

THESIS FOR THE DEGREE OF LICENTIATE OF ENGINEERING

Interaction of trehalose with water and protein for the understanding of biological stabilization

CHRISTOFFER OLSSON



CHALMERS

Department of Physics
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2016

Interaction of trehalose with water and protein for the understanding of biological stabilization
CHRISTOFFER OLSSON

© CHRISTOFFER OLSSON, 2016

Department of Physics
Chalmers University of Technology
SE-412 96 Göteborg, Sweden

Chalmers Reproservice
Göteborg, Sweden 2016

Interaction of trehalose with water and protein for the understanding of biological stabilization

Christoffer Olsson
Department of Physics
Chalmers University of Technology
SE-412 96 Göteborg, Sweden

Abstract

The understanding of biomolecular interactions with water and co-solutes opens up further understanding for the mechanism behind biomolecular stabilization. This is highly important for developing technologies aimed to preserve biological material. Such techniques include cryopreservation of for example pharmaceuticals or human organ transplants. For these purposes, the disaccharide trehalose has been shown to be an outstanding stabilizing agent during cryostorage or storage of desiccated materials. However, the stabilizing role of trehalose is still not fully understood; why does trehalose perform better than other molecules?

To partly answer this question, this work investigated two important molecular systems. First, the structural properties of aqueous trehalose were studied using neutron diffraction combined with EPSR modeling. Secondly, ternary protein–trehalose–water systems were investigated using calorimetric experiments to obtain indirect evidence for different structural properties.

The aqueous trehalose study provided a direct proof of strong trehalose–water interactions, and consequently a strong perturbation of the bulk-water structure. Furthermore, this study found that the trehalose molecules are highly unlikely to cluster to each other, which is hypothesized to be the reason for why trehalose is able to interact so strongly with water. The results from the calorimetric study gave support to the preferential hydration model. This study also showed that the protein stability is not necessarily coupled to the glass transition temperature of the trehalose–protein–water-matrix.

Keywords: *trehalose, protein, biomolecule, water, amorphous, cryopreservation, neutron scattering, neutron diffraction, EPSR, DSC*

Publications included in this thesis

Paper I. Structure of aqueous trehalose solution by neutron diffraction and structural modeling

Christoffer Olsson, Helén Jansson, Tristan Young, and Jan Swenson

Submitted

Paper II. The Role of Trehalose for the Stabilization of Proteins

Christoffer Olsson, Helén Jansson, and Jan Swenson

J. Phys. Chem. B **2016**, *120*, 4723-4731

My contributions to the papers:

Paper I

I prepared all the samples and performed all the diffraction experiments together with the co-authors. I performed the data processing, EPSR simulations, data analysis, and I was the main author of the manuscript.

Paper II

I performed sample preparations and DSC measurements. Furthermore, I did most of the data analysis, and I wrote the first draft of the manuscript.

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 1 |
| 2 | Materials | 5 |
| 2.1 | Liquids and amorphous materials | 5 |
| 2.2 | Water | 10 |
| 2.3 | Trehalose | 14 |
| 2.4 | Biomolecules | 16 |
| 2.5 | Stabilization of biomolecules | 19 |
| 3 | Experimental and Computational Methods | 27 |
| 3.1 | Neutron Scattering | 27 |
| 3.2 | Empirical Potential Structure Refinement Modeling | 33 |
| 3.3 | Differential Scanning Calorimetry | 37 |
| 4 | Experimental Procedures | 39 |
| 4.1 | Neutron Diffraction Experiments | 39 |
| 4.2 | EPSR analysis | 42 |
| 4.3 | DSC Experiments | 43 |
| 5 | Summary of Appended Papers | 45 |
| 6 | Conclusions and Outlook | 49 |
| | Acknowledgments | 53 |
| | Bibliography | 55 |
| | Papers I-II | 61 |

1

Introduction

In the end of May 2013, pictures of a frozen wholly mammoth began to pop up in the newspapers. A team of Russian and South Korean scientists had discovered this extinct animal in a sheet of ice in Siberia; an occurrence which has become relatively more common recently. What was so spectacular about this particular mammoth however, was that it was possible to extract blood and seemingly fresh meat from it, after it had been frozen for over 43 000 years.¹ This fascinating discovery could very well mean that it could be possible to extract DNA from an intact mammoth cell, which could furthermore lead to the cloning of an animal that went extinct around 4000 years ago. Shortly after this discovery, I began my PhD-studies aimed to study the fundamentals of cryopreservation. Thus, this story about the mammoth served as a great source of inspiration; what made this particular mammoth-carcass able to be so well preserved for such a long time? How come biological material in general degenerate over time? What are the mechanisms responsible for suppressing these degenerations? There are multitudes of answers to these questions, most of them reaching far beyond the scope of any single licentiate thesis, but at least some of these questions will be addressed here.

The process of preserving biological material has huge importance in a large number of areas. Preservation of biological components for human transplants, such as different tissues or blood is one important life-saving area. Improved preservation techniques may have a special increase of demand in the future due to advances in e.g. tissue engineering,² where appropriate storage of the created

tissues is needed. In other areas such as pharmaceuticals or food-production, it is of course also important to obtain a long shelf-life for all the different products. Similarly, in several areas within biotechnology it is vital to keep cell cultures or other biological material viable for longer periods of time.

Two of the most common methods of preserving biological material are cryopreservation, and freeze-drying.³ Cryopreservation is the term for preserving materials at low temperatures. This is advantageous due to the decrease of motions of the stored complex biomolecular structures; motions which normally are required for biological function, but are also capable of disrupting or breaking these structures. However, a huge disadvantage of this temperature decrease is the formation of ice, which often cause great damage to the preserved material.⁴ Freeze-drying on the other hand attempts to immobilize the biological material by removing the water around it. Without an aqueous medium, most biological processes cease and the biomolecules become stabilized. This process can however also damage the biological molecules and is far from optimal in many scenarios.^{5,6} The problems connected to cryopreservation or freeze-drying are commonly, at least in part, dealt with by the addition of protective molecules (cryoprotectants and lyoprotectants). There are many such molecules, each typically well suited for one of the two methods, but there is one molecule that excels in stabilizing biological material when it comes to both mentioned methods: trehalose.

Trehalose is a sugar-molecule, very similar to more common sugar-molecules such as sucrose. Although sucrose and other sugar molecules typically exhibit good stabilization properties,⁷ trehalose almost always has the superior properties (see e.g. Ref. 8). Why trehalose possess these superior properties is still unclear, and it is the aim of this thesis to provide insights into this question.

For this purpose, the work in this thesis is presented in two papers. One paper concerning the structural properties of trehalose in an aqueous solution (33 wt% trehalose), as studied by neutron diffraction and empirical potential structure refinement-modeling, the other paper concerning the interaction of trehalose with water and protein, as studied by mainly differential scanning calorimetry.

In the structural diffraction study, the interaction between water and trehalose was studied. It was shown that plenty of hydrogen bonds between trehalose

and water was formed, and also that trehalose has a very small probability of forming intermolecular clusters.

In the differential scanning calorimetry study, evidence for the preferential hydration model was found. The study also includes glass transition and denaturation temperatures for a wide concentration range of both trehalose and protein.

This thesis contains some background information about the materials and the experimental techniques that are used in the studies. Chapter 2 first gives an overview of liquids and amorphous materials and how these can be investigated. This is followed by a more detailed description of water, protein, and trehalose. Subsequently, some models regarding the interactions between these materials, and how protein stabilization occurs, are presented.

Chapter 3 describes the theories behind the different techniques, and in chapter 4 it is explained how these techniques were applied in the presented studies. Chapter 5 briefly describes the obtained results from the papers, and chapter 6 gives a summary of this work and an outlook on how to proceed with answering the questions concerning the stabilizing role of trehalose.

2

Materials

2.1 Liquids and amorphous materials

When discussing the structure of a liquid or an amorphous material, a specific framework is necessary to define what is meant by structure. As opposed to the structure of a crystal, where all the atoms and molecules have a more or less well-defined position relative to one another, an amorphous material lack most of such structural ordering. In fact, the definition of an amorphous material is a material which lack long-range order between atoms, i.e. there are no repeating units that can be found throughout the material.⁹ Of course, the lack of long-range order does not mean that there is no structure at all. On shorter length scales, molecules in an amorphous material can exhibit very complex and interesting structure which determines the property of the material, such as clusters, vesicles and ring structures to name a few example. The materials studied in this thesis appear in two types of amorphous phases, liquids and glasses. This section gives a brief general overview of these two phases; concerning their general properties and how to describe their structures.

2.1.1 From liquids to glasses

Typically, when a liquid is cooled the viscosity increases, partly due to the decrease in kinetic motion of the individual particles in the liquid, and partly due to an increase in density. When the motion becomes slower the particles

propagate through the material at a slower rate – they become caged by the surrounding molecules. Thus when a force is applied to the liquid it takes longer time for the particles to react to that force. This phenomenon is what we typically experience as a high viscosity. When the liquid is cooled below its freezing point it becomes energetically favorable for the particles to form lattices of particles, a process known as crystallization. However, for crystallization to occur the particles need some time to reorient themselves into the correct position of a lattice site. Thus if the freezing occurs very fast (quenching) the viscosity increases rapidly – the movements of the particles quickly slow down – and the time required for the particles to arrange themselves into a crystal lattice-structure may grow to extremely high time spans. When these time spans become larger than the experimental time spans (i.e. such that particle motions are too slow to react during the time span of the experiment) the material can go through what is called the *glass transition*; the material becomes vitrified.

The glass transition is often detected by e.g. calorimetric measurements, and can be seen as a gradual change of the enthalpy (H) dependent on temperature (Figure 2.1). This change in the slope of the enthalpy shows that the glass transition behaves rather differently compared to a "normal" transition, such as a crystallization process, where the enthalpy changes abruptly at a certain temperature (melting/crystallization-temperature T_m/T_c).*

A transition that exhibit a discontinuity in the directly observable thermodynamic quantities, such as enthalpy (and also volume and entropy), is called a first-order transition. A glass transition however displays a change in the slope of these thermodynamic properties (Figure 2.1), which means that the derivatives of these thermodynamic quantities – such as the thermal expansion $\alpha_T = \frac{\delta \ln V}{\delta T}$, or the heat capacity $C = \frac{\delta H}{\delta T}$ – are discontinuous at the glass transition. This type of transition is a so called second-order transition.^{9,10}

The temperature at which this discontinuity occur often serves as the definition of the glass transition temperature (T_g). An important difference between this transition temperature and that of e.g. a crystallization temperature, is that the glass transition temperature depends on the history of the material. For

*Other thermodynamic properties, such as the thermal expansion behaves similar to the heat capacity.

example, T_g will typically become lower if the material is cooled at a slower rate. This is due to the extra time available for the material to reach equilibrium at a slower cooling rate. If the material is cooled at a greater rate, the particles in the material become immobilized before they are able to reach a new equilibrium.⁹

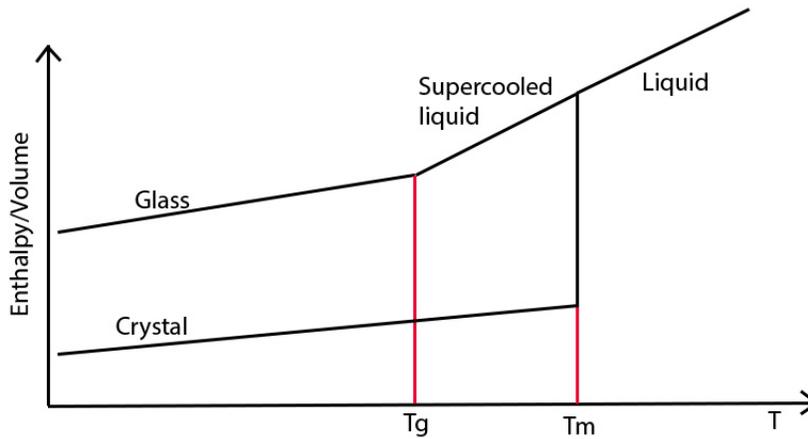


Figure 2.1: Graph showing how a substance varies in volume or enthalpy depending on temperature if it transitions from a liquid into either a glass or a crystal.

A common definition for the glass transition is that it occurs when a material reaches a viscosity of 10^{13} poise. At this viscosity the material can be regarded as a solid for all practical purposes.

Another common definition – directly related to the previously mentioned – for the glass transition is that it occur at the temperature where the characteristic α -relaxation process of the material, which is related to cooperative motions of the particles, reaches a relaxation time of 100 seconds. As previously mentioned, when a particle in a liquid is perturbed by a force it can only move if there are available sites for the particle to move to. The α -relaxation time (τ_α) is thus the characteristic time it takes for a particle to make that move. In a low viscosity liquid, τ_α is short and thus the particles collectively moves in response to an applied force, but as τ_α increases the longer it takes for these collective motions to occur, and thus the material appear as more and more viscous. τ_α can typically be probed using techniques such as dielectric spectroscopy.

2.1.2 Structure of Liquids and Glasses

In order to discuss the structure of a liquid there is a need for establishing a formalism that can accurately describe this. The most common and simple way to do this is via the pair correlation function, $g(r)^*$. This function is essentially the probability of finding another atom within a spherical shell between r and $r + dr$ at a distance r from the center of any arbitrary atom, as illustrated in Figure 2.2.

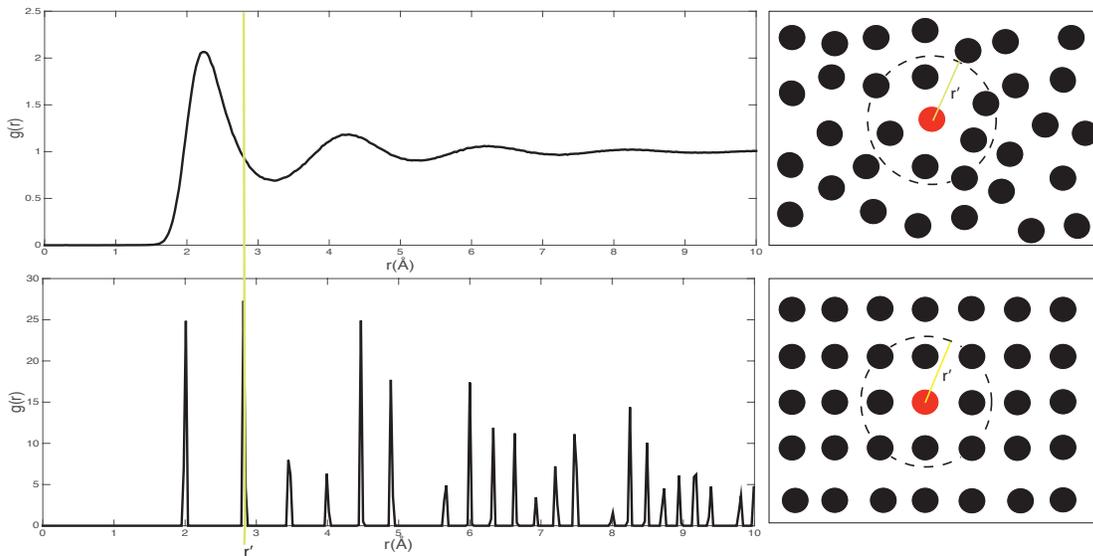


Figure 2.2: Pair correlation function schematic. Upper figure shows the pair correlation function of a hard-sphere liquid. Lower figure shows pair correlation function of a crystal structure.

In a crystalline material this function shows sharp peaks at specific r values, corresponding to the well-defined atomic distances within a crystal (Bottom figure in Figure 2.2). The broadening of these peaks is mainly due to structural thermal fluctuations of the atoms. The pair correlation function for an amorphous material (Upper figure in Figure 2.2) on the other hand oscillates around one at relatively large distances from the center. This reflects the random orientation of atoms at large distances from a central atom, where it is expected to find a number of atoms within a given volume, $n(r)$, equal to that of the average number density, ρ , of the material ($n(r) = \rho g(r)$). The peaks in an amorphous material at short distances are generally also broader compared

*More specifically, the *static* pair correlation function.

to those in a crystalline structure. This is mainly due to that the bonding sites are less well-defined in an amorphous material, but also partly due to a higher degree of structural thermal fluctuations of the atoms.⁹

The pair correlation function originates from the autocorrelation function of the local atomic density $n(\mathbf{r})$, defined as:¹¹

$$n(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \quad (2.1)$$

Where δ is the Dirac delta function, and \mathbf{r}_i is the position of atom i . The autocorrelation of $n(\mathbf{r})$ is*:

$$G(\mathbf{r}) = \frac{1}{N} \int n(\mathbf{r}')n(\mathbf{r}' + \mathbf{r})d\mathbf{r} = \frac{1}{N} \sum_{ij} \delta(\mathbf{r} + \mathbf{r}_j - \mathbf{r}_i) \quad (2.2)$$

and by separating the summation in two terms, one where $i = j$ and one where $i \neq j$, this can be reduced to:

$$G(\mathbf{r}) = \delta(\mathbf{r}) + \frac{1}{N} \sum_{i \neq j} \delta(\mathbf{r} + \mathbf{r}_j - \mathbf{r}_i) = \delta(\mathbf{r}) + \rho g(\mathbf{r}) \quad (2.3)$$

where the relationship $\rho g(\mathbf{r}) = \frac{1}{N} \sum_{i \neq j} \delta(\mathbf{r} + \mathbf{r}_j - \mathbf{r}_i)$ is the definition of $g(\mathbf{r})$ and refers to the pair correlation function (i.e. the correlation between any two different atoms), whereas the first Dirac delta term in equation 2.3 is the self-correlation part, i.e. how an atom correlates with itself.

Multicomponent systems

When a material contains more than one atom type it is often useful to reduce the total pair correlation function into a weighted sum of so called partial pair correlation functions of different atom pairs. These are often denoted as $g_{\alpha\beta}(\mathbf{r})$, where α and β represent two different atom types, and is interpreted as a probability of finding an atom of type β at a distance r from an atom of type α (or vice versa since $g_{\alpha\beta}(\mathbf{r}) \equiv g_{\beta,\alpha}(-\mathbf{r})$). The sum, describing the

*This autocorrelation function can be interpreted as a measure of the degree of correlation between two atoms separated by a distance \mathbf{r} .

autocorrelation function is then:¹¹

$$G(\mathbf{r}) = \delta(\mathbf{r}) + \rho g(\mathbf{r}) = \sum_{\alpha} c_{\alpha} \delta(\mathbf{r}) + \rho \sum_{\alpha, \beta \geq \alpha} (2 - \delta_{\alpha\beta}) c_{\alpha} c_{\beta} g_{\alpha\beta}(\mathbf{r}) \quad (2.4)$$

The partial pair correlation functions are very useful for the analysis of an amorphous material since they completely describe the average structure of the material, although as a one-dimensional representation of the real three-dimensional structure. Various information which can be extracted will be discussed in more detail in section 4.2. Section 3.1.1 describes how to obtain the pair correlation functions by neutron scattering and modeling techniques.

2.2 Water

Equipped with the tools to characterize the structure of a liquid, the formalism described above will be exemplified in this section using a relatively simple compound, one of the most abundant substances on earth: water. This is – to no surprise – a well-studied material, due to its presence in a vast number of chemical and biological reactions. For the scope of this thesis, it is important to highlight some of the properties of water, since it is the medium in which all the studied materials take place. Furthermore, the structure of water is also highly relevant for a part of the study presented in paper I.

2.2.1 Properties

Water consists of one oxygen atom covalently bonded to two hydrogen atoms. The hydrogen atoms are bonded to the oxygen in a triangular shape, where the H–O–H angle is around 104.5° on average.¹² This geometry stems from that the electrons in the water molecule slightly prefer to adhere to the oxygen atoms rather than to the hydrogens, which gives the molecule its dipole moment. Thus, water molecules attract each other through hydrogen bonds (HBs) between hydrogens and oxygens due to their slight charge differences. Compared to the covalent intramolecular bonds, hydrogen bonds are relatively weak. However, these hydrogen bonds are widely believed to be the main reason for a series of peculiar properties of water.

Water is – in many aspects – not a normal material, but possesses several anomalous properties. Most commonly known is probably the density maximum of water at 4°C; this property means that for a certain temperature interval, the density of H₂O *decreases* with decreasing temperature, which of course is an unusual behavior.¹² Another anomalous property of water is its high specific heat capacity (4.18J g⁻¹K⁻¹ at 25°) which is one of the highest heat capacities out of all known substances.*

To understand these anomalous properties, we ought to investigate the hydrogen bonded networks and structures formed within water at its different phases.

2.2.2 Structure

The intermolecular structure of ice is typically shown as in Figure 2.3 a. Each water molecule has four other water molecules as first-order neighbors (see section 4.2.1),¹² forming a tetrahedral structure in three dimensions. Ice is in a crystalline (relatively**) well-structured state; each water molecule is in a well-defined lattice point, connected symmetrically to four other water molecules. Liquid water is not as symmetrical since the water molecules translate and rotate due to thermal fluctuations. One widely accepted view of water structure has been that water molecules form on *average* a tetrahedral network, where each water molecule, on average, binds to 4.4 other water molecules.¹³ These bonds constantly break and form with a lifetime of ~1 ps at room temperature.¹⁴ Recent studies have however shown that most water molecules (80%) are involved in only two, relatively strong, hydrogen bonds, and only 20% are in a tetrahedral conformation (at 25°C).¹⁵

The structure of regular bulk water can be studied by the partial pair correlation functions, O–O, O–H, and H–H (Figure 2.4). Particularly $g_{OO}(r)$ – the partial pair correlation function between two water-oxygen atoms – tell us something interesting regarding the coordination numbers of water at 298K compared to ice at 220K. The first sharp peak at 2.8Å indicates the typical distance to a water molecules first-order neighbor. Notably, this peak which

*There are of course a lot more anomalous properties of water which are not mentioned here. For more information, the reader is referred to e.g. Ref. 12

**Ice is actually neither specifically structured compared to other crystalline materials

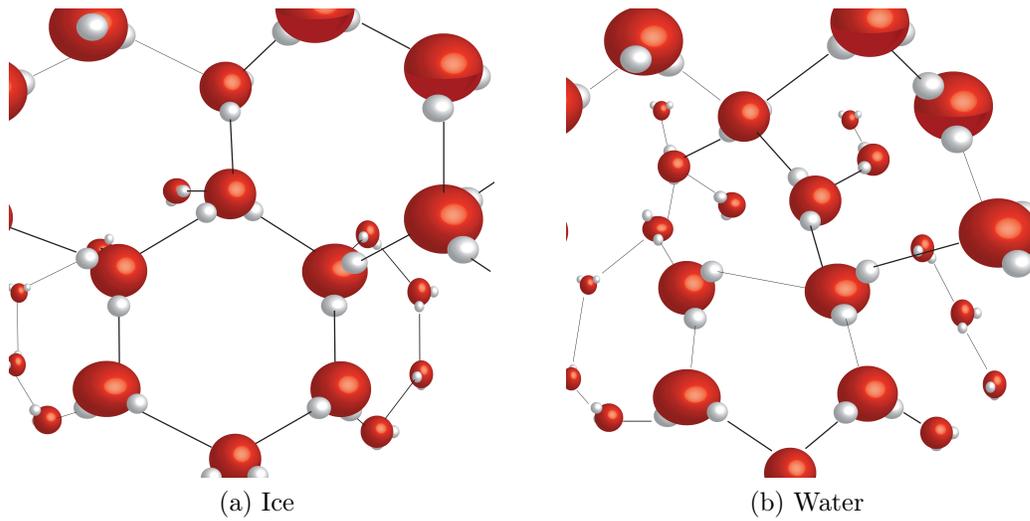


Figure 2.3: Schematic of ice and water structure.

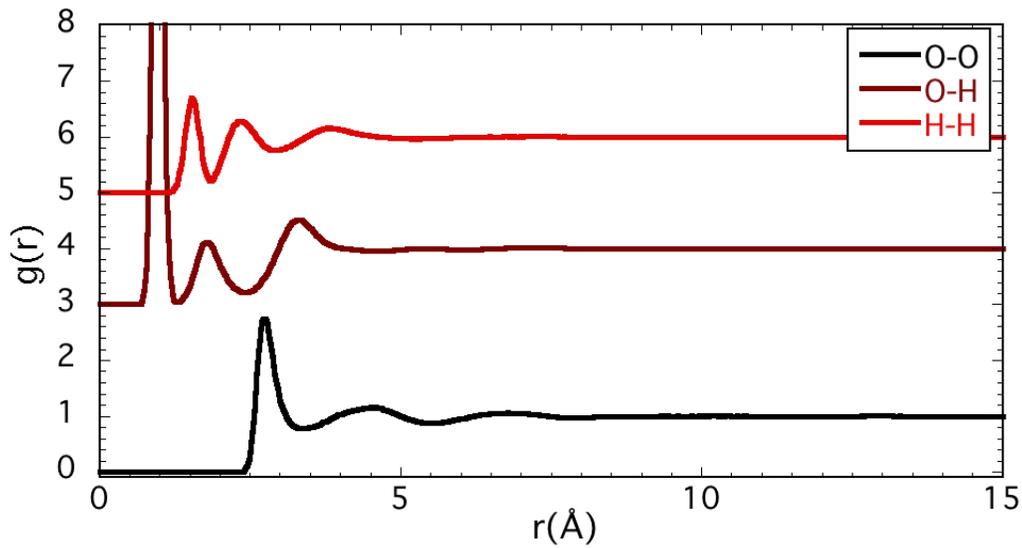
shows the first-order distance moves only to slightly higher values and becomes somewhat broader as the ice melts. A much more pronounced effect however is seen for the second peak (i.e. the second-order coordination shell), at about 4.5\AA , which is highly pronounced in ice structure and is significantly less pronounced in the liquid water. This second peak is seen as a signature of tetrahedral structure in water and ice. In ice, the water molecules are pretty well coordinated at the second-order neighboring (and to third and fourth etc. as well, but to a lower and lower extent). In liquid water this coordination is still present, however less defined, particularly at higher temperatures.*^{16,17}

From these correlation functions and other similar derived quantities, it has been shown that, in liquid water, a water molecule is typically surrounded by a tetrahedral structure of neighboring water molecules. The molecules in these structures are not stationary but rather fluctuate in their positions. First order neighbors exchange with second-order neighbors, and interstitial water molecules break up the otherwise four-coordinated water into five-coordinated water.^{18,19}

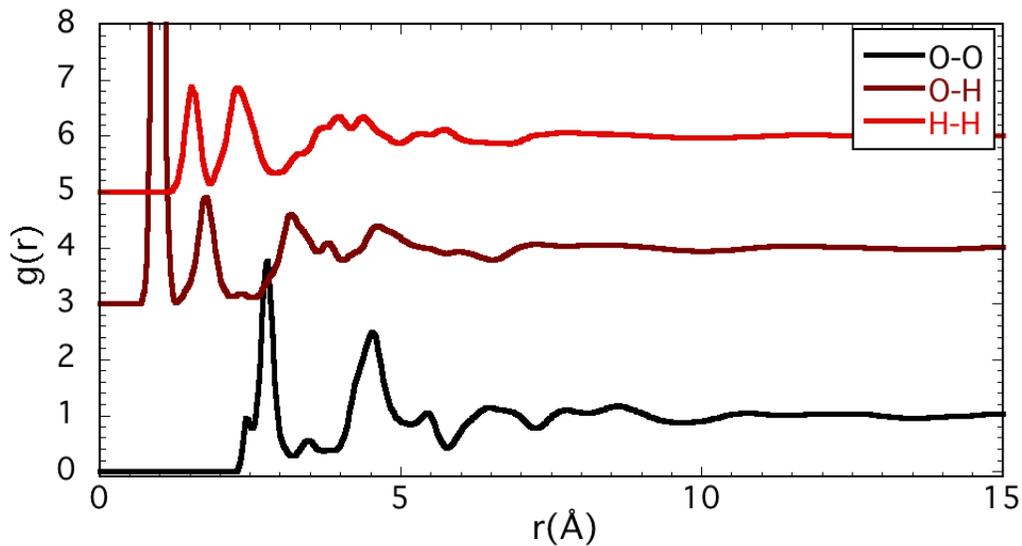
The abnormal density maximum of water can thus be explained by noting that the existence of destructured water arrangements increase with increasing temperature, and the destructured water is more densely packed than the

*A similar trend is found for pressure dependence; higher pressure changes the partial pair correlation functions similar to higher temperatures.^{16,17}

structured ones, even though both types of structures expand in volume with increasing temperature. Thus, at a certain point (at 4°C) the sum of the existence of density-increasing (destructured) water structures and the density increasing effect of lowering the temperature reaches a maximum. Below this temperature, more tetrahedral structures begin to form, thus lowering the density more than the decreasing temperature increases it.¹²



(a) Water at 298K



(b) Ice at 220K

Figure 2.4: Partial correlation functions for O–O, O–H, and H–H correlations of (a) water at 298K and (b) ice at 220K. Data obtained from the ISIS disordered materials database.^{17,20}

2.3 Trehalose

Trehalose appear to be present in a wide range of different extremophiles, and it has been shown in countless studies that trehalose possess an extraordinary ability to stabilize biological material against many different types of environmental stresses, such as desiccation, extreme temperatures or extreme pressure (see for example Ref.²¹ wherein the extremophile tardigrade is studied, a small microscopic animal capable of surviving some extreme environments). Understanding the mechanism of behind trehalose stabilization could thus mean a new understanding of the function of biological material in general, and what makes biological materials stabilized.

2.3.1 Chemical Properties

Trehalose is a disaccharide consisting of two glucose rings, linked by a α -1,1-glycosidic linkage. Its chemical formula is $C_{12}H_{22}O_{11}$, and it has a molecular weight of 342.296. A simple cartoon of it can be seen in Figure 2.5. This chemical formula is identical to a number of other disaccharides, such as sucrose or maltose. In fact, these molecules are almost identical in structure, apart from the positioning of the different hydroxyl and hydroxy-methyl groups on the glucose rings. The similar disaccharides share many properties with trehalose, however the small structure differences have been shown to play a highly important role for the molecular functions studied in the work presented in this thesis.

The melting point of water-free (anhydrous) trehalose is 203°C . However, trehalose easily binds to two water molecules – forming dihydrate crystals – which is the most stable form of trehalose at ambient conditions. This form has a melting temperature of 97°C and – after being melted – dissociates its water molecules and forms the anhydrous solid phase at about 130°C .²²

Comparing with similar disaccharides, trehalose has the highest T_g .²³ This property is highly dependent on water concentration, as the water molecules are able to plasticize the sugar matrix. It is also widely believed that trehalose not only forms a glass at the highest temperature, but also that it is superior in remaining in a stable glassy state.^{24,25} This effect stem from the fact that trehalose can transform into the dihydrate phase when exposed to water. A

small water addition to any other glassy disaccharide matrix destabilizes the entire glass by homogeneously distributing the water molecules. This process leads to a decrease in T_g which subsequently leads to a lower stabilization effect (see e.g. Ref. 26 and 27). In a glassy trehalose matrix however, the molecules which are immediately exposed to water acts as a buffer by absorbing it and thus excluding the water from the remaining glass-matrix. This process makes T_g more stable and therefore provides a more reliable stabilization.²⁵

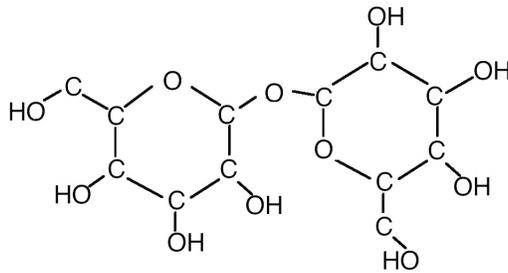


Figure 2.5: Sketch of a trehalose molecule.

2.3.2 Structure

The structure of trehalose in an aqueous solution is highly dependent on the water concentration. As the amount of water decreases, trehalose tend to fold across its glycosidic linkage, forming intramolecular bonds between the two glucose rings.²⁸ This folding prevents intermolecular bindings between the water and trehalose due to that many of the hydroxyl groups on the trehalose are occupied in the intramolecular fold. With an increasing amount of water, the folded structure breaks up and subsequently makes more hydroxyl groups available for further interactions between water and trehalose.²⁸

Intermolecular trehalose–trehalose interactions is another important property (also highly dependent on water concentration). In the dehydrated state (< 0.5 water molecules per trehalose), trehalose has been shown²⁹ to exhibit a homogeneous amorphous network, a property which could be partly responsible for its superior stabilizing mechanism during desiccation. This finding suggests that trehalose is better at hosting larger molecules (such as proteins), and anchor them to the homogeneous glassy matrix via the residual protein water molecules.²⁹

At high hydration levels (>38 water molecules per trehalose), trehalose also shows a high degree of homogeneity (compared to other disaccharides), in the sense that the trehalose molecules do not cluster substantially.³⁰ Furthermore, in paper I presented here it is shown, by neutron diffraction experiments, that the clustering effect of trehalose is much smaller than previously reported.^{30,31} In fact, according to the model in paper I, the trehalose molecules appeared to repel each other to a small extent.

2.3.3 Effect on water structure

It was shown through a series of articles by Branca and Magazu et.al³²⁻³⁷ that the structure and dynamics of water is highly altered by the presence of trehalose. This effect has however been under some discussion, which motivated the work done in paper I. In that paper it was also shown that trehalose perturb the bulk behavior of water by breaking up the tetrahedral network. By extensive – although not necessarily strong – hydrogen bonding, the water molecules are forced to reorient toward the trehalose hydroxyl groups, thus allowing for more interstitial water molecules in the first coordination shell of water followed by a clear destruction of the second coordination shell (see Figure 3,4, and 5 in paper I).

The destruction of the water structure is most likely the main reason for why trehalose is good at depressing ice formation and why aqueous solutions of trehalose exhibits such a high T_g compared to similar disaccharides.

2.4 Biomolecules

Biological molecules are a class of molecules typically defined as the molecules involved in biological systems and processes. They can be very complex, such as proteins or DNA, or much simpler, such as simple sugars involved in e.g. metabolism. When discussing biological stability, it is typically the more complex molecules and structures that are of interest since these are the ones more prone to breaking from environmental stress. The focus of this thesis is on the stabilization of proteins, but for a broader discussion it is important to some-

times deviate into the role of stability of other complex structures such as cell membranes.

2.4.1 Proteins

Proteins are the molecules that drives most biological processes. They are essential for the basic functions behind the process of life, thus the loss of these functions are catastrophic for the organism hosting them. Proteins are made up from chains of different amino acids, which link together according to different sequences in the DNA chain (i.e. genes). These protein chains typically fold to a subset of configurations that give the proteins their particular geometrical structures and functions. The folds and creases on the protein surfaces exhibit different chemical properties –such as being hydrophilic or hydrophobic– which typically determines their functions.³⁸



Figure 2.6: 3D model of myoglobin. The structure of the illustrated figure was obtained from the 1MBN entry from the protein data bank.³⁹

Myoglobin

In paper II, myoglobin was studied in solution with water and trehalose, since it is a well-studied relatively "simple" protein. Myoglobin was the first protein which structure was determined (via X-ray crystallography in 1958),⁴⁰

and has since then been often used as a model protein. It is a relatively small ($\sim 17600\text{g/mole}^*$) globular protein consisting of eight α -helices and a heme group (see Figure 2.6). The protein stores oxygen in the heme group, specifically in muscle tissue.³⁸

2.4.2 Water at Protein Surfaces

The view of proteins is often simplified as being fixed structures, which functions are only determined by their surface geometry. However, a functioning protein is almost always normally in an aqueous environment. In fact, the most prevalent molecule by mass of most organisms is water.⁴¹ Protein functions thus occur in the presence of water, and different configurations of a protein are normally slaved to this environment.^{42–44} Thus, the understanding how proteins function is not always merely a question of their static geometry, but on their dynamic behavior, and how this is affected by hydration or variation in local environment, by e.g. addition of different solutes. This is of course particularly important when considering the effects of protein stability in different environments.

Water at protein surfaces behaves – as when water binds to any other type of solute – very different from that in bulk. Hydration water, i.e. water confined by interactions to protein surfaces does not participate in ice formation. Such water has also much slower dynamics than bulk water. For water contents up to about 17 wt% the protein does not exhibit any function, and the water that adsorbs to the protein surface mainly bind to the most hydrophilic sites. As the hydration increases to 30 wt%, the protein is generally hydrated with a single water layer and the function then resumes. The addition of water enables dynamical behavior of the protein; at first locally (so called β -relaxations, related to e.g. rotational motions of protein side-chains or dipole reorientations), but as more water is added the more collective motions – coupled to the α -relaxations – can emerge. Full protein function does however not return until hydration of about 50 wt%, i.e. equal amounts of water and protein,⁴⁵ at which point the protein flexibility associated with adding water stop increasing substantially.

*For myoglobin obtained from horse heart-muscles

2.5 Stabilization of biomolecules

Although the work done in this thesis is focused on the effects of trehalose, the discussion will first describe why some solutes in general are able to stabilize biomolecules, and secondly, in section 2.5.3 the discussion will focus on what makes trehalose such an excellent protective solute.

2.5.1 Protein stability

Protein dynamic is indeed a criterion for the functionality of proteins, as was previously discussed in section 2.4. However, with great flexibility and motions comes greater risks for denaturation to occur. The probability of reaching the denatured state increases with temperature (and therefore an increase in the protein dynamics), and at a specific point (T_d for denaturation temperature), the protein undergoes an irreversible phase transition into the denatured state. Thus, a decrease of the dynamics of the protein is essential for the long-term survivability of protein structures. One method of doing this, is to directly slow down the dynamics by decreasing the temperature, so called cryopreservation. This method is common for storing large and complex biological materials, such as cells or tissues, but it is also used for storing proteins for example.² Another common method of slowing down the dynamics is to reduce the water concentration, which – as pointed out in section 2.4 – reduces the conformational freedoms of proteins. This can be done by many different methods, one of the most common methods include freeze-drying (lyophilization). Lyophilization is suitable for storing less complex biological materials, such as proteins. It is done by first freezing the material dissolved in water, followed by keeping the frozen material in a vacuum, in which the water subsequently sublimates. One advantage of storing biological material in a lyophilized state is that this material can be kept at a much higher temperature (e.g. room temperature).⁶ There are however problems with both of these two different methods which have to be overcome for improving biological storage techniques.

Cryoprotection

Depressing ice formation is probably the most important aspect of a successful preservation of biological material at low temperatures.* Ice crystals can grow extra- and intra-cellularly puncturing or completely bursting membranes. Another serious problem with ice formation is that it effectively concentrates different chemicals in the cell to high, sometimes toxic levels. This is due to that ice does not dissolve any substantial amount of solute itself. Similarly, if extracellular water freezes, an effective osmotic stress arises due to the concentration of solutes at the cell membrane. This creates a concentration gradient that effectively dries out the cells.^{47,48} Furthermore, a crystalline ice structure can destroy the native configuration of proteins by forcing these into unfavorable conformations through direct interactions between ice and protein.⁴⁹ It is therefore desirable to freeze the protein solution in such a way that it avoids crystallization during cooling, by first entering the supercooled regime and then the glassy state. The glassy state is thus – at least in principle – a way of obtaining a solid but liquid-like environment that very much adapts its structure to that of the protein. Hence the risk for unfavorable steric constraints from a crystal structure is avoided.

Reaching the glassy state and avoiding crystallization is commonly done by the addition of cryoprotectants. These are typically non-reactive molecules that interfere with water crystallization and induce glass formation. There are however some general problems with cryoprotectants. To achieve a vitrified state, high concentrations of cryoprotectants are typically required, which can in itself be toxic. Some cryoprotectants are more toxic than others, but generally, all cryoprotective compounds are "toxic" at a high enough concentration.⁴⁸ Another problem is that the loading of cryoprotectants into a cell may induce osmotic stresses; if a cryoprotectant does not enter through the cell membrane fast enough, the buildup of cryoprotectant extracellularly dries out the cell, possibly leading to a volume collapse. Similarly, after successful cryostorage, the cryoprotectants have to be washed away during reheating, and

*There are methods that do not involve the avoidance of crystallization, but instead attempts to control the crystallization process by e.g. producing nucleation sites for the crystal structures to grow on (see e.g. Ref. 2). Some organisms use a version of this method by use of antifreeze proteins (see e.g. Ref. 46 for further information). These methods are however beyond the scope of this thesis.

a buildup of high intracellular cryoprotectant concentration may occur during that stage. This leads to an expansion of the cell, which again may lead to that the cell bursts.^{48,50}

Although avoiding ice formation may be crucial for successful cryopreservation, it has been shown to not be a sufficient property. Some co-solutes which depress ice formation and increase T_g offer little to no stabilization at low temperatures.⁵¹ So even though crystallization prevention and glass forming abilities seem to be near essential for successful storage of biomolecules, there has to be some other mechanisms that some cryoprotectants offer that others do not.

Lyophilization

Removing water is another method of immobilizing biological molecules, such as proteins. The reduction of water concentration leads to a reduction of molecular dynamics and biomolecular functionality, which has a stabilizing effect. However, when a biomolecule become desiccated, the loss of its function can also have negative effects on its stability.⁵² If enough water disappears it could lead to the aggregation and denaturation of protein^{53,54} and the destruction of different kinds of cell membranes.⁵⁵ This denaturation of protein from either freeze-drying or just air-drying appears to occur when the first hydration shell disappears.^{53,54}

In order to obtain a successful freeze-drying, one thus again ought to add protective molecules to ensure stability at a desiccated stage. It has been pointed out however that the stabilizing mechanism of lyophilization-protectants is fundamentally different from, and much more complex than, that of cryoprotection.⁵⁶⁻⁵⁸ This aspect can be indirectly evidenced by the fact that most cryoprotective molecules are not capable of desiccation-protection. Molecules that actually have this capability of both cryo- and desiccation-protection are disaccharides.^{7,59,60} Disaccharides may thus tell us a bit more about stabilization properties in general than other co-solutes.

2.5.2 Stabilizing co-solutes

A well-proven thermodynamic theory of stabilization/destabilization mechanisms through the addition of solutes has been developed by Timasheff et al. (see for example Ref. 61 or Ref. 62). According to this theory, the stability of the functional state of a protein is proportional to the difference in free energy between the functional state and the denatured state. The greater the difference, the more energy is required to break the functional state.

Stabilizing co-solutes in dilute solutions with protein have been found to be generally preferentially excluded from the protein surfaces (preferential hydration).^{63,64} It was pointed out that a preferentially excluded solute increases the free energy of the material.^{63,64} This increase is proportional to the surface area of the protein, and since the surface area is larger for the denatured protein, the free energy of the denatured state will increase more than that of the functional protein state.

2.5.3 Bioprotective properties of trehalose

Trehalose emerges in more and more areas as a particularly excellent biological stabilizer. Not only as a protectant against cold or heat, but also as a lyophilization stabilizer. It also has the ability to stabilize both lipid bilayers and proteins. The stabilization of proteins will however be the main focus of this thesis. What exactly it is that makes trehalose different from any other co-solute is not entirely understood, although there are some established theories, which will be discussed below.

Vitrification

Stabilizing co-solutes are typically associated with being glass formers. By encapsulating proteins in a glassy matrix, the dynamics of a biomolecule slow down due to the decrease of solvent motions, but without the negative effects associated with crystallization (e.g. unfavorable molecular geometry). Compared to other similar disaccharides, trehalose has been reported to have the highest glass transition temperature, and to be an excellent glass former. This observation led to the so called *vitrification hypothesis* – as first proposed by Green and Angell²³ – which suggests that the main reason for trehalose's ex-

ceptionality resides in its glass forming properties. The reason for this high glass forming property was previously discussed in e.g. section 2.3.3, and is further discussed in paper I. Furthermore, when exposed to moisture, trehalose remain a stable T_g for higher water contents than other disaccharides,^{24,25,65} which was discussed in section 2.3.1.

This view – that trehalose’s protective properties are due to its extraordinary glass forming properties – was however challenged by a range of studies where some researchers showed that, by using even better glass formers – such as dextran – trehalose was still better at preventing degradation of biomolecules.^{66,67} Thus, although the excellent glass forming properties of trehalose may be an important (if not the most important) aspect of its stabilizing properties, it is not sufficient to explain what makes trehalose special. Rather, there ought to be some more intricate interaction between trehalose and biomolecules. Several different models exist that describe this type of interaction, and some of them will be presented here. Specifically the water replacement model, and the preferential hydration model.

Water replacement model

It has been proposed that one important mechanism for further trehalose stabilization stems from direct trehalose–protein interactions. This model presumes that there is a preferential interaction between protein and trehalose, and the hydration layer would then be (at least partially) substituted with trehalose. This is hypothesized to yield a stabilizing effect due to that the protein is kept in its configuration through a direct coupling to a rigid trehalose matrix (see e.g. Ref. 59). Indeed, water replacement with trehalose seems to occur in the stabilization of lipid bilayers,⁶⁸ however it is less clear regarding protein stabilization. The water replacement model at first showed that the mechanism of stabilization during freeze-drying differed from that of e.g. cryoprotection. The cryoprotection mechanism had been proposed to act according to the preferential hydration model^{61–64} mentioned above, but this model was based on studies with relatively high water contents. The water replacement model on the other hand was primarily concerned with the stabilization mechanism at low water contents (e.g. during lyophilization). According to this model, when the water content is extremely low, then the water replacement by trehalose

occur. This view was however later disputed by e.g. Belton and Gil et al.,⁶⁹ who proposed an alternative theory: the water entrapment model, a variation of the preferential hydration model.

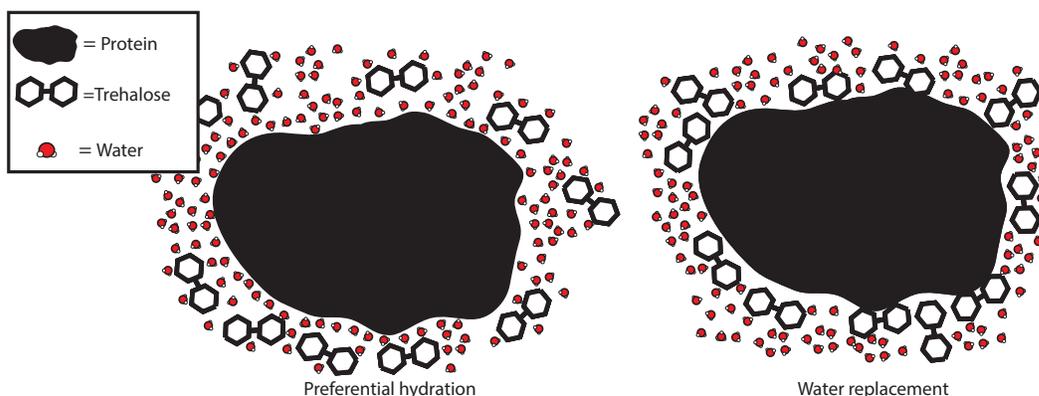


Figure 2.7: Sketch of preferential hydration model (left) and water replacement model (right). According to the preferential hydration model, the native structure of the protein is maintained through direct interaction between water and protein. In the water replacement theory, the hydration shell of the protein is partially replaced by trehalose molecules.

Preferential hydration model

Rather than direct interactions between trehalose and protein, it has been suggested that water molecules prefer to bind directly to the protein surfaces in a solution. This effect is highly important, considering that if a layer of water molecules surrounds the protein, its native solvated state is preserved, even though it may be embedded in a rigid trehalose matrix. The water entrapment theory is a special case of the preferential hydration theory, when the trehalose-protein matrix has a lower water concentration. It has been shown that this interaction remains even when the solution is freeze-dried,⁶⁹ indicating that trehalose entraps a hydration layer around the proteins (although some trehalose-protein interaction may become prevalent for cases of extreme desiccation). Support for this model was also found in this work, as presented in paper II.

Other important mechanisms

As some sugar molecules interact with amino acids they may each give of molecular constituents to form complexes at certain (mostly elevated) temperatures. This reaction is called the Maillard reaction and is the cause for the often desired browning of food, such as toasting of bread. If a sugar and a protein undergo a Maillard reaction, then it follows that there is a subsequent loss in the functionality of the protein.^{70,71} It has been shown⁷² that trehalose exhibits barely no Maillard reaction with amino acids at all, as opposed to e.g. sucrose which is highly prone to undergo this type of reaction. It has been proposed that this ability to avoid reactions between sugar and protein is one of the reasons behind the superior protective abilities of trehalose.

It has furthermore been shown that trehalose occupies a larger sample volume than similar disaccharides (37.5 v/v% for trehalose, 13 v/v% for sucrose, and 14 v/v% for maltose at a molar concentration of 1.5M),⁷³ which is coupled to the fact that trehalose forms larger hydrated volumes. It was indicated that if other protective co-solutes (such as sucrose or maltose) had the same volumetric concentration as trehalose then they will exhibit similar prevention of protein unfolding at the same level as trehalose, albeit at a much higher molarity*.⁷³

In summary, there are many different aspects to be concerned about regarding the question of why trehalose has such an excellent stabilizing ability. In the discussion above, just a few important hypotheses have been brought up which has been relevant to the current thesis; particularly the issue of the structure of aqueous trehalose or the issue of water replacement/entrapment-theories.

There are probably different protective mechanisms for different types of environmental stresses. For example, the mechanism that stabilizes during cryoprotection is not the same mechanism which is important for lyophilization. Some factors are however very important for a molecule to be able to stabilize

*The molarities required to reach the same volumetric concentration was in the order of 4M for sucrose and maltose. Since such concentrations of disaccharides are not soluble the unfolding prevention property was actually shown using glycerol with the same volume concentration as trehalose(37.5 v/v%)

biological material. It needs to be able to slow down the surrounding water dynamics and perturb the water structure, yielding a stable glass without crystallization of the water. It should also be non-reactive with the protein, and be able to protect many different types of biomolecules (such as both proteins and lipid bilayers). Trehalose seem to possess all these qualities. It interacts strongly with water, thus it easily perturb crystallization and enables glass formation. Moreover, it prefers to interact with water over protein, thus leaving the protein hydration layer intact. Trehalose has also been shown to remain in a glassy state by buffering water molecules, and furthermore does not undergo Maillard reactions with proteins.

3

Experimental and Computational Methods

In the work presented in this thesis, the main methods used have been neutron diffraction combined with empirical potential structure refinement (EPSR) modeling – used for the investigation of the structure of aqueous trehalose – and differential scanning calorimetry (DSC) for indirect structural investigations of aqueous trehalose containing protein. These three methods will be the main focus of the following chapter, containing a brief overview of the theory behind these techniques. The implementation of these techniques will be further discussed in chapter 4.

3.1 Neutron Scattering

The use of scattering techniques for the study of all types of materials (including biological material) have been around for a long time. The first images of a protein (myoglobin) for example was, as previously mentioned, produced by Kendrew et.al., using X-ray scattering.⁴⁰ Scattering methods in general has grown a lot in use during the past decades. For the investigation of biological materials (and many other kinds of materials) there are often a lot to be gained from using neutrons rather than X-rays, since the neutrons are electrically neutral and are capable of penetrating deep into the material. However, perhaps the biggest advantage of using neutrons is that hydrogen – which is extremely

abundant in biological materials – has a high cross section for neutron interaction, and is thus easier to study using neutrons. The cross section of atoms using X-rays is proportional to the number of electrons in each atom, and thus hydrogens are objects which are difficult to detect. The cross section of atoms as seen by neutrons however, vary more spuriously for different elements and isotopes. Furthermore, since the neutrons interact with atomic nuclei, this technique gives the possibility of performing isotope substitutions. By substituting e.g. hydrogen with deuterium, which has an extremely different total cross section, it is possible to alter the contrast between atoms depending on its isotope composition (see also section 4.1.3), without altering the structural and dynamical properties of the material substantially.

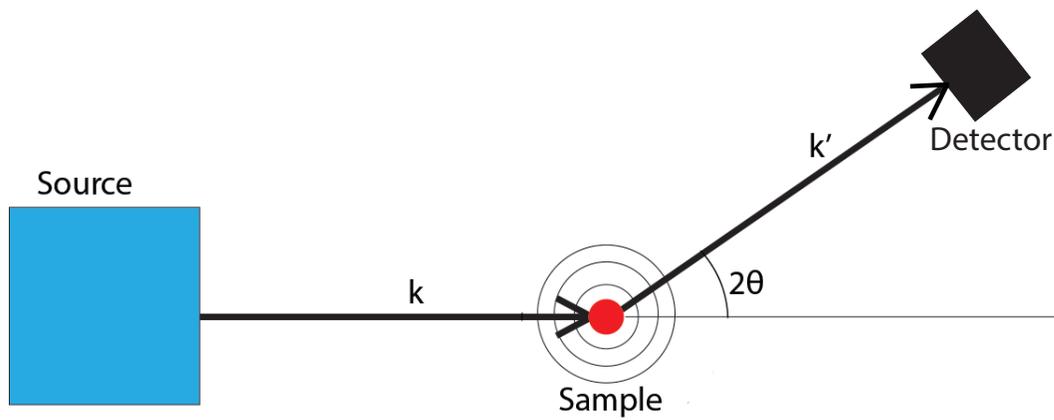


Figure 3.1: Schematic of a neutron scattering event. Incident neutrons impinge on the sample as a plane wave, with wave-vector k , and scatