto \mathrm{d}\gamma/\mathrm{d}T$ due to a temperature dependence of the chemical potential. Re-analysing the data with that in mind should give a less convoluted picture.

For the urea/water system simulations showed that the slope change of $S_T(T)$ from positive to negative with addition of urea is connected to its aggregation at higher concentrations. For some systems (*e.g.* ethanol/water) the opposite behaviour has been observed. It would be interesting to get a microscopic picture as to why this is the case.

Models for the prediction of S_T are currently not successful for polar mixtures. They could be improved by taking into account the temperature dependence of the activity coefficient γ, which seems to be negligible in non-polar mixtures, but should be much larger in aqueous solutions due to the temperature dependence of hydrogen bond strength. Finding such an expression and including it into a model like that of Hartmann *et al.* [25, 26] might improve the predictions for polar mixtures considerably.

We made a connection between change in thermodiffusion behaviour and entropy change in the hydration shell of a protein upon complex formation and could show qualitative agreement for streptavidin-biotin. This hypothesis has to be tested more rigorously for different systems and it is to be seen if a correlation can be established that would make it possible to estimate the entropic contribution of the hydration shell $\Delta S^{\mathrm{hydration}}$ from the temperature dependence of thermodiffusion. It would also be rather interesting to investigate in more detail streptavidin-biotin and the dependence of its thermodiffsuion on temperature and biotin concentration.

Since many proteins are charged, another aspect for future research would be to widen the investigation of the hydration shell towards charged systems. As a first start one could connect to the cyclodextrin studies and study processes at different pH.

So far the accumulation in pores has only been predicted in finite element calculation. It would be interesting to see, whether it could be realized experimentally. It would also be interesting to transfer the concept of pore accumulation to organic catalytic systems.

Acknowledgement

When I think about the writing of this thesis, I think about the last three years of my life. And while there were many challenges and stretches of stressful times, it was generally a rewarding and exciting experience. I was able to learn a lot, to travel and get to know many interesting and wonderful people.

First and foremost I want to thank Prof. Simone Wiegand. She has been there for me in the last three years, be it scientific questions, career advise or stolen bicycles. I can't imagine a more dedicated and supportive supervisor.

I thank Prof. Jan K. G. Dhont for allowing me to work in his fantastic group, for his continued support and interest, and for inspiring discussions.

I thank Prof. Annette Schmidt for her support, her interest in my work, and for giving me the possibility to do my PhD at the University of Cologne. I would also like to thank Prof. Hans-Günther Schmalz for taking the chair of my examination committee.

I thank Dr. Hartmut Kriegs for his technical support in aligning the setup, coding the evaluation programs and helping with any trouble the computer came up with.

My sincere thanks go to all the people who collaborated with me - Dr. Dzmitry Afanasenkau, Prof. Fernando Bresme, Dr. Silvia Di Lecce, Prof. Jörg Fitter, Dr. Bernd König, Dr. Andreas Stadler, Mona Sarter, Prof. Rio Kita, Jana Hovankova, and Kazuya Eguchi - for the pleasant teamwork, enlightening discussions and their contribution to this work.

I thank the International Helmholtz Research School of Biophysics and Soft Matter *BioSoft* for accepting me as a fellow and broadening my scientific horizon, and especially Dr. Thorsten Auth for all the work he puts into the organisation.

I thank my wonderful parents for their unconditional love and support. I thank my sister

Moni and her family for my cat and for always welcoming me when I wanted to get away from work. I thank my twin Christiane for always understanding me.

Last, but certainly not least, I thank all my friends in the ICS-3 for so many things scientific and non-scientific: discussions, help, distractions, constructive criticism, coffee breaks, lunches, pub and game nights, outings, and coming to my concerts. The last three years have been a happy time for me and the supportive atmosphere of this group and the kindness of the brilliant people around me had a big part in that.

ICS-3 October 2017

Bibliography

[1] C. Ludwig. Diffusion zwischen ungleich erwärmten Orten gleich zusammengesetzter Lösungen. *Sitz. ber. Akad. Wiss. Wien Math.-naturw. Kl.*, 20:539, 1856.

[2] C. Soret. Sur l'etat d'équilibre que prend au point de vue de sa concentration une dissolution saaline primitivement homogéne dont deux parties sont portées a des températures différentes. *Arch. Geneve*, 3:48–64, 1879.

[3] S.R. de Groot and P. Mazur. *Non-equilibrium Thermodynamics*. Dover, New York, 1984.

[4] W. Köhler and K. I. Morozov. The Soret effect in liquid mixtures - a review. *J. Non-Equilib. Thermodyn.*, 41:151–197, 2016.

[5] H. Brenner and J. R. Bielenberg. A continuum approach to phoretic motions: Thermophoresis. *Physica a - Statistical Mechanics and Its Applications*, 355:251–273, 2005.

[6] K. Clusius and G Dickel. Zur trennung der chlorisotope. *Naturwissenschaften*, 27:148–149, 1939.

[7] W. M. Rutherford. Separation of isotopes in thermal-diffusion column. *Sep. Purif. Methods*, 4:305–350, 1975.

[8] C. Debuschewitz and W. Köhler. Molecular origin of thermal diffusion in benzene plus cyclohexane mixtures. *Phys. Rev. Lett.*, 87:055901, 2001.

[9] S. Hartmann, W. Köhler, and K. I. Morozov. The isotope Soret effect in molecular liquids: a quantum effect at room temperatures. *Soft Matter*, 8:1355–1360, 2012.

[10] J. Schirdewahn, A. Klemm, and L. Waldmann. Thermodiffusion in D_2-HT und anderen Wasserstoffgemischen. *Zeitschrift für Naturforschung Part a - Astrophysik, Physik und physikalische Chemie*, 16:133–144, 1961.

[11] K. E. Grew, F. A. Johnson, and W. E. J. Neal. The thermal diffusion factor and temperature. *Proceedings of the Royal Society of London Series a - Mathematical and Physical Sciences*, 224:513–526, 1954.

[12] L. Waldmann. *Transporterscheinungen in Gasen von mittlerem Druck*, volume 3/12 of *Handbuch der Physik*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1958.

[13] I. Prigogine, L. Debrouckere, and R. Amand. Recherches sur la thermodiffusion en phase liquide .1. *Physica*, 16:577–598, 1950.

[14] S. Chapman. The kinetic theory of simple and composite monatomic gases viscosity, thermal conduction, and diffusion. *Proceedings of the Royal Society of London Series a - Containing Papers of a Mathematical and Physical Character*, 93:1–20, 1916.

[15] A. R. Allnatt and A. V. Chadwick. Thermal diffusion in crystalline solids. *Chem. Rev.*, 67:681–705, 1967.

[16] H. Brenner. Elementary kinematical model of thermal diffusion in liquids and gases. *Phys. Rev. E*, 74:036306, 2006.

[17] J. Rauch, M. Hartung, A. F. Privalov, and W. Köhler. Correlation between thermal diffusion and solvent self-diffusion in semidilute and concentrated polymer solutions. *J. Chem. Phys.*, 126:214901, 2007.

[18] P. Blanco, M. M. Bou-Ali, J. K. Platten, P. Urteaga, J. A. Madariaga, and C. Santamaria. Determination of thermal diffusion coefficient in equimolar n-alkane mixtures: Empirical correlations. *J. Chem. Phys.*, 129:174504, 2008.

[19] P. Blanco, H. Kriegs, B. Arlt, and S. Wiegand. Thermal diffusion of oligosaccharide solutions: The role of chain length and structure. *J. Phys. Chem. B*, 114:10740–10747, 2010.

[20] P. Blanco and S. Wiegand. Study of the Soret effect in monosaccharide solutions. *J. Phys. Chem. B*, 114:2807–2813, 2010.

[21] R. Haase. Zur Thermodynamisch-Phänomenologischen Theorie der Thermodiffusion. *Zeitschrift für Physik*, 127:1–10, 1949.

[22] L. J. T. M. Kempers. A comprehensive thermodynamic theory of the Soret effect in a multicomponent gas, liquid, or solid. *J. Chem. Phys.*, 115:6330–6341, 2001.

[23] K. G. Denbigh. The heat of transport in binary regular solutions. *Trans. Faraday Soc.*,

48:1–8, 1952.

[24] W. M. Rutherford and H. G. Drickamer. Theory of thermal diffusion in liquids and the use of pressure to investigate the theory. *J. Chem. Phys.*, 22:1157–1165, 1954.

[25] S. Hartmann, G. Wittko, W. Köhler, K. I. Morozov, K. Albers, and G. Sadowski. Thermophobicity of liquids: Heats of transport in mixtures as pure component properties. *Phys. Rev. Lett.*, 109:065901–1–065901–4, 2012.

[26] S. Hartmann, G. Wittko, F. Schock, W. Gross, F. Lindner, W. Köhler, and K. I. Morozov. Thermophobicity of liquids: Heats of transport in mixtures as pure component properties-the case of arbitrary concentration. *J. Chem. Phys.*, 141, 2014.

[27] M. Eslamian and M. Z. Saghir. Microscopic study and modeling of thermodiffusion in binary associating mixtures. *Phys. Rev. E*, 80:061201, 2009.

[28] P. A. Artola, B. Rousseau, and G. Galliero. A new model for thermal diffusion: Kinetic approach. *J. Am. Chem. Soc.*, 130:10963–10969, 2008.

[29] L. J. Tichacek, W. S. Kmak, and H. G. Drickamer. Thermal diffusion in liquids - the effect of non-ideality and association. *J. Phys. Chem.*, 60:660–665, 1956.

[30] P. A. Artola and B. Rousseau. Thermal diffusion in simple liquid mixtures: what have we learnt from molecular dynamics simulations? *Mol. Phys.*, 111:3394–3403, 2013.

[31] Bjørn Hafskjold. Computer simulations of thermal diffusion in binary fluid mixtures. *LNP Vol. 584: Thermal Nonequilibrium Phenomena in Fluid Mixtures*, 584:3, 2002.

[32] F. Bresme, A. Lervik, and J. Armstrong. Non-equilibrium molecular dynamics. *Experimental Thermodynamics, Vol X: Non-Equilibrium Thermodynamics with Applications*, 10:105–133, 2016.

[33] P. A. Artola and B. Rousseau. Microscopic interpretation of a pure chemical contribution to the Soret effect. *Phys. Rev. Lett.*, 98:125901–1–125901–4, 2007.

[34] G. Galliero, M. Bugel, B. Duguay, and F. Montel. Mass effect on thermodiffusion using molecular dynamics. *J. Non-Equilib. Thermodyn.*, 32:251–258, 2007.

[35] G. Galliero and S. Volz. Thermodiffusion in model nanofluids by molecular dynamics simulations. *J. Chem. Phys.*, 128:064505, 2008.

[36] M. C. Yang and M. Ripoll. Thermophoretically induced flow field around a colloidal particle. *Soft Matter*, 9:4661–4671, 2013.

[37] A. Perronace, C. Leppla, F. Leroy, B. Rousseau, and S. Wiegand. Soret and mass diffusion measurements and molecular dynamics simulations of n-pentane-n-decane mixtures. *J. Chem. Phys.*, 116:3718–3729, 2002.

[38] M. M. Zhang and F. Müller-Plathe. Reverse nonequilibrium molecular-dynamics calculation of the Soret coefficient in liquid benzene/cyclohexane mixtures. *J. Chem. Phys.*, 123:124502, 2005.

[39] P. Polyakov, M. Zhang, F. Müller-Plathe, and S. Wiegand. Thermal diffusion measurements and simulations of binary mixtures of spherical molecules. *J. Chem. Phys.*, 127:014502, 2007.

[40] S. Iacopini and R. Piazza. Thermophoresis in protein solutions. *Europhys. Lett.*, 63:247–253, 2003.

[41] Y. Kishikawa, S. Wiegand, and R. Kita. Temperature dependence of Soret coefficient in aqueous and nonaqueous solutions of pullulan. *Biomacromolecules*, 11:740–747, 2010.

[42] Z. Wang, H. Kriegs, and S. Wiegand. Thermal diffusion of nucleotides. *J. Phys. Chem. B*, 116:7463–7469, 2012.

[43] R. Piazza, S. Iacopini, and B. Triulzia. Thermophoresis as a probe of particle-solvent interactions: The case of protein solutions. *Phys. Chem. Chem. Phys.*, 6:1616–1622, 2004.

[44] M. Jerabek-Willemsen, C. J. Wienken, D. Braun, P. Baaske, and S. Duhr. Molecular interaction studies using microscale thermophoresis. *ASSAY Drug Dev. Technol.*, 9:342–353, 2011.

[45] M. Jerabek-Willemsen, T. André, W. Wanner, H.M. Roth, S. Duhr, P. Baaske, and D. Breitsprecher. Microscale thermophoresis: Interaction analysis and beyond. *J. Mol. Struct.*, pages 101–113, 2014.

[46] Susanne A. I. Seidel, Patricia M. Dijkman, Wendy A. Lea, Geert van den Bogaart, Moran Jerabek-Willemsen, Ana Lazic, Jeremiah S. Joseph, Prakash Srinivasan, Philipp Baaske, Anton Simeonov, Ilia Katritch, Fernando A. Melo, John E. Ladbury, Gideon Schreiber, Anthony Watts, Dieter Braun, and Stefan Duhr. Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions. *Methods*, 59:301–315, 2013.

[47] C. J. Wienken, P. Baaske, U. Rothbauer, D. Braun, and S. Duhr. Protein-binding assays

in biological liquids using microscale thermophoresis. *Nat. Commun.*, 1:1–7, 2010.

[48] Friederike M. Möller, Michael Kieß, and Dieter Braun. Photochemical microscale electrophoresis allows fast quantification of biomolecule binding. *J. Am. Chem. Soc.*, 138:5363–5370, 2016.

[49] N. Osterman and D. Braun. Thermooptical molecule sieve on the microscale. *Appl. Phys. Lett.*, 106, 2015.

[50] S. Iacopini, R. Rusconi, and R. Piazza. The "macromolecular tourist": Universal temperature dependence of thermal diffusion in aqueous colloidal suspensions. *Eur. Phys. J. E*, 19:59–67, 2006.

[51] D. Vigolo, S. Buzzaccaro, and R. Piazza. Thermophoresis and thermoelectricity in surfactant solutions. *Langmuir*, 26:7792–7801, 2010.

[52] J. K. G. Dhont, S. Wiegand, S. Duhr, and D. Braun. Thermodiffusion of charged colloids: Single-particle diffusion. *Langmuir*, 23:1674–1683, 2007.

[53] J. K. G. Dhont and W. J. Briels. Single-particle thermal diffusion of charged colloids: Double-layer theory in a temperature gradient. *Eur. Phys. J. E*, 25:61–76, 2008.

[54] K. A. Eslahian, A. Majee, M. Maskos, and A. Wurger. Specific salt effects on thermophoresis of charged colloids. *Soft Matter*, 10:1931–1936, 2014.

[55] K. A. Eslahian and M. Maskos. Hofmeister effect in thermal field-flow fractionation of colloidal aqueous dispersions. *Colloids and Surfaces a - Physicochemical and Engineering Aspects*, 413:65–70, 2012.

[56] R. Sugaya, B. A. Wolf, and R. Kita. Thermal diffusion of dextran in aqueous solutions in the absence and the presence of urea. *Biomacromolecules*, 7:435–440, 2006.

[57] K. Maeda, N. Shinyashiki, S. Yagihara, S. Wiegand, and R. Kita. Ludwig-Soret effect of aqueous solutions of ethylene glycol oligomers, crown ethers, and glycerol: Temperature, molecular weight, and hydrogen bond effect. *J. Chem. Phys.*, 143:124504, 2015.

[58] A. Königer, B. Meier, and W. Köhler. Measurement of the Soret, diffusion, and thermal diffusion coefficients of three binary organic benchmark mixtures and of ethanol-water mixtures using a beam deflection technique. *Philos. Mag.*, 89:907–923, 2009.

[59] B. C. Reed. Liquid thermal diffusion during the manhattan project. *Phys. Perspect.*, 13:161–188, 2011.

[60] Gary H. Thompson, Marcus N. Myers, and J. Calvin Giddings. An observation of a field-flow fractionation effect with polystyrene samples. *Separation Science*, 2:797–800, 1967.

[61] J. C. Giddings, V. Kumar, P. S. Williams, and M. N. Myers. Polymer separation by thermal field-flow fractionation - high- speed power programming. *Adv. Chem. Ser.*, 227:3–21, 1990.

[62] M. F. Riley and A. Firoozabadi. Compositional variation in hydrocarbon reservoirs with natural convection and diffusion. *AlChE J.*, 44:452–464, 1998.

[63] D. E. Rosner, R. S. Israel, and B. La Mantia. "heavy" species Ludwig-Soret transport effects in air-breathing combustion. *Combust. Flame*, 123:547–560, 2000.

[64] F. S. Gaeta, U. Bencivenga, P. Canciglia, S. Rossi, and D. G. Mita. Temperature-gradients and prebiological evolution. *Cell Biophysics*, 10:103–125, 1987.

[65] P. Baaske, F. M. Weinert, S. Duhr, K. H. Lemke, M. J. Russell, and D. Braun. Extreme accumulation of nucleotides in simulated hydrothermal pore systems. *P. Natl. Acad. Sci. USA*, 104:9346–9351, 2007.

[66] C. B. Mast and D. Braun. Thermal trap for dna replication. *Phys. Rev. Lett.*, 104:188102–1–188102–4, 2010.

[67] Simone Wiegand and Werner Köhler. Measurement of transport coefficients by an optical grating technique. *LNP Vol. 584: Thermal Nonequilibrium Phenomena in Fluid Mixtures*, 584:189–210, 2002.

[68] P. Blanco, H. Kriegs, M. P. Lettinga, P. Holmqvist, and S. Wiegand. Thermal diffusion of a stiff rod-like mutant Y21M-fd-virus. *Biomacromolecules*, 12:1602–1609, 2011.

[69] S. Wiegand, H. Ning, and H. Kriegs. Thermal diffusion forced rayleigh scattering setup optimized for aqueous mixtures. *J. Phys. Chem. B*, 111:14169–14174, 2007.

[70] K. F. Palmer and D. Williams. Optical-properties of water in near-infrared. *J. Opt. Soc. Am.*, 64:1107–1110, 1974.

[71] A. Becker, W. Köhler, and B. Müller. A scanning michelson interferometer for the measurement of the concentration and temperature derivative of the refractive- index of liquids. *Ber. Bunsen-Ges. Phys. Chem. Chem. Phys.*, 99:600–608, 1995.

[72] Rafael D. Camerini-Otero, Richard M. Franklin, and Loren A. Day. Molecular weights,

dispersion of refractive index increments, and dimensions from transmittance spectrophotometry. Bacteriophages R17, T7, and PM2, and tobacco mosaic virus. *Biochemistry*, 13:3763–3773, 1974.

[73] V. V. Sechenyh, J. C. Legros, and V. Shevtsova. Experimental and predicted refractive index properties in ternary mixtures of associated liquids. *J. Chem. Thermodyn.*, 43:1700–1707, 2011.

[74] NanoTemper Technologies GmbH. https://nanotempertech.com/.

[75] A. Leo, C. Hansch, and D. Elkins. Partition coefficients and their uses. *Chem. Rev.*, 71:525–616, 1971.

[76] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.*, 23:3–25, 1997.

[77] L. Mamy, D. Patureau, E. Barriuso, C. Bedos, F. Bessac, X. Louchart, F. Martin-Laurent, C. Miege, and P. Benoit. Prediction of the fate of organic compounds in the environment from their molecular properties: A review. *Critical Reviews in Environmental Science and Technology*, 45:1277–1377, 2015.

[78] W. L. Jorgensen. Free-energy calculations - a breakthrough for modeling organic-chemistry in solution. *Acc. Chem. Res.*, 22:184–189, 1989.

[79] U. Avico, E. C. Signoretti, and P. Zuccaro. Activity-coefficients and partition-coefficients as parameters for the evaluation of biological-activity - a linear relationship between log-gamma and log-p. *Farmaco-Edizione Scientifica*, 35:590–595, 1980.

[80] M. G. Gonzalez-Bagnoli. *Modeling the Thermal Diffusion Coefficients*. PhD thesis, 2004.

[81] M. G. Gonzalez-Bagnoli, A. A. Shapiro, and E. H. Stenby. Evaluation of the thermodynamic models for the thermal diffusion factor. *Philos. Mag.*, 83:2171–2183, 2003.

[82] R. C. Dougherty. Temperature and pressure dependence of hydrogen bond strength: A perturbation molecular orbital approach. *J. Chem. Phys.*, 109:7372–7378, 1998.

[83] S. Liese, M. Gensler, S. Krysiak, R. Schwarzl, A. Achazi, B. Paulus, T. Hugel, J. P. Rabe, and R. R. Netz. Hydration effects turn a highly stretched polymer from an entropic into an energetic spring. *ACS Nano*, 11:702–712, 2017.

[84] R. Kita, P. Polyakov, and S. Wiegand. Ludwig-Soret effect of poly(N-isopropyl-acrylamide): Temperature dependence study in monohydric alcohols. *Macromolecules*, 40:1638–1642, 2007.

[85] M. Sarter. private communication.

[86] M. Gonzalez, L. A. Bagatolli, I. Echabe, J. L. R. Arrondo, C. E. Argarana, C. R. Cantor, and G. D. Fidelio. Interaction of biotin with streptavidin - thermostability and conformational changes upon binding. *J. Biol. Chem.*, 272:11288–11294, 1997.

[87] J. N. Song, Y. L. Li, C. G. Ji, and J. Z. H. Zhang. Functional loop dynamics of the streptavidin-biotin complex. *Sci. Rep.*, 5, 2015.

[88] S. Meskers, J. M. Ruysschaert, and E. Goormaghtigh. Hydrogen-deuterium exchange of streptavidin and its complex with biotin studied by 2D-attenuated total reflection fourier transform infrared spectroscopy. *J. Am. Chem. Soc.*, 121:5115–5122, 1999.

[89] Z. Q. Zhong, X. S. Li, and Y. Zhao. Enhancing binding affinity by the cooperativity between host conformation and host-guest interactions. *J. Am. Chem. Soc.*, 133:8862–8865, 2011.

[90] R. Sugaya, B. A. Wolf, and R. Kita. Thermal diffusion of dextran in aqueous solutions in the absence and the presence of urea. *Biomacromolecules*, 7:435–440, 2006.

Appendix

SUPPORTING INFORMATION:
Accumulation of formamide in hydrothermal pores to form prebiotic nucleobases

Doreen Niether, Dzmitry Afanasenkau, Jan K.G. Dhont,
and Simone Wiegand

22$^{\text{nd}}$ February, 2016

Contents

1   Refractive index contrast measurements

The contrast factors, e.g. the change of the refractive index in dependence of temperature, T, and concentration, ω, are measured with an interferometer [1] and an Abbe refractometer (Anton Paar ABBEMAT RXA 158), respectively. For the calculation of S_{T} from the IR-TDFRS measurements, the contrast factors were interpolated from these measurement series for the correct temperatures and concentrations. $(\partial n/\partial T)_{p,\omega}$ is negative in the measured concentration and temperature range, the absolute value increases with higher formamide concentration and decreases with increasing temperature (Fig. 1).

Measurements of the refractive index were conducted for 11 concentrations ranging from pure water ($\omega = 0$) to pure formamide ($\omega = 1$) at 20, 35, 50 and 65°C. The measured values (Fig. 2) were fitted with a 2nd order polynomial and the slope of the resulting curves is $(\partial n/\partial \omega)_{p,T}$. The inset in Fig. 2 shows the behavior of the contrast factor: $(\partial n/\partial \omega)_{p,T}$ increases at higher formamide concentrations and decreases with rising temperature.
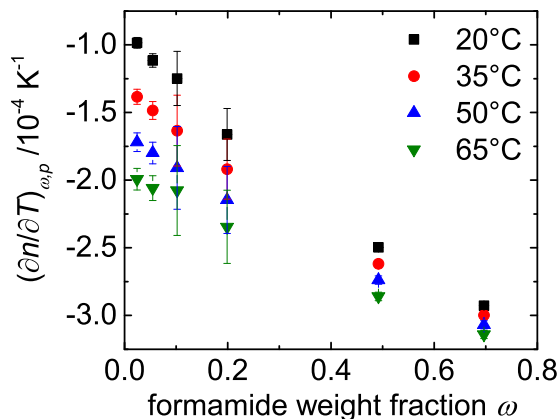

Figure 1: Results of the interferometric measurement of the contrast factor $(\partial n/\partial T)_{p,\omega}$ as function of concentration, ω.
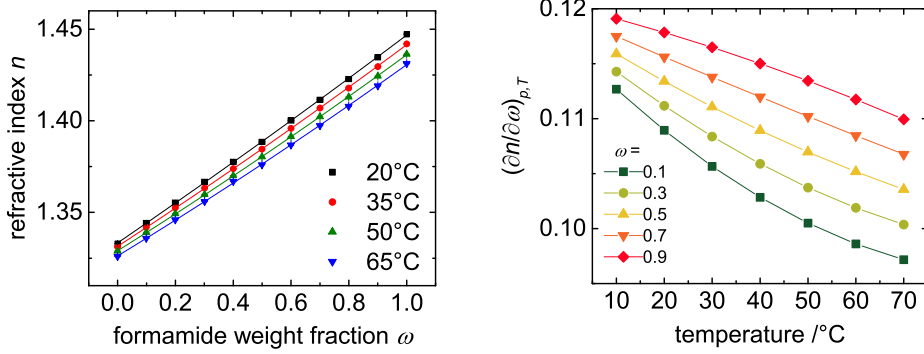
2

Figure 2: (left) Concentration dependence of refractive index n at different temperatures. (right) Contrast factors $(\partial n/\partial \omega)_{p,T}$ determined from a polynomial fit as function of temperature.

2 Temperature and concentration dependence of the various quantities used as an input to the numerical calculations

The temperature and concentration dependence of the thermo- and mass diffusion ceofficient are obtained from IR-TDFRS measurements [2], which are shown in Fig.2. For all investigated concentrations the mass diffusion coefficient shows only a slight decrease with increasing formamide concentration and increases significantly with rising temperatures, so that for the calculations we include only the temperature dependence and neglect the concentration dependence.

From a fit to the diffusion coefficients, D, in Fig.2 and of the Soret coefficients, S_{T}, in Fig.3 (main text) we obtain the following expressions, which are used for the calculations using the units given in Table 2,

$$
\begin{aligned}
D_T(T) &= 3.36833 \cdot 10^{-5} \\
&\quad + (1.13331 \cdot 10^{-5} - 3.36833 \cdot 10^{-5})/(1 + (T/38.42736)^{4.08469}) , \\
S_{\mathrm{T}}(\omega, T) &= (-2.55765 \cdot 10^{-4} + 2.34093 \cdot 10^{-5} \cdot T - 1.02968 \cdot 10^{-7} \cdot T^2) \\
&\quad + \frac{\omega^{(0.6916 - 0.01459 \cdot T + 2.4822 \cdot 10^{-4} \cdot T^2)}}{(317.62149 + 1.03712 \cdot T + 0.18358 \cdot T^2) + \omega^{(0.6916 - 0.01459 \cdot T + 2.4822 \cdot 10^{-4} \cdot T^2)}} .
\end{aligned}
$$

(1)

(2)

The specific mass density was measured over a range from 10 to 70°C for 5 mixtures with different weight fractions as well as pure water and formamide (see Table 1). Figure 4 compares the measured densities with the concentration dependent density at 25°C taken from literature [3], which agrees well
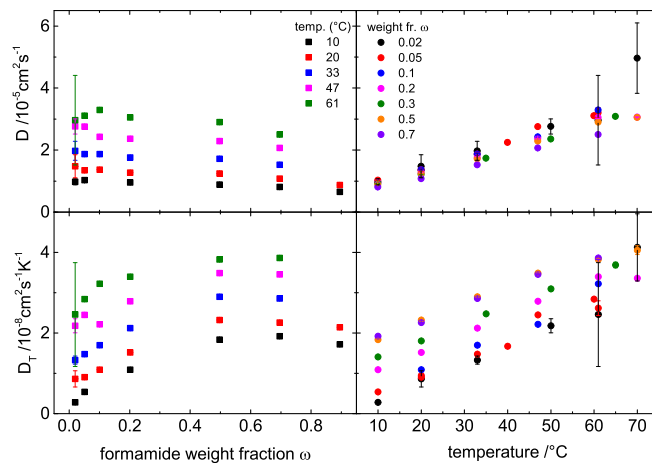
Figure 3: Results of IR-TDFRS measurements. The mass diffusion coefficient D and thermal diffusion coefficient D_T as functions of the formamide weight fraction (left) and temperature (right).
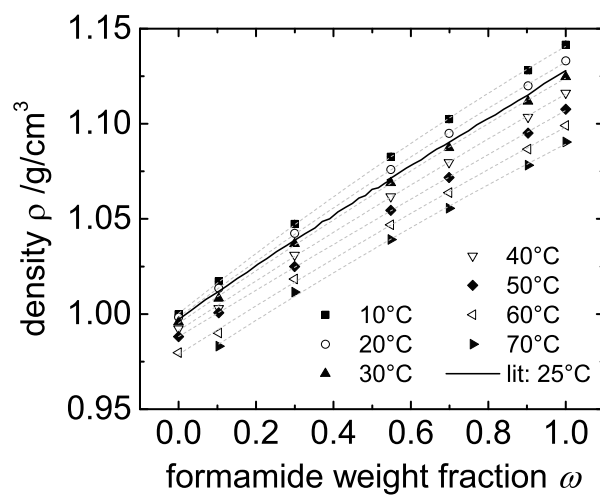

Figure 4: Plot of the density values given in TABLE 1. The solid line represents literature values at 25°C [3].

4

weight fraction ω	density ρ / g/cm^3						
	10°C	20°C	30°C	40°C	50°C	60°C	70°C
0.0000	0.99982	0.99832	0.99577	0.99236	0.98813	0.97965	-
0.1038	1.01727	1.01360	1.00824	1.00306	1.00065	0.99004	0.98309
0.2997	1.04737	1.04238	1.03694	1.03110	1.02490	1.01833	1.01142
0.5481	1.08271	1.07593	1.06900	1.06182	1.05446	1.046900	1.03916
0.6988	1.10244	1.09500	1.08740	1.07968	1.07181	1.06380	1.05563
0.9031	1.12822	1.12001	1.11178	1.10350	1.09514	1.08671	1.07825
1.0000	1.14151	1.13313	1.12465	1.11617	1.10765	1.09907	1.09044

Table 1: Density of formamide/water mixtures measured as a function of formamide weight fraction at temperatures from 10-70°C.

with our data. From the experimental data, the following expression for the mass density is obtained,

$$\rho(\omega, T) = 1.00409 - 8.88391 \cdot 10^{-5} \cdot T - 4.97372 \cdot 10^{-6} \cdot T^2 \qquad (3)$$
$$+ \left(0.03902 \cdot \exp\left(-\frac{T}{22.80603}\right) + 0.11532\right) \cdot \omega \,.$$

Experimental data for the shear viscosity, $\eta(\omega, T)$, of the mixture were taken from literature [4, 5],

$$\eta(\omega, T) = (0.03808 + 1.43946 \cdot 10^{-4} \cdot T) \cdot \exp\left(\frac{\omega}{0.20487 + 0.00189 \cdot T}\right) \quad (4)$$
$$+ (1.21076 - 0.01371 \cdot T) \,.$$

For the thermal conductivity, γ, the temperature dependence is negligible for both formamide [6] and water [7, 8]. The concentration dependence can be described by a linear function [9], and is accurately approximated by,

$$\gamma(\omega) = 0.5932 - 0.24653 \cdot \omega. \qquad (5)$$

Experimental data for the specific heat capacity, C_P, of the formamide/water mixture are taken from Ref.[10], and are given by,

$$C_P(\omega, T) = 4168.383 + (-2441.1484 + 13.23161 \cdot T) \cdot \omega + 339.319 \cdot \omega^2 \quad (6)$$

The above expressions for the temperature and concentration dependencies are implemented in the COMSOL software for the calculation of the concentration- and flow profiles within the pores.

5

T	ω	ρ	η	γ	C_p	S_{T}	D
°C	w.f.	g/cm³	mPa·s	W/(m·K)	J/(kg·K)	K⁻¹	cm²/s

Table 2: Units used for the temperature, concentration, density, dynamic viscosity, thermal conductivity, specific heat capacity, Soret coefficient and mass diffusion coefficient in the specified expressions.
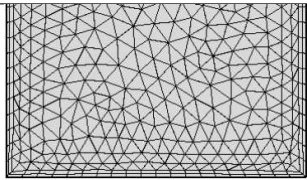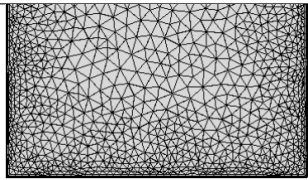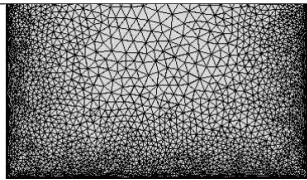
Mesh size	fine	finer	extra fine
Max. element size [m]	0.00117	8.14E-4	4.4E-4
Min. element size [m]	6.6E-6	2.75E-6	1.6E-6
Max. element growth rate	1.3	1.25	1.2
Curvature factor	0.3	0.25	0.25
Res. of narrow regions	1	1	1

Table 3: Mesh sizes used for the simulations.

3 Numerical calculations

To calculate the accumulation of formamide in a hydrothermal pore we solved a combination of Navier-Stokes-, heat transfer-, and thermodiffusion equations using a commercially available finite element software (COMSOL Multiphysiscs 5.1). We carefully checked the consistency of the stationary and time dependent solutions, by varying the mesh size and length of the time steps (the mesh sizes that were used are shown in Table 3). The *finer* and *extra fine* mesh gave identical results for the stationary states, while the *extra fine* mesh size was used for the time-dependent calculations in order to suppress an unphysical overshoot of the concentration for later times, just before the stationary state is reached.

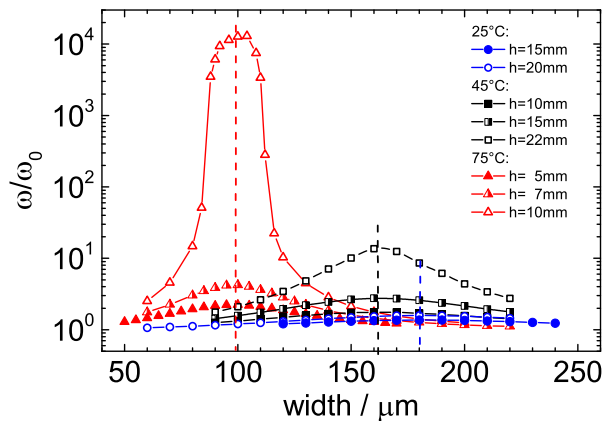The calculations were done for various aspect ratios of the pore at the

Figure 5: The accumulation-fold, ω/ω_0, vs. width of the pore for T equal to 25°C (the blue symbols), 45°C (black) and 75°C (red). Dashed lines mark the optimal widths of 180, 160 and 100 μm at the respective temperatures of $T = 25, 45$ and 75°C.

optimal pore width. The optimal pore width depends on temperature. We used three different average temperatures for our calculations (25°C, 45°C and 75°C) and determined the optimal pore widths (180, 160 and 100 μm) for each temperature (see fig.5). With increasing temperature the optimal width decreases. For the highest temperature, $T = 75$°C, we observe a strong decrease of the width, which might be related to the asymmetry of the velocity profile, which is discussed below.

Figure 6(left) shows the velocity profiles for the studied temperatures, Fig. 6(middle) the maximum velocity as function of the width and Fig. 6(right) the velocity versus the width of the pore. The latter contains information on the flow direction indicated by a sign change and illustrated by arrows in Fig. 6(left). Fig. 6(right) shows exemplary for one point how the maximum velocity is determined. The velocity profile in the pore is not dependent on width or height in the investigated range. At lower temperatures it is also independent of temperature. At very high temperatures the velocity profile shows an increased asymmetry with an up-flow stream that is much faster and narrower than the down-flow stream. The asymmetry is probably related to non-linear effects in the convective flow occurring for large temperature gradients and might also be responsible for the strong decrease of the optimal width at this high temperature. At $T = 75$°C compared to the low temperatures the maximum velocity increases especially for the larger widths
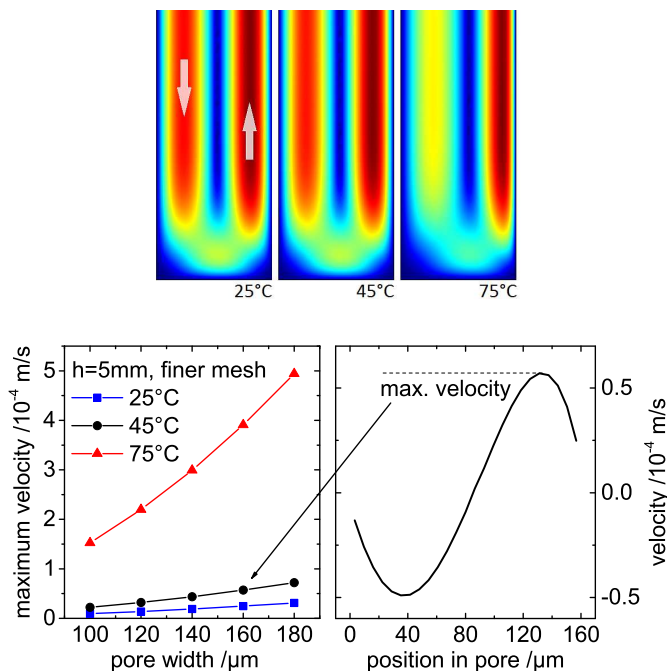
7

Figure 6: (left) Relative velocity profiles at $T =$ 25, 45 and 75°C. For each particular temperature dark blue and red refer to velocity=0 and highest velocity in the study, respectively. Note that the color code between different temperatures can not be compared. (middle) Maximum velocity vs. width at $T =$ 25, 45 and 75°C. (right) Velocity profile in the pore with added information about flow direction.

and leads to more mixing due to convection, therefore the width needs to be reduced to achieve large accumulation-folds.

Additonally we investigated how the accumulation-fold depends on the temperature difference, ΔT, across the pore. Figure 7(left) shows the accumulation-fold at $T = 75$°C as function of the aspect ratio for three different temperature differences across the pore. As expected a larger ΔT leads to a stronger accumulation, so that the plateau is reached at smaller aspect ratios.

Figure 7(right) shows the time dependence of the accumulation for different mean temperatures. It turns out that the time to reach the plateau, τ_{plateau} increases with increasing temperature. The physical reason for this behavior is that the convection process becomes stronger (fig. 6) due to a decreasing viscosity with increasing temperature and leads to a stronger mixing so that a longer time is required to reach the plateau.

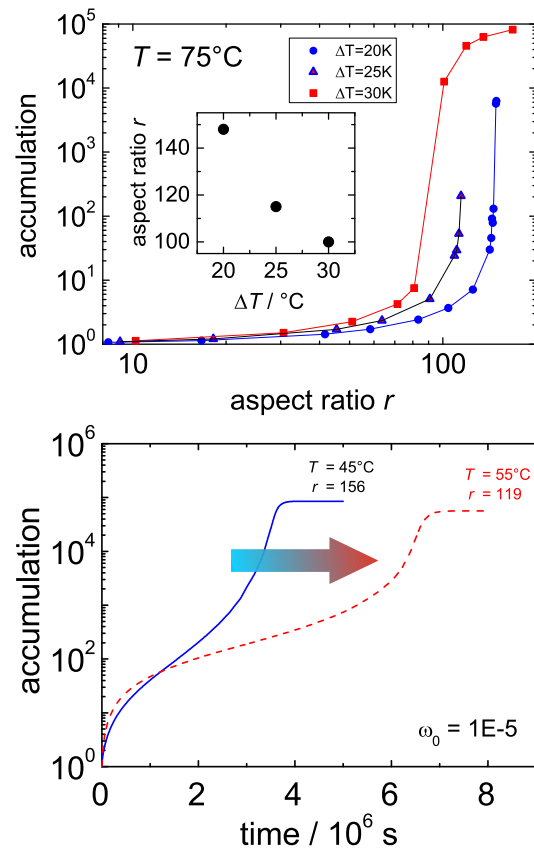Figure 7: (left)The accumulation-fold vs. aspect ratio at $T = 75°$C for three different temperature differences ΔT. The inset shows the steep decrease of the aspect ratio, r, with increasing temperature difference, ΔT. (right) Accumulation as a function of time for two mean temperatures at an aspect ratio close to the plateau. The initial concentration was $\omega_0 = 1 \cdot 10^{-5}$.

9

References

[1] A. Becker, W. Köhler, and B. Müller. A scanning michelson interferometer for the measurement of the concentration and temperature derivative of the refractive- index of liquids. *Ber. Bunsen-Ges. Phys. Chem. Chem. Phys.*, 99(4):600–608, 1995.

[2] S. Wiegand, H. Ning, and H. Kriegs. Thermal diffusion forced rayleigh scattering setup optimized for aqueous mixtures. *J. Phys. Chem. B*, 111:14169–14174, 2007.

[3] E. P. Egan and B. B. Luff. Heat of solution heat capacity and density of aqueous formamide solutions at 25 degrees c. *J. Chem. Eng. Data*, 11, 1966.

[4] S. Akhtar, A. N. M. O. Faruk, and M. A. Saleh. Viscosity of aqueous solutions of formamide, n-methylformamide and n,n-dimethylformamide. *Phys. Chem. Liq.*, 39(3):383–399, 2001.

[5] *Landolt-Bönstein New Series IV/25*, chapter Viscosity of the mixture (1)water, (2) formamide. Springer, 2008.

[6] Yu. A. Ganiev and Yu. L. Rastorguev. Thermal conductivity of organic liquids. *J. Eng. Phys.*, 15:519–525, 1968.

[7] Hugh D. Young. *University Physics*. Addison Wesley, 7th ed. edition, 1992.

[8] M. N. Hill, editor. *The Sea*, volume 1. Wiley, 1962.

[9] A. Tobitani and T. Tanaka. Predicting thermal-conductivity of binary-liquid mixtures on basis of coordination-number. *Can. J. Chem. Eng.*, 65(2):321–328, 1987.

[10] R. F. Checoni and P. L. O. Volpe. Measurements of the molar heat capacities and excess molar heat capacities for water plus organic solvents mixtures at 288.15 k to 303.15 k and atmospheric pressure. *J. Solution Chem.*, 39(2):259–276, 2010.

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

Doreen Niether, Silvia Di Lecce, Fernando Bresme,
and Simone Wiegand

October 18, 2017

Contents
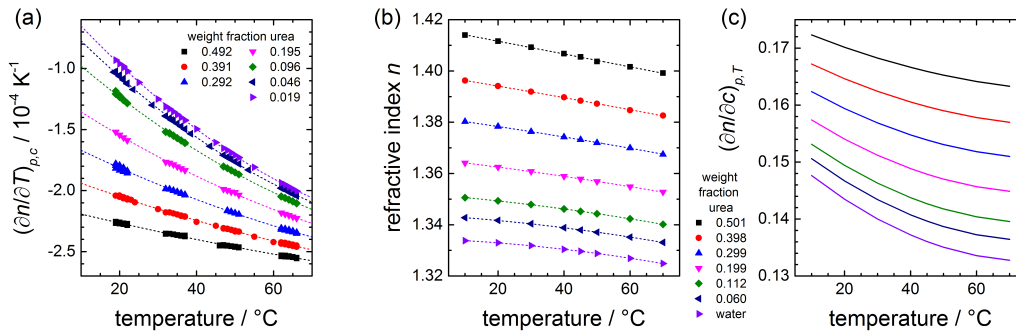
1

1 Refractive index contrast measurements

Figure 1: (a) Results of the interferometric measurement of the contrast factor $(\partial n/\partial T)_{p,c}$ for different concentrations of urea in water. (b) Refractive index n for different urea concentrations as a function of temperature. (c) Contrast factor $(\partial n/\partial c)_{p,T}$ calculated from refractive index measurements.

In order to calculate the Soret coefficient S_T from the intensity of the diffracted read-out beam, it is necessary to know the dependence of the refractive index on temperature and concentration, $(\partial n/\partial T)_{p,c}$ and $(\partial n/\partial c)_{p,T}$. The former was measured interferometrically [1] and the latter with an Abbe refractometer (Anton Paar ABBEMAT RXA 158). For the calculation of S_T from the IR-TDFRS measurements, the contrast factors were interpolated from these measurement series for the correct temperatures and concentrations. $(\partial n/\partial T)_{p,c}$ is negative in the measured concentration and temperature range, the absolute value increases with higher urea concentration and with increasing temperature (Fig. 1a). Measurements of the refractive index were conducted for 7 concentrations ranging from pure water ($c = 0$) to an aqueous solution with 50 wt% urea (Fig. 1b). Fig. 1c shows the derivative $(\partial n/\partial c)_{p,T}$ calculated from the function $n(c, T)$ we determined by fitting the experimental values. $(\partial n/\partial c)_{p,T}$ increases at higher urea concentrations and decreases with rising temperature.

2 Denaturation of proteins using urea and formamide

The denaturation of proteins by addition of organic solvents has often been investigated [2, 3, 4, 5, 6, 7, 8, 9, 10]. The question is whether, as in the urea, the denaturation of proteins by formamide occurs in the same concentration range where we observe the changes in the temperature dependent slope of the Soret coefficient in the formamide/water solutions. For formamide (FA) solutions this change occurs for a weight fraction of 0.2 FA corresponding to 4.6 mol/L or a volume fraction of 0.18. Khabiri et al. [9] investigated the influence of formamide 5% (v/v), acetone 20% (v/v) and isopropanol 10% (v/v) on the structure of the haloalkane dehalogenases DhaA, LinB, and DbjA. With the exception of LinB in acetone, the structures of studied enzymes were stabilized in water-miscible organic solvents. The volume fraction of 5% is well below the concentration at which the temperature dependent slope changes, so that we do not expect an effect. Asakura et al. [10] did not observe a denaturation of hemoglobin in the presence of formamide even at very high formamide concentration. Fuchs et al. [7] determine only change of the melting curve. In fig. 2 we see the Soret coefficient as function of concentration. The inversion of the slope can be observed as an 'intersection point', where the T-dependence of S_{T} is close to zero. If we compare the formamide and the urea results, that concentration is w.f. = 0.2 and w.f. = 0.3 for formamide and urea, respectively. In both cases we find a correlation with the denaturation range of the two compounds.

Table 1: UREA

Molarity	weight fraction	reference
4-5	0.23-0.28	[11]
5-6	0.28-0.33	[12]
> 5	> 0.28	[13]
5	0.28	[14]
6	0.33	[15]
5-6	0.28-0.33	[16]
4-6	0.23-0.33	[17]
5.2	0.29	average

Table 2: FORMAMIDE

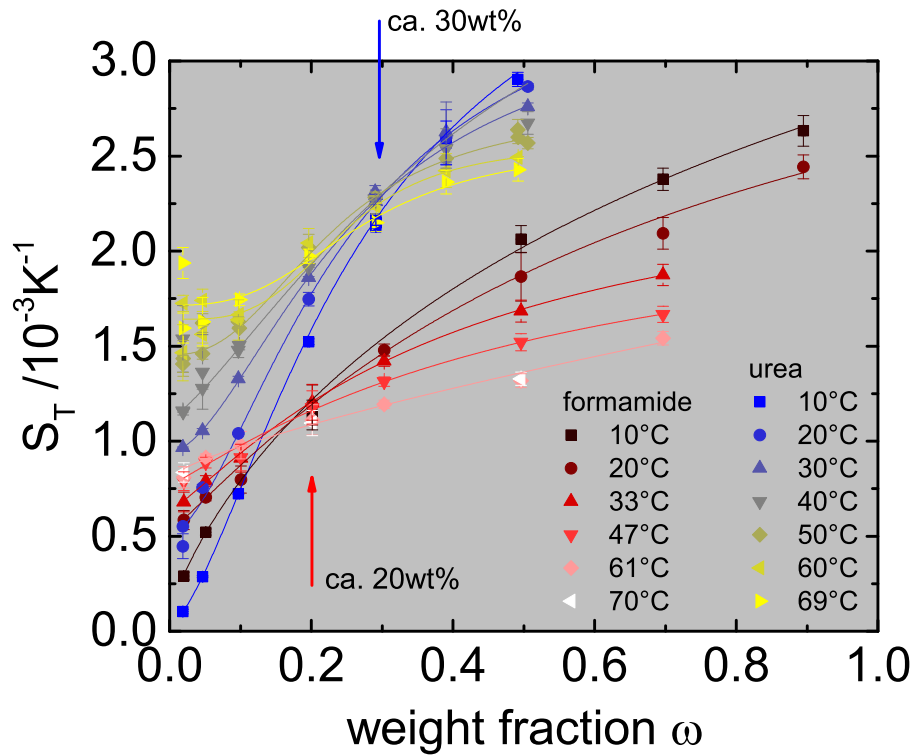| Molarity	weight fraction	reference
1-2	0.05-0.09	[4]
5.9	0.26	[5]
9.3	0.40	[5]
5-6	0.22-0.26	[8]
4.9	0.21	average

Figure 2: Soret coefficient S_T against concentration for urea and formamide in water. The intersection points at w.f. = 0.3 and w.f. = 0.2 show the concentration where the T-dependence of S_T is inverted.

3 Density of aqueous urea solutions

Densities were measured with an Anton Paar DMA 4500 densimeter with an error of 0.0002 g/cm^3. Solutions were prepared with Urea ($\geq$ 99%, Fluka, Sigma-Aldrich, 89555 Steinheim, Germany) and Millipore water. For degassing, the sample solutions were sealed in their flasks with Parafilm and kept in an ultra-sonic bath at 70°C for 2h. No formation of bubbles was observed during the measurements, except for the 5 wt% solution at 60 and 65°C.
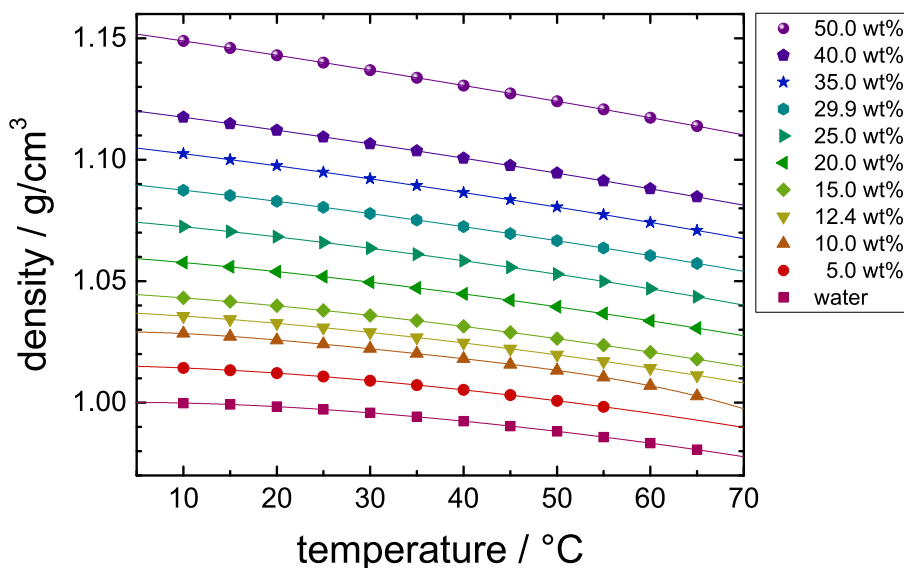
Figure 3: Density ρ against temperature for the measured urea solutions and water. Values can also be found in Table 3.

The following empirical equation (1) was found by fitting the experimental data and can be used to interpolate the temperature and concentration dependant density of aqueous urea solutions. Note that the units of ρ, ω, and T are g/cm^3, weight fractions, and °C, respectively.

$$\rho(\omega, \mathrm{T}) = (0.99965 + 0.30714 \cdot \omega) + (5.92438 \cdot 10^{-5} - 0.00187 \cdot \omega + 0.00141 \cdot \omega^2) \cdot \mathrm{T} +$$
$$(-8.31302 \cdot 10^{-6} + 2.38684 \cdot 10^{-5} \cdot \omega - 2.5325 \cdot 10^{-5} \cdot \omega^2) \cdot \mathrm{T}^2 +$$
$$(5.34609 \cdot 10^{-8} - 7.13453 \cdot 10^{-8} \cdot \omega) \cdot \mathrm{T}^3 + (-2.52222 \cdot 10^{-10} + 2.08659 \cdot 10^{-10} \cdot \omega) \cdot \mathrm{T}^4$$
$$(1)$$

5

Table 3: Measurement results for the density ρ of aqueous urea solutions with weight fractions of urea from w.f. $= 0$ to w.f. $= 0.5$ and at temperatures between 10 and 65°C.

| w.f. | ρ / g/cm^3 | | | | | |
	10°C	15°C	20°C	25°C	30°C	35°C
0	0.9998	0.9992	0.9983	0.9972	0.9958	0.9942
0.0503	1.0143	1.0133	1.0121	1.0107	1.0091	1.0072
0.0997	1.0285	1.0272	1.0257	1.0241	1.0222	1.0202
0.1242	1.0356	1.0343	1.0326	1.0309	1.0289	1.0268
0.1497	1.0431	1.0416	1.0399	1.0380	1.0359	1.0337
0.1999	1.0577	1.0559	1.0539	1.0518	1.0496	1.0472
0.2497	1.0724	1.0704	1.0682	1.0660	1.0635	1.0610
0.2991	1.0875	1.0853	1.0829	1.0804	1.0779	1.0752
0.3500	1.1025	1.1001	1.0975	1.0949	1.0922	1.0894
0.4001	1.1175	1.1149	1.1122	1.1094	1.1066	1.1036
0.5001	1.1489	1.1460	1.1430	1.1400	1.1369	1.1338
	40°C	45°C	50°C	55°C	60°C	65°C
0	0.9924	0.9903	0.9882	0.9858	0.9833	0.9806
0.0503	1.0053	1.0031	1.0008	0.9983	–	–
0.0997	1.0180	1.0157	1.0133	1.0104	1.0070	1.0027
0.1242	1.0246	1.0222	1.0196	1.0170	1.0142	1.0112
0.1497	1.0314	1.0289	1.0263	1.0236	1.0208	1.0179
0.1999	1.0448	1.0422	1.0394	1.0366	1.0337	1.0307
0.2497	1.0584	1.0557	1.0529	1.0499	1.0468	1.0436
0.2991	1.0724	1.0696	1.0667	1.0637	1.0605	1.0574
0.3500	1.0865	1.0835	1.0805	1.0774	1.0742	1.0709
0.4001	1.1007	1.0976	1.0945	1.0913	1.0880	1.0847
0.5001	1.1306	1.1273	1.1241	1.1207	1.1173	1.1138

4   Concentration scales and conversion for aqueous urea solutions

The weight fraction of urea ω_{urea} is the ratio of the mass of urea m_{urea} and the total mass of the solution $(m_{urea} + m_{water})$:

$$\omega_{urea} = m_{urea}/(m_{urea} + m_{water}) \tag{2}$$

The mole fraction of urea χ_{urea} is defined as the ratio of the number of urea molecules N_{urea} and the total number of molecules $(N_{urea} + N_{water})$:

$$\chi_{urea} = N_{urea}/(N_{urea} + N_{water}) \tag{3}$$

With $N = m/M$, $\omega_{water} = 1 - \omega_{urea}$, and the molar masses for urea and water, $M_{urea} =$60.05526 g/mol and $M_{water} =$18.01528 g/mol, mole fractions can be calculated with

$$\chi_{urea} = (\omega_{urea}/M_{urea})/[(\omega_{urea}/M_{urea}) + (\omega_{water}/M_{water})]. \tag{4}$$

The number of water molecules per urea molecule can be calculated from χ with $N_{urea} = 1$ and

$$N_{water} = (N_{urea}/\chi_{urea}) - N_{urea}. \tag{5}$$

The molar concentration c is defined as the number of urea molecules in the volume of the solute

$$c_{urea} = N_{urea}/V, \tag{6}$$

where the volume V is given by

$$V = (m_{urea} + m_{water})/\rho. \tag{7}$$

In our case, where the density ρ is given in g/cm^3, c in the unit M = mol/L can be calculated from ω with

$$c_{urea} = (\omega_{urea}/M_{urea}) \cdot \rho \cdot 1000. \tag{8}$$

Note that the maximal value for the molar concentration at a mole fraction of $\chi_{urea} = 1$ is calculated as $c_{urea} =$21.56 M, which, considering that the density used for the calculation is extrapolated from the aqueous solution, is in reasonable agreement with the value of 21.98 M calculated for pure solid urea with a density of $\rho_{urea} = 1.32$ g/cm^3.

Table 4: Conversion table for concentrations of aqueous urea solutions in weight fraction ω, mole fraction χ, molar concentration c and molecular ratio urea:water in a concentration range from $\omega = 0$ to $\omega = 0.5$.

χ	ω	c / M	urea:water		ω	χ	c / M	urea:water
0.01	0.03258	0.54622	1:99		0.05	0.01554	0.84306	1:63.34
0.05	0.14926	2.58461	1:19		0.1	0.03226	1.70936	1:30
0.1	0.27028	4.83616	1:9		0.2	0.06976	3.51320	1:13.33
0.15	0.37039	6.80652	1:5.67		0.3	0.11392	5.41460	1:7.78
0.2	0.45456	8.54050	1:4		0.4	0.16666	7.41662	1:5
0.25	0.52633	10.07544	1:3		0.5	0.23076	9.52234	1:3.33
c / M	χ	ω	urea:water		c / M	χ	ω	urea:water
1	0.0185	0.0592	1:53.01		6	0.1286	0.3298	1:6.77
2	0.0381	0.1165	1:25.28		7	0.1551	0.3796	1:5.45
3	0.0587	0.1721	1:16.04		8	0.1834	0.4282	1:4.45
4	0.0806	0.2261	1:11.41		9	0.2139	0.4757	1:3.67
5	0.1039	0.2787	1:8.63		10	0.2468	0.5221	1:3.05

5 Data from computer simulations

5.1 Comparison of simulation and experiment

To test the accuracy of our forcefields for urea-water solutions we performed a number of simulations and computed the density and diffusion coefficient as a function of urea concentration. Figure 4 shows the concentration dependence of the density with urea concentration at 25^0C. Both experiments and simulations shows an increase of the solution density with the urea concentration. Our computed values are in excellent agreement with both experimental [18] and simulated [19] results obtained at the same thermodynamic state, in the latter case using a different force-field.

Figure 5 shows the diffusion coefficient D for several concentrations as a function of temperature. The simulation data for the diffusion coefficient were calculated using equilibrium simulations ensemble and via the mean square displacement and Einstein equation [20].

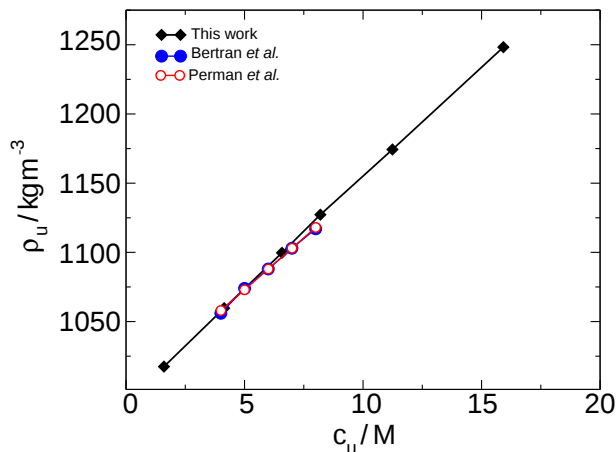Figure 4: Concentration dependence of density for urea solutions at 25^0C. The simulation results are compared with both experimental [18] and simulation [19] data using a different forcefield. All the data correspond to the same thermodynamic state.
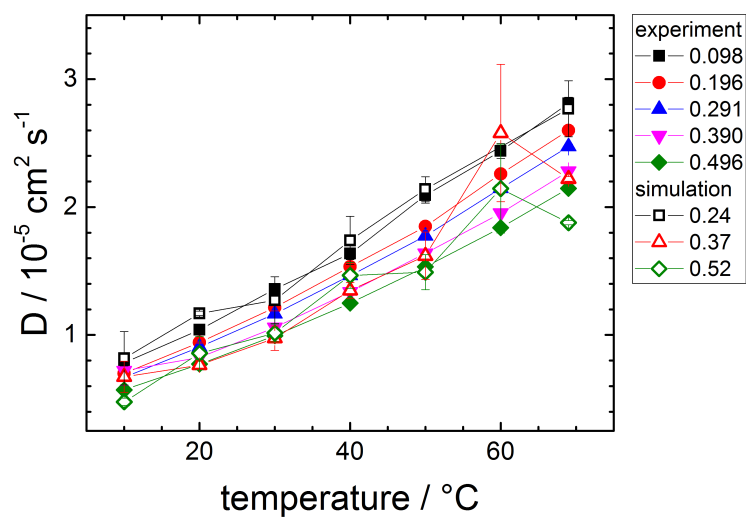
Figure 5: Diffusion coefficient D against temperature for different concentrations from IR-TDFRS measurements (lines) and simulations (symbols).

5.2   Temperature and mole fraction profiles

In order to compute the Soret coefficient we performed Non-Equilibrium Molecular Dynamics (NEMD) simulations as described in the Methods section.
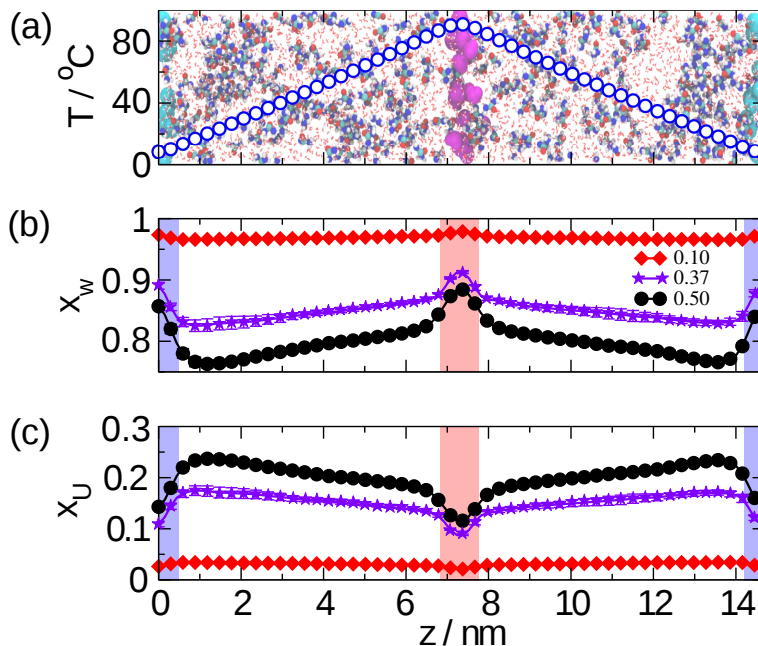

Figure 6: Temperature profile (a) water mole fraction (b) and urea mole fraction (c) along the simulated box for different urea concentrations. The red and blue squares in panel (b) and (c) indicate the position of the hot and cold thermostatted layers, respectively.

Figure 6 (a) show a typical snapshot of the system simulated in this work, along with a representative temperature profile. The hot and cold thermostatted molecules are represented in magenta and cyan, respectively. The temperatures in the cold and hot layers were set to 2^0C and 102^0C, respectively, and result in a well defined temperature profile along the simulated box. The size of the water molecules, in red, was decreases to allow a better visualization of the urea molecules. Figure 6 (b) and (c) shows the mole fraction of water and urea, respectively, for different urea concentrations. Both, the mole fraction of urea and water changes along the simulated box and with temperature as a results of the thermal gradient applied. The data were acquired at the stationary state, *i.e* when there is no net mass flux. The variation in mole fraction along the simulated box was used to compute the Soret coefficient (see equation (4) in the Methods section). We neglected the data next to the thermostatted layers.

11

5.3 Soret coefficient and chemical potential

Table 5 shows the computed Soret coefficient values for the three concentrations considered in this work (concentrations 10 wt% (1.7 M), 37 wt% (6.7 M) and 50 wt% (9.4 M), shown in fig. 1b of the main text), along with the fitting parameters used to model our data using the Iacopini and Piazza equation [21] (equation 2 in the main text) or the expression $S_T = S_T^\infty + S_T^0 \cdot \exp(-T/T_0)$ in case of negative temperature dependence (0.50 - U1). The labels U1 and U2 in the table, indicate the different water-urea interactions used in this work (see the main text for more details). The chemical potential of urea was computed as described in the Methods section and then fitted to a linear equation ($\mu_{tot}(T) = mT + q$) to calculate the entropy and enthalpy. The fitting parameters are reported in the last column of Table 5. The fitting of the chemical potential was obtained using the Kelvin scale, while that one for the Soret coefficient using the Celsius scale. Note that the values of the Soret coefficient are independent of the scale used.

Table 5: Soret coefficient S_T for different concentrations of water computed via NEMD simulations. In order to describe the temperature dependence of S_T the adjustable parameter sets S_T^∞, T^* and T_0 or S_T^∞, S_T^0 and T_0 are used. Additionally, the adjustable parameters m and q are listed to describe the temperature dependence of the chemical potential of urea in the water-urea solutions.

| conc. $w.f.$	T / $^\circ$C	S_T / $10^{-3}\mathrm{K}^{-1}$	S_T^∞ / $10^{-3\circ}\mathrm{C}^{-1}$	T^* / $^\circ$C	T_0 / $^\circ$C
0.1	29.63	0.236 ± 0.358	4.632	29.437	22.765
	40.01	0.966 ± 0.229			
	50.28	3.624 ± 0.386			
	60.52	3.326 ± 0.379			
	70.85	3.719 ± 0.391			
0.37	22.52	2.925 ± 0.332	4.588	15.734	6.6907
	34.48	4.296 ± 0.487			
	46.64	4.625 ± 0.539			
	59.11	4.380 ± 0.516			
	72.01	4.718 ± 0.596			
$0.50 - \mathrm{U2}$	27.24	2.006 ± 0.194	3.065	16.102	10.3838
	38.09	2.826 ± 0.300			
	48.93	2.881 ± 0.213			
	59.96	2.872 ± 0.220			
	71.27	3.246 ± 0.252			
			S_T^∞ / $10^{-3\circ}\mathrm{C}^{-1}$	S_T^0 / $10^{-3\circ}\mathrm{C}^{-1}$	T_0 / $^\circ$C
$0.50 - \mathrm{U1}$	26.66	6.435 ± 0.711	3.676	8.753	23.4695
	37.28	5.571 ± 0.624			
	48.07	4.938 ± 0.561			
	59.10	3.948 ± 0.464			
	70.41	4.354 ± 0.526			
conc. $w.f.$				m / kJmol^{-1} K^{-1}	q / kJmol^{-1}
0.1				0.112	-73.624
0.37				0.136	-78.071
$0.50 - \mathrm{U1}$				0.139	-78.522
$0.50 - \mathrm{U2}$				0.142	-82.505

5.4 Aggregation of urea molecules
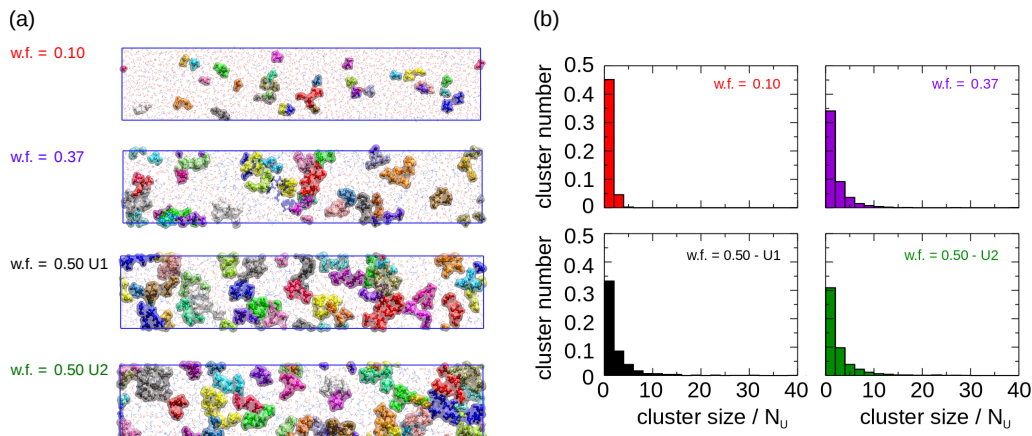
Figure 7: (a) Snapshot of urea water mixture showing the aggregation of urea molecules at different concentrations. Clusters with different size are represented with different colors. (b) Cluster size distribution computed through the average of 100 configuration over 15 ns.

To explain the thermodiffusive behaviour of the aqueous urea solutions we quantified the clustering of urea molecules. We performed a 3D Voronoi tessellation of the solutions at 30^0C using the voro++ package [22]. Each urea molecule has been represented according to the position of the C=O group which has been considered as the center of the molecule to construct the cluster.

Figure 7 illustrates the urea aggregation as a function of the urea concentration. In the snapshots on the left side of Figure 7 (a), clusters with different size are represented with different colors. Water and urea molecules not aggregated are represented with lines. Urea molecules aggregate creating an amount of clusters which increases, as well as their size, with concentration. Furthermore, we show, in Figure 7 (b), the cluster size distribution representing the normalized number of clusters against the cluster size, in terms of number of urea molecules, N_U. This graphical representation of the numerical data was obtained by computing the average cluster number and size of 100 configurations over 15 ns. At the highest concentration considered, 50 wt%, we find that the increase in water-urea interactions (U2) results in a decrease of both the size and number of cluster. Indeed, we quantified the aggregation by computing the mean size of the clusters, $< S_{U,cluster} >$, using:

$$< S_{U,cluster} >= \int_{s_{min}}^{s_{max}} s n_s(s) ds, \tag{9}$$

14

where $n_s(s)$ represent the number of cluster with size s, and s_{min} and s_{max} are the lowest and highest cluster size observed for these simulated systems, respectively.

Table 6: Mean size of the clusters, $< S_{U,cluster} >$ as a function of different weight fractions and water-urea interactions.

force field	w.f.	$< S_{U,cluster} > / N_u$
U1	0.10	2.20
U1	0.37	3.16
U1	0.50	4.83
U2	0.50	3.76

Table 5 shows the mean size of the clusters, $< S_{U,cluster} >$ for the different simulated systems. $< S_{U,cluster} >$ decreases with the increase of the water-urea interaction (see w.f. = 0.5 U1 and U2), loosing about one urea molecule per cluster and being closer to the cluster size obtained at lower concentration, namely w.f. = 0.37. At this concentration the Soret coefficient increases with temperature. We find that the increase in the water-urea interactions results in a reduction of clustering, hence better solubility in water, which correlates well with our hypothesis that stronger solvation leads to a less thermophobic state.

References

[1] A. Becker, W. Köhler, and B. Müller. *Ber. Bunsen. Phys. Chem.*, 99(4):600–608, 1995.

[2] T. Samejima and K. Shibata. *Archives of Biochemistry and Biophysics*, 93:407, 1961.

[3] L. Casiraghi G. Invernizzi, R. Grandori, and M. Lotti. *Journal of Biotechnology*, 141:42–46, 2009.

[4] R. D. Blake and S. G. Delcourt. *Nucleic Acids Research*, 24:2095–2103, 1996.

[5] Y. L. Khmelnitsky, V. V. Mozhaev, A. B. Belova, M. V. Sergeeva, and K. Martinek. *European Journal of Biochemistry*, 198:31–41, 1991.

[6] C. Tibbetts, K. Johansson, and L. Philipson. *Journal of Virology*, 12:218–225, 1973.

[7] J. Fuchs, D. Dell'Atti, A. Buhot, R. Calemczuk, M. Mascini, and T. Livache. *Analytical Biochemistry*, 397:132–134, 2010.

[8] L. S. Yilmaz, A. Loy, E. S. Wright, M. Wagner, and D. R. Noguera. *Plos One*, 7(8), 2012.

[9] M. Khabiri, B. Minofar, J. Brezovsky, J. Damborsky, and R. Ettrich. *Journal of Molecular Modeling*, 19:4701–4711, 2013.

[10] T. Asakura, K. Adachi, and E. Schwartz. *Journal of Biological Chemistry*, 253:6423–6425, 1978.

[11] M. C. Stumpe and H. Grubmüller. *J. Phys. Chem. B*, 111:6220–6228, 2007.

[12] N. Samanta, D. Das Mahanta, and R. K. Mitra. *Chemistry-an Asian Journal*, 9:3457–3463, 2014.

[13] L. B. Sagle, Y. J. Zhang, V. A. Litosh, X. Chen, Y. Cho, and P. S. Cremer. *J. Am. Chem. Soc.*, 131:9304–9310, 2009.

[14] A. Idrissi, M. Gerard, P. Damay, M. Kiselev, Y. Puhovsky, E. Cinar, P. Lagant, and G. Vergoten. *J. Phys. Chem. B*, 114:4731–4738, 2010.

[15] Y. Hayashi, Y. Katsumoto, S. Omori, N. Kishii, and A. Yasuda. *J. Phys. Chem. B*, 111:1076–1080, 2007.

[16] D. Bandyopadhyay, S. Mohan, S. K. Ghosh, and N. Choudhury. *J. Phys. Chem. B*, 118:11757–11768, 2014.

[17] A. Tischer and M. Auton. *Protein Science*, 22:1147, 2013.

[18] Edgar Philip Perman and Trevor Lovett. *Trans. Faraday Soc.*, 22:1–19, 1926.

[19] Celso A. Bertran, José J. V. Cirino, and Luiz C. G. Freitas. *J. Braz. Chem. Soc.*, 13:238 – 244, 00 2002.

[20] M.P. Allen and D.J. Tildesley. *Computer Simulation of Liquids*. Oxford Science Publ. Clarendon Press, 1989.

[21] S. Iacopini and R. Piazza. *Europhys. Lett.*, 63:247–253, 2003.

[22] Chris H. Rycroft. *Chaos*, 19:041111, 2009.

SUPPORTING INFORMATION:
Role of Hydrogen Bonding of Cyclodextrin-Drug Complexes Probed by Thermodiffusion

Doreen Niether, Tsubasa Kawaguchi, Jana Hovancová, Kazuya Eguchi, Jan K.G. Dhont, Rio Kita, and Simone Wiegand
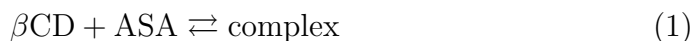
August 16, 2017

Contents

1 Determination of the association constant by NMR self diffusion coefficient measurements

The complex formation between β-cyclodextrin (β-CD) and acetylsalicylic acid (ASA) is an equilibrium reaction

$$\beta CD + ASA \rightleftarrows complex \tag{1}$$

with the association coefficient K_a defined as

$$K_a = \frac{[\text{complex}]}{[\beta CD_{\text{free}}][ASA_{\text{free}}]} = \frac{[\text{complex}]}{x_{\text{free}}{}^2} \tag{2}$$

Note that the last part of the equation is only valid in an equimolar system. The amount of free compound x_{free} is

$$x_{\text{free}} = (1 - p_{\text{com.}}) \cdot x_0 \tag{3}$$

where $p_{\text{com.}}$ is the fraction of compound bound in complexes relative to its total amount x_0:

$$p_{\text{com.}} = \frac{[\text{complex}]}{x_0} \tag{4}$$

The fraction $p_{\text{com.}}$ can also be determined by comparing the self diffusion coefficients D_{ASA} and $D_{\beta CD}$ measured in the binary solutions of ASA/D_2O and β-CD/D_2O with the corresponding self diffusion coefficients $D_{\text{ASA,mix}}$ and $D_{\beta CD,\text{mix}}$ observed in the ternary mixture β-CD/ASA/D_2O. In the ternary mixture, we observe an averaged diffusion coefficient ($D_{\beta CD,\text{mix}}$ and $D_{\text{ASA,mix}}$) determined by the diffusion coefficients of free compounds (D_{ASA} and $D_{\beta CD}$) and complexes ($D_{\text{com.}}$). We can write

$$D_{\beta CD,\text{mix}} = (1 - p_{\text{com.},\beta CD}) D_{\beta CD} + p_{\text{com.},\beta CD} \cdot D_{\text{com.}} \tag{5}$$
$$D_{\text{ASA,mix}} = (1 - p_{\text{com.,ASA}}) D_{\text{ASA}} + p_{\text{com.,ASA}} \cdot D_{\text{com.}}$$

For our equimolar system the fractions of β-CD and ASA, $p_{\text{com.},\beta CD}$ and $p_{\text{com.,ASA}}$, respectively, are equal to $p_{\text{com.}}$ and we can solve the linear eqs. 5 for $p_{\text{com.}}$:

$$p_{\text{com.}} = 1 - \frac{D_{\beta CD,\text{mix}} - D_{\text{ASA,mix}}}{D_{\beta CD} - D_{\text{ASA}}} \tag{6}$$

Plugging eq. 4 and 3 into eq. 2 we get

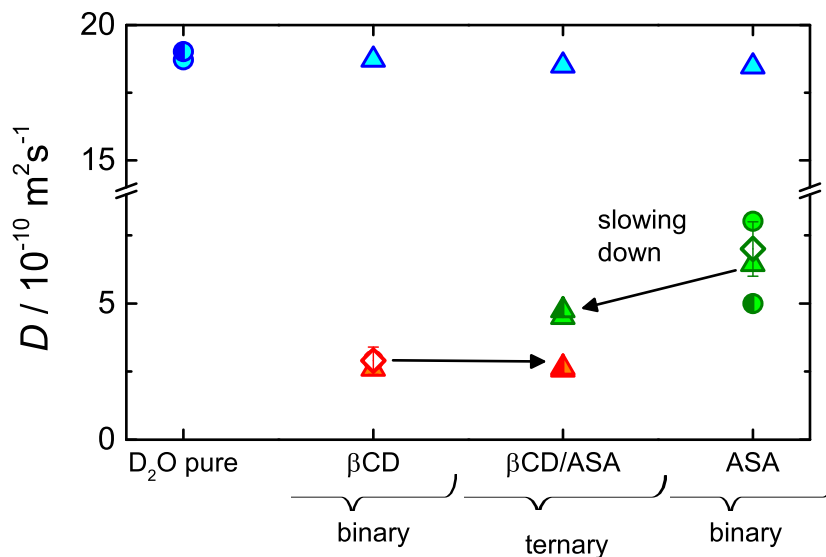$$K_a = \frac{p_{\text{com.}}}{x_0 \cdot (1 - p_{\text{com.}})^2} \tag{7}$$

2

Figure 1: Self diffusion coefficients of D_2O (blue symbol), β-CD (1.0 wt%, red symbol) and ASA (1.0 wt%, green symbol) at $T = 25°$C measured by NMR (triangles) in the binary mixtures of β-CD/D_2O (1.0 wt%), ASA/D_2O (1.0 wt%) and in the ternary mixture β-CD/ASA/D_2O. The self diffusion coefficients determined for a diffusion time of 5 ms (fully filled symbol) and 1000 ms (half filled) agree within the error bars. For comparison we show the collective diffusion coefficients (open diamond) determined in the TDFRS experiment in H_2O, which agree also within their error bars with the self diffusion coefficients determined by NMR. Additionally we display literature values of the self diffusion coefficient of H_2O (blue circle) in D_2O determined by tracer diffusion (half filled blue circle) and diaphragma method (filled blue circle) [1], of β-CD in H_2O (filled red circle) [2] and of ASA in H_2O (filled green circle) and (half filled circle) by diffusion cell measurements [3] and a DOSY experiment [4], respectively. Note that the self diffusion of ASA significantly slows down by 40 % in the ternary mixture indicating the complex formation with β-CD.

Figure 1 shows the measured self diffusion coefficients of D_2O (blue triangles), ASA (green triangles) and β-CD (red triangles) at $T = 25°$C by pulsed field gradient NMR. While the diffusion of water D_2O agrees within 2% with the value found for the self diffusion coefficient found in pure D_2O the diffusion of ASA and β-CD is slowed down in the ternary mixture by $\approx 40\%$ and $\approx 1\%$, respectively. We can understand that the diffusion of β-CD is less effected by the complex formation by looking at the following

arguments: First, ASA is located in the inside of the cavity, so that the outer shape of the complex is similar to β-CD and the mass change is probably negligible, because ASA with a molar mass of $\approx 180\text{g/mol}$ replaces nine D_2O molecules with the same approximate mass. On the other hand the complex formation substantially slows down the diffusion of ASA, because the mass of β-CD+D_2O is more than seven-times the mass of ASA. The differences in the diffusion coefficients can be used to determine the association constant according to Eqs. 6 and 7.

The diffusion coefficient of the complex $D_{com.}$ can be calculated by

$$D_{com.} = \frac{b \cdot D_{ASA} - D_{\beta CD}}{b - 1} \tag{8}$$

with the ratio of diffusion changes upon mixing

$$b = \frac{D_{\beta CD, mix} - D_{\beta CD}}{D_{ASA, mix} - D_{ASA}}. \tag{9}$$

For our experiment however, the experimental errors for the diffusion coefficients $D_{\beta CD, mix}$ and $D_{\beta CD}$ are of the same order of magnitude as the difference $D_{\beta CD, mix} - D_{\beta CD}$ and so the error of b is too large to give $D_{com.}$ with any certainty. The calculated value is $D_{com.} = (2.4 \pm 20) \cdot 10^{-10} \text{ m}^2/\text{s}$.

2   Calculation of S_{T} for pure complex

We assume that the Soret coefficient S_{T}^{mix} measured for the mixture of CD and ASA in water is the weighted average of the Soret coefficients of the CD, ASA and the complex:

$$S_T^{mix} = \frac{p_{com.} \cdot S_T^{com.} + (1 - p_{com.}) \cdot S_T^{CD} + (1 - p_{com.}) \cdot S_T^{ASA}}{2 \cdot (1 - p) + p} \tag{10}$$

The theoretical value $S_{\mathrm{T}}^{com.}$ for a solution where 100% of the CD and ASA are forming a complex is then given by

$$S_T^{com.} = \frac{[2 \cdot (1 - p) + p] \cdot S_T^{mix} + (1 - p_{com.}) \cdot S_T^{CD} + (1 - p_{com.}) \cdot S_T^{ASA}}{p_{com.}}. \tag{11}$$

The fraction of complex $p_{com.}$ is dependent on temperature. To estimate $p_{com.}$ in our temperature range, we used the change of chemical shift in ^{1}H-NMR measurements of the complex as an indicator of the amount of complex present. We looked at the chemical shift of the H_3 and H_5 peaks of β-CD and found the results consistent with ROESY experiments. The difference in chemical shift between the free CD and the complex $\Delta\delta$ at 25 and 60°C is shown in fig. 2. The measurement at 25 °C agrees well with literature [5]. The reduction of $\Delta\delta$ at higher temperatures indicates a decreased amount of inclusion complex. To quantify the results, we estimate the fraction of
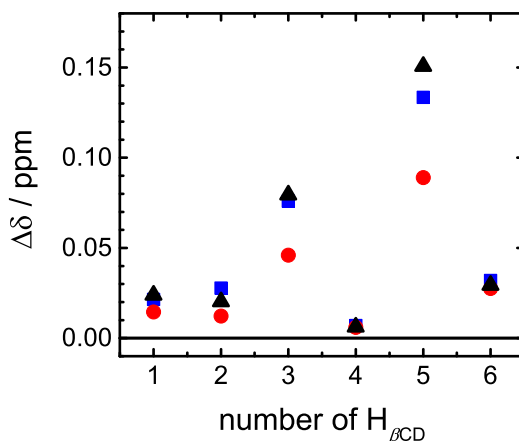

Figure 2: Chemical shift difference $\Delta\delta = \delta_{\beta CD+ASA} - \delta_{\beta CD}$ with adding ASA at 25°C (blue square), 60°C (red circle), and reference values[5] (black triangle, at 24°C) for each βCD peaks(see text).

complex at 60°C $p_{com.}(60°C)$ from $p_{com.}(25°C) = 0.45$ calculated from the diffusion measurements

$$p_{com.}(60°C) = \frac{\delta(60°C)}{\delta(25°C)} \cdot p_{com.}(25°C). \tag{12}$$

Interpolating the fraction of complex $p_{com.}(T)$ for the temperature range of the IR-TDFRS measurements, we are able to calculate $S_T^{com.}$ from equation 11. The result for β-CD is shown in fig. 3 (blue symbols). The calculations show that for the mixture of CD and ASA, the free ASA leads to a decrease in the observed temperature dependence of the thermodiffusion ΔS_T. Note that the large error of $p_{com.}(25°C)$ and the assumption of a linear temperature dependence of $p_{com.}(T)$ result in a large uncertainty regarding the fraction of complex present, however, the variations in the observed temperature range are relatively small so that this should not influence the qualitative results. To illustrate this, the observed S_T of the mixture was calculated with a constant $p_{com.} = 0.45$ over the whole temperature range (green symbols).

The observed temperature dependence of the thermodiffusion of the com-
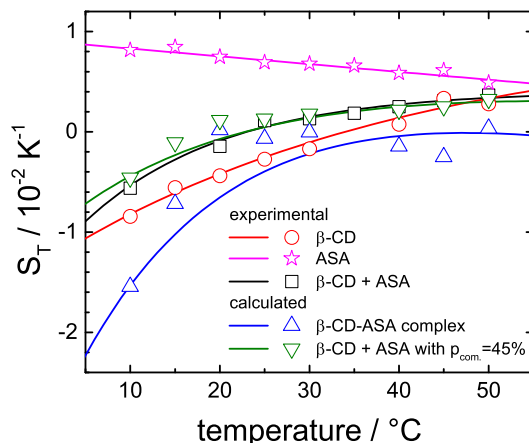
Figure 3: $S_T^{com.}$ against temperature assuming 100% β-CD-ASA complex (blue symbols) calculated from the experimental results of β-CD (red symbols), ASA (pink symbols) and mixture of both (black symbols) in water. The symbols show experimental values or values calculated directly from them, lines show fits of these values according to eq. (2) of the main text or values calculated from those fits. The green symbols show S_T for the mixture of CD and ASA in the case of a constant $p_{com.}$, illustrating the relatively small influence on the results due to the weak temperature dependence of the complex fraction found for this temperature range.

plex $\Delta S_{\mathrm{T}}^{com.}$ can now be calculated from that of the mixture of free ASA and CD $\Delta S_{\mathrm{T}}^{mix}$ (see fig. 4). To estimate the $\Delta S_{\mathrm{T}}^{com.}$ of α- and γ-CD, we used the $p_{com.}(T)$ determined for β-CD, which might deviate slightly from the actual values. The estimation of $\log P$ for the complexes is explained in the following section.
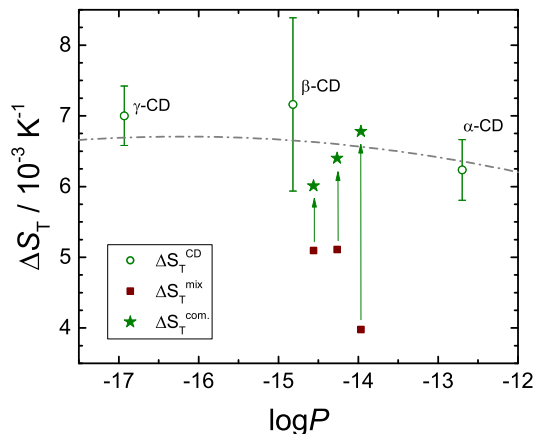
Figure 4: $\Delta S_{\mathrm{T}} = S_{\mathrm{T}}(50°\mathrm{C}) - S_{\mathrm{T}}(20°\mathrm{C})$ against $\log P$ for the empty CDs (green circles), the mixtures of ASA, CD and CD/ASA-complex (red squares) and calculated for the CD/ASA-complex alone (green stars). The dash-dotted line is the fit also shown in Figure 8 of the main text.

3   Donor/acceptor concept

The donor/acceptor concept was proposed by Maeda *et al.* [6], who investigated the thermodiffusion of ethylene glycols and crown ethers, which can be viewed as cyclic ethylene glycols. They found that the Soret coefficient S_T correlates with the difference of donor and acceptor groups in the molecule. Hydrogen bonds are formed between positively polarized hydrogen and the free electron pair of a negatively polarized atom. In the case of ethylene glycols and the cyclodextrins investigated by us, the hydrogen always belongs to a hydroxyl group (-OH), denoted as hydrogen donor, and the free electron pair always belongs to ether oxygen (-O-), denoted as hydrogen acceptor. Maeda *et al.* chose to count each hydroxyl group as one donor and each ether bridge as one acceptor (method (a) in tab. 1 and fig. 5). There are some problems that arise counting the donor and acceptor groups for our investigation:

- Each oxygen atom has two free electron pairs and could therefore be counted as two acceptors.

- The oxygen of a hydroxyl group can act as a hydrogen bond acceptor as well.
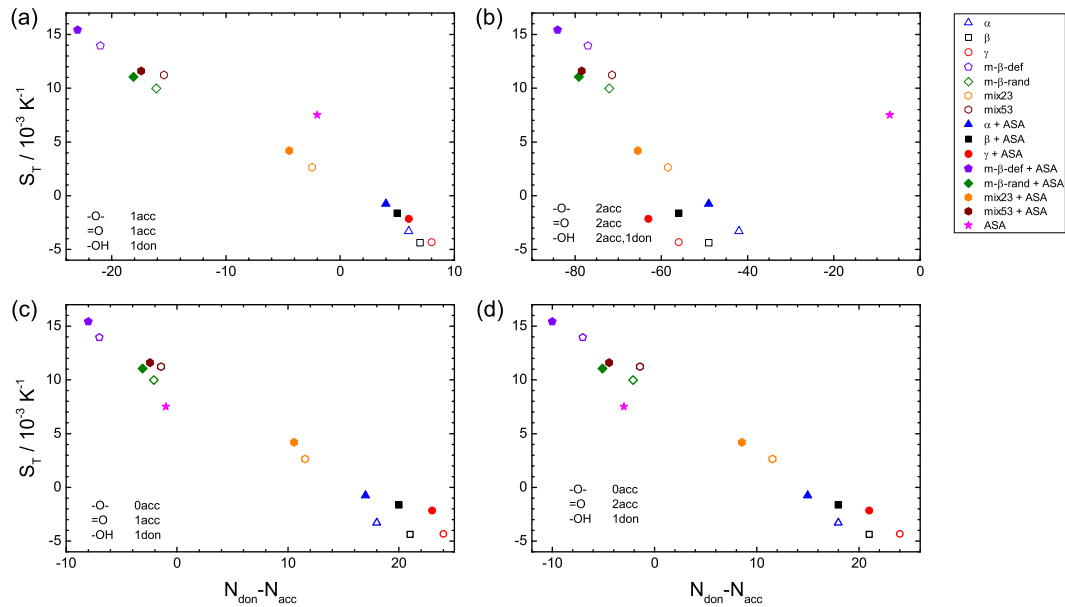

Figure 5: S_T against difference of donor and acceptor groups $N_{don} - N_{acc}$ with different counting methods.

- We are expanding the concept to include ASA, which contains an acid and an ester group.

- We assume that we can count donor and acceptor groups for the complex as the sum of both components.

To find out which way of counting donors and acceptors is the best, we compare different approaches. The number of acceptor sites for the different counting methods are listed in tab. 1 and graphs of S_{T} against $N_{don} - N_{acc}$ can be found in fig. 5. We see that there is a linear correlation for the cyclodextrins, no matter what counting method is applied, but the aspirin fits the data set more or less well. The bad agreement of method (b) shows that the hydroxyl group should indeed be only considered as a hydrogen bond donor. Method (a), which was introduced by Maeda *et al.* and employed by us in the main work, works as well as method (c) and (d) for the investigated systems. This is probably the case, because in the cyclodextrins the number of ether bridges does not change without also changing the number of hydroxyl groups. For the ethylene glycols, where the increments are defined by additional ether groups, (c) and (d) would not work.

Since $\log P$ is usually additive and it shows a linear dependence for homologous groups. From the linear fit of $\log P$ versus $N_{don} - N_{acc}$, the $\log P$ of the complexes with aspirin can be estimated (see fig. 6).
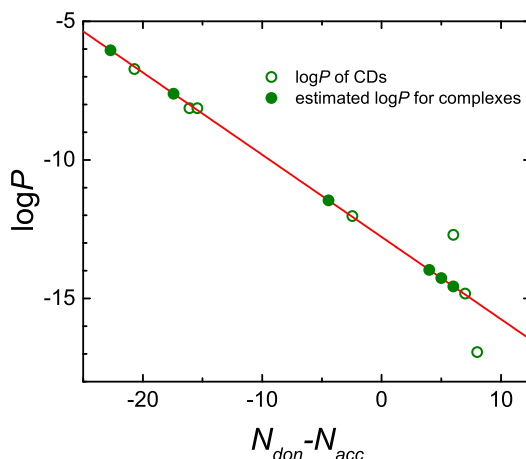

Figure 6: Estimation of the 1-octanol/water partition coefficient $\log P$ for the CD-ASA complexes based on the linear correlation with the difference between hydrogen bond donor and acceptor sites.

9

Table 1: Number of donor and acceptor sites in the molecule or complex according to different counting methods denoted in the header (letters (a) to (d) refer to fig. 5). The number of donors N_{don} does not vary. The last line gives the slope error of the linear fit for each method.

	N_{acc}				N_{don}
	(a)	(b)	(c)	(d)	
	-O- 1acc	-O- 2acc	-O- 0acc	-O- 0acc	
	=O 1acc	=O 2acc	=O 1acc	=O 2acc	
	-OH 1don	-OH 2acc,1don	-OH 1don	-OH 1don	
α-CD	12	60	0	0	18
β-CD	14	70	0	0	21
γ-CD	16	80	0	0	24
m-β-def	28	84	14	14	7
m-β-rand	25.55	81.55	11.55	11.55	9.45
mix23	18.725	74.725	4.725	4.725	16.275
mix53	25.214	81.214	11.214	11.214	9.786
α-CD + ASA	15	68	2	4	19
β-CD + ASA	17	78	2	4	22
γ-CD + ASA	19	88	2	4	25
m-β-def + ASA	31	92	16	18	8
m-β-rand + ASA	28.55	89.55	13.55	15.55	10.45
mix23 + ASA	21.725	82.725	6.725	8.725	17.275
mix53 + ASA	28.214	89.214	13.214	15.214	10.786
ASA	3	8	2	4	1
lin. fit error	6%	46%	5%	5%	

4   Influence of pH

To estimate the influence of the pH change due to addition of aspirin (ASA) we conducted experiments in buffer for 1 wt% β-CD and 1 wt% β-CD with equimolar ASA. Acetate buffer (0.1 M) was prepared with 1.1118 g CH_3COOH, 122.9 mg CH_3COONa, and 200 mL Millipore water. The pH is

$$\text{pH} = \text{pK}_a + \log \frac{[A^-]}{[HA]} = 3.67. \tag{13}$$

The pH of both solutions (β-CD and β-CD + ASA) was confirmed with a quick test (Macherey-Nagel pH-fix 0-14) as pH 3-4.

Figure 7 shows the results of the IR-TDFRS measurements. The Soret coefficient S_T in buffer is slightly increased compared to the systems in water (7a). Comparing the thermal diffusion coefficient D_T (7b) shows very little change between the solutions in water and in buffer. Note that D_T for the cyclodextrin in water and in buffer lie on top of each other, although these solutions deviate strongly in pH. The error bars for β-CD + ASA solution in buffer at higher temperatures are large, because the additional component of the buffer makes data evaluation difficult. In conclusion, these measurements show that the change in pH due to the addition of aspirin has no influence on D_T. There is, however, a slight increase of S_T with decreasing pH. This indicates a slower diffusion at lower pH values corresponding to a higher ionic strength I. For large I hydrodynamic interactions $H(q)$ influence the col-
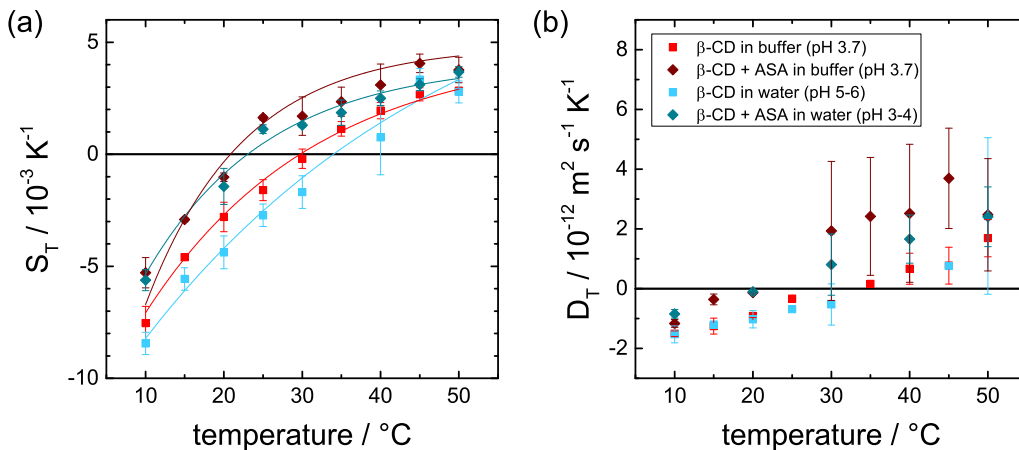
Figure 7: IR-TDFRS results of β-CD and β-CD + ASA in acetate buffer with pH 3.7 (red symbols) and in water (blue symbols).

lective diffusion coefficient, which can be expressed as $D(q) = D_0 \cdot H(q)/S(q)$ with the Stokes-Einstein diffusion coefficient D_0 [7]. In the long wave length limit $S(q)$ can be approximated by $S(0) \equiv \lim_{q \to 0} S(q) = \rho_0 k_B T \chi_T$ and is proportional to the isothermal osmotic compressibility χ_T, which describes the deviations from an ideal solution and increases with increasing salt content, so that the collective diffusion becomes slower [7].

5 Refractive index contrast factors

In order to evaluate the results of IR-TDFRS the so-called contrast factors, the change of refractive index n with concentration $(\partial n/\partial c)_{p,T}$ and with temperature $(\partial n/\partial T)_{p,c}$ are needed.

The refractive index of the cyclodextrin and cyclodextrin-aspirin solutions with cyclodextrin concentrations of 0.25, 0.5, 0.75 and 1wt% was measured in a temperature range from 10 to 50°C with an Abbe refractometer (Anton Paar, Abbemat RXA 156). A linear fit of the refractive index against concentration was done, the slope gives $(\partial n/\partial c)_{p,T}$. We did not observe a temperature dependence of the slope and averaged $(\partial n/\partial c)_{p,T}$ of the different temperatures.

The change of refractive index with temperature $(\partial n/\partial T)_{p,c}$ was measured interferrometrically in a tempeature range from 10 to 55°C.

Table 2: Change of refractive index n with concentration (weight fraction) averaged over a temperature range from 10 to 50 °C for 1 wt% cyclodextrin solution and solutions with aspirin (ASA) added in equimolar amounts.

	$(\partial n/\partial c)_{p,T}$ only CD	$(\partial n/\partial c)_{p,T}$ CD + ASA
α-CD	0.1283	0.1692
β-CD	0.1251	0.1596
γ-CD	0.1322	0.1764
m-β-def	0.1338	0.1672
m-β-rand	0.1277	0.1422
mix23	0.1294	0.1571
mix53	0.1276	0.1528
β-CD in buffer	0.1250	0.1480

Table 3: Change of refractive index n with temperature for 1 wt% cyclodextrin solution and solutions with aspirin (ASA) added in equimolar amounts.

| | $(\partial n/\partial T)_{p,c}$ | / | 10^{-4}K^{-1} | | |
	10°C	20°C	30°C	40°C	50°C
α-CD	-0.59966	-0.92454	-1.21558	-1.47279	-1.69617
β-CD	-0.52348	-0.88071	-1.19210	-1.45768	-1.67743
γ-CD	-0.51967	-0.88053	-1.19246	-1.45549	-1.66959
m-β-def	-0.54482	-0.91006	-1.22337	-1.48474	-1.69417
m-β-rand	-0.55563	-0.91046	-1.22026	-1.48503	-1.70477
mix23	-0.540754	-0.90586	-1.22147	-1.4876	-1.70424
mix53	-0.54515	-0.91883	-1.23629	-1.49751	-1.70251
β-CD in buffer	-0.59455	-0.94144	-1.24203	-1.49632	-1.7043
α-CD + ASA	-0.54549	-0.91155	-1.22549	-1.48733	-1.69706
β-CD + ASA	-0.53723	-0.90826	-1.22244	-1.47976	-1.68023
γ-CD + ASA	-0.54484	-0.91312	-1.22771	-1.48859	-1.69577
m-β-def + ASA	-0.54761	-0.91397	-1.22757	-1.48842	-1.69652
m-β-rand + ASA	-0.52852	-0.90029	-1.21905	-1.4848	-1.69755
mix23 + ASA	-0.53925	-0.90513	-1.21952	-1.48241	-1.69382
mix53 + ASA	-0.54478	-0.91218	-1.22641	-1.48747	-1.69536
β-CD + ASA in buffer	-0.59703	-0.94134	-1.24234	-1.50003	-1.71443

6   Temperature dependence of S_T

The temperature dependence of the Soret coefficient $S_T(T)$ for many different types of systems follows the empirical scaling formula,

$$S_T(\tilde{T}) \, = \, S_T^\infty \, \left\{ 1 - \exp \left[A \left(1 - \tilde{T} \right) \right] \right\}, \qquad (14)$$

that has been proposed in Ref.[8]. Here, S_T^∞ is the Soret coefficient at infinite temperature, $\tilde{T} = T/T^*$ is a reduced temperature, and A a freely adjustable parameter. Vigolo *et al.* found that A is constant, which indicates that, at least for similar systems, the fit parameters T^* and T_0 of eq. 2 of the main text [9] are not independent from each other, because $A = T^*/T_0 = const$.

Figure 8 shows the reduced master plot of our investigated systems with an overall fit (solid line) according to eq. 14 with $A = 10.6 \pm 0.1$. The systems with a high degree of methylation deviate systematically from the overall fit, which might be related to the fact that the system is less hydrophilic, which can also be related to the formation of micro-heterogeneous structures [10]. For comparison, we also show the master curve of the surfactant system
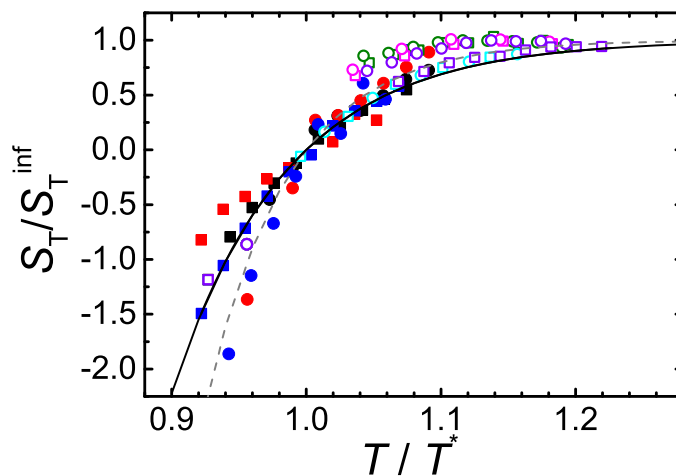

Figure 8: Normalized Soret coefficient $S_T(\tilde{T})/S_T^\infty$ versus $\tilde{T} = T/T^*$ for CDs without ASA (square) and the equimolar mixtures of CD and ASA (circle): α-CD (black), β-CD (red), γ-CD (blue), β/methyl(def.)-β-CD(23%) (light blue), β/methyl(def.)-β-CD(53%) (magenta) methyl(rand.)-β-CD(55%) (violet) and methyl(def.)-β-CD(67%) (olive). For the mixtures containing methylated CDs we used open symbols and percentages give the degree of methylation. The solid line refers to a fit according to Eq. 14 and the dashed line corresponds to the fit of the surfactant system [8].

15

investigated in Ref. [8] with $A \simeq 16$ (dashed line), which describes a large fraction of the data surprisingly well, despite the quite different chemical structure of the ionic surfactant from the CDs. Vigolo *et al.* [8] concluded from their findings that the temperature dependence of S_T is independent of the thermoelectric effects they were investigating. The small deviation between their surfactant system and our CDs when rescaled suggests that the temperature dependence of $S_\mathrm{T}(T)$ is quite general for dilute aqueous systems. From the correlation with hydrophilicity of the solute discussed in the main text, we can infer that the temperature dependence of S_T depends on the interaction strength between solute and solvent and that this general temperature dependence most probably reflects the temperature dependence of hydrogen bond strength.

7 Sample composition

Aqueous solutions of CDs and CD-ASA complexes were alsways prepared with a concentration of $1.00 wt\%$ CD and, in case of the complex solutions, an equimolar amount of ASA. The amount of ASA for the complex solutions can be found in tab. 4. The compositions of the mixed solutions are given in fig. 5.

Table 4: Molar mass M, mole fraction χ and weighted portion of ASA in the investigated solutions. All solutions were prepared with 5 mL water and 50 mg cyclodextrin.

cyclo-dextrin	M / g/mol	χ / 10^{-4}	$m(ASA)$ / mg equimolar
α	973	1.85	9.2
β	1135	1.59	7.9
γ	1297	1.39	6.9
m-β-rand.	1303	1.38	6.9
m-β-def.	1331	1.35	6.7

Table 5: Masses of methyl-β-CD, defined β-CD, and ASA for the preparation of the mixtures in 5 mL water to adjusted the methylation degree.

	m (β-CD) / mg	m (m-β-def.) / mg	m (ASA) / mg
mix 23%	8.4	41.7	-
mix 23% + ASA	8.5	42.2	7.3
mix 53%	31.1	18.9	-
mix 53% + ASA	31.1	18.9	7.5
ASA	-	-	9.2

References

[1] H. Weingärtner. Diffusion in liquid-mixtures of light and heavy-water. *Ber. Bunsen Ges. Phys. Chem.*, 88:47–50, 1984.

[2] G. M. Pavlov, E. V. Korneeva, N. A. Smolina, and U. S. Schubert. Hydrodynamic properties of cyclodextrin molecules in dilute solutions. *Eur. Biophys. J.*, 39:371–379, 2010.

[3] L. J. Edwards. The dissolution and diffusion of aspirin in aqueous media. *Trans. Faraday Soc.*, 47:1191–1210, 1951.

[4] Alexandria Kate Rogerson. *New Techniques in Diffusion-Ordered NMR Spectroscopy.* PhD thesis, 2013.

[5] T. Loftsson, B. J. Olafsdottir, H. Frioriksdottir, and S. Jonsdottir. Cyclodextrin complexation of NSAIDs - physicochemical characteristics. *Eur. J. Pharm. Sci.*, 1:95–101, 1993.

[6] K. Maeda, N. Shinyashiki, S. Yagihara, S. Wiegand, and R. Kita. Ludwig-soret effect of aqueous solutions of ethylene glycol oligomers, crown ethers, and glycerol: Temperature, molecular weight, and hydrogen bond effect. *J. Chem. Phys.*, 143:124504, 2015.

[7] G. Nägele. *The physics of colloidal soft matter*, volume 14. Institute of Fundamental Technological Research, Polish academy of Science, Warsaw, 2004.

[8] D. Vigolo, S. Buzzaccaro, and R. Piazza. Thermophoresis and thermoelectricity in surfactant solutions. *Langmuir*, 26:7792–7801, 2010.

[9] R. Piazza. *Philos. Mag.*, 83:2067–2085, 2003.

[10] D. Niether, D. Afanasenkau, J.K.G. Dhont, and S. Wiegand. Accumulation of formamide in hydrothermal pores to form prebiotic nucleobases. *Proc. Natl. Acad. Sci. U.S.A.*, 113:4272–4277, 2016.

Declaration of Individual Contribution

Article no. 1

Title: Accumulation of Formamide in Hydrothermal Pores to Form Prebiotic Nucleobases
Corresponding Thesis Chapter: Chapter 2
Status: Published. DOI: 10.1073/pnas.1600275113

D. Niether, D. Afanasenkau, J. K. G. Dhont, S. Wiegand, Proc. Natl. Acad. Sci. USA 113(16), 4272 - 4277 (2016);

The independent contribution of the candidate:
I did all experiments and the data analysis 100% by myself. In compiling the simulation model and running the simulations I was involved ~ 35%. I wrote about 60% of the manuscript.

Article no. 2

Title: Heuristic Approach to Understanding the Accumulation Process in Hydrothermal Pores
Corresponding Thesis Chapter: Chapter 3
Status: Published. DOI: 10.3390/e19010033

D.Niether, S.Wiegand, Entropy 19, 33 (2017);

The independent contribution of the candidate:
I did all experiments and the data analysis 100% by myself. I ran about 30% of the simulations. Interpretation of the simulation results was done 100% by me. I wrote about 80% of the manuscript.

Article no. 3

Title: Unravelling the Hydrophobicity of Urea in Water Using Thermodiffusion: Implications for Protein Denaturation
Corresponding Thesis Chapter: Chapter 4
Status: Published. DOI: 10.1039/C7CP05843H

D. Niether, S. Di Lecce, F. Bresme, S. Wiegand, Physical Chemistry Chemical Physics 20, 1012 - 1020 (2018);

The independent contribution of the candidate:
I did all TDFRS experiments and data analysis 100%. The MD simulations were run and analysed by our collaborators Dr. Di Lecce and Prof. Dr. Bresme. I contributed about 60% to the interpretation of the results and the writing of the manuscript.

Article no. 4

Title: Thermophoresis of Cyclic Oligosaccharides in Polar Solvents
Corresponding Thesis Chapter: Chapter 5
Status: Published. DOI: 10.1140/epje/i2016-16086-5

K. Eguchi, **D. Niether**, S. Wiegand, R. Kita, Eur. Phys. J. E 39(9), 86 (2016);

The independent contribution of the candidate:
Mr. Eguchi, the master student I supervised, performed 80% of the experiments, I performed 20%. I contributed 80% of data analysis, interpretation of results and writing of the manuscript.

Article no. 5

Title: Role of Hydrogen Bonding of Cyclodextrin-Drug Complexes Probed by Thermodiffusion

Corresponding Thesis Chapter: Chapter 6

Status: Published. DOI: 10.1021/acs.langmuir.7b02313

D. Niether, T. Kawaguchi, J. Hovancova, K. Eguchi, J.K.G. Dhont, R. Kita, S. Wiegand, Langmuir 33(34), 8483-8492 (2017);

The independent contribution of the candidate:

I performed about 45% of the experiments and data analysis, the master students Ms. Hovancova and Mr. Eguchi contributed 50 and 5%, respectively. Both were supervised by me. Mr. Kawaguchi conducted the NMR experiments in this work. Interpretation and connection of the NMR data to the TDFRS data was done by me to about 80%. I contributed about 80% to the writing of the manuscript.

Article no. 6

Title: Thermodiffusion as a Probe of Protein Hydration for Streptavidin and the Streptavidin-Biotin Complex

Corresponding Thesis Chapter: Chapter 7

Status: Published. DOI: 10.1063/1.5021914

D. Niether, M. Sarter, B. König, M. Zamponi, J. Fitter, A. Stadler, S. Wiegand, *AIP Conference Proceedings* 1929, 020001 (2018)

The independent contribution of the candidate:

I performed and evaluated 100% of the thermodiffusion experiments, but not the neutron scattering experiments. The parts detailing QENS technique and evaluation were written by Ms. Sarter and Dr. Stadler, the rest of the mauscript was written by me (amounting to about 85%).

Erklärung zur Dissertation

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit - einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen - noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluß des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. A. Schmidt betreut worden.

Nachfolgend genannte Teilpublikationen liegen vor:

1. **D. Niether**, D. Afanasenkau, J. K. G. Dhont, S. Wiegand, Accumulation of formamide in hydrothermal pores to form prebiotic nucleobases, *Proc. Natl. Acad. Sci. USA* 113(16), 4272 - 4277 (2016)

2. K. Eguchi, **D. Niether**, S. Wiegand, R. Kita, Thermophoresis of cyclic oligosaccharides in polar solvents, *Eur. Phys. J. E* 39(9), 86 (2016)

3. **D.Niether**, S.Wiegand, Heuristic Approach to Understanding the Accumulation Process in Hydrothermal Pores, *Entropy* 19, 33 (2017)

4. **D. Niether**, T. Kawaguchi, J. Hovancova, K. Eguchi, J.K.G. Dhont, R. Kita, S. Wiegand, Role of Hydrogen Bonding of Cyclodextrin-Drug Complexes Probed by Thermodiffusion, *Langmuir* 33(34), 8483-8492 (2017)

5. **D. Niether**, S. Di Lecce, F. Bresme, S. Wiegand, Unravelling the hydrophobicity of urea in water using thermodiffusion: implications for protein denaturation, *Physical Chemistry Chemical Physics* 20, 1012 - 1020 (2018)

6. **D. Niether**, M. Sarter, B. König, M. Zamponi, J. Fitter, A. Stadler, S. Wiegand, Thermodiffusion as a Probe of Protein Hydration for Streptavidin and Streptavidin-Biotin Complex, *AIP Conference Proceedings* 1929, 020001 (2018)

DOREEN NIETHER

Lebenslauf

Johannesplatz 14, 52428 Jülich
Tel.: 0174 9558521 | e-mail: d.niether@icloud.com
Staatsangehörigkeit: deutsch

❖ AKADEMISCHE AUSBILDUNG

seit 2015   **Ph.D.** Physikalische Chemie. Department für Chemie, **Universität zu Köln** und Institut für Komplexe Systeme – Weiche Materie, Forschungszentrum Jülich, Deutschland. *"Thermophoresis of biological and biocompatible systems"*

2009-2013   **M.Sc.** Chemie. Institut für Chemie, **Freie Universität Berlin**, Deutschland *"Photooxidation schwerflüchtiger organischer Schadstoffe"*

2005-2009   **B.Sc.** Chemie. Institut für Chemie, **Freie Universität Berlin**, Deutschland

❖ SCHULBILDUNG

2002-2005   Leonard Bernstein Oberschule, Berlin Hellersdorf

1998-2002   Pascal-Oberschule, Berlin Lichtenberg

1992-1998   Goethe-Grundschule, Neuenhagen bei Berlin

❖ AUSGEWÄHLTE PUBLIKATIONEN

2018   D. Niether *et al.*, *"Unravelling the hydrophobicity of urea in water using thermodiffusion: implications for protein denaturation"*, *Physical Chemistry Chemical Physics* 20, 1012 - 1020 (2018)

2017   D. Niether *et al.*, *"Role of Hydrogen Bonding of Cyclodextrin-Drug Complexes Probed by Thermodiffusion"*, *Langmuir* 33(34), 8483-8492 (2017)

2016   D. Niether *et al.*, "*Accumulation of formamide in hydrothermal pores to form prebiotic nucleobases*", *Proc. Natl. Acad. Sci. USA* 113(16), 4272 - 4277 (2016)

❖ AUSGEWÄHLTE PRÄSENTATIONEN

2017   *116th General Assembly of the German Bunsen Society for Physical Chemistry*, Kaiserslautern, Deutschland

2016   *12th International Meeting on Thermodiffusion*, Madrid, Spanien

XXV Sitges Conference on Statistical Mechanics, Barcelona, Spanien

❖ SPRACHEN

Deutsch (Muttersprache), Englisch (fließend), Französisch (Grundkenntnisse)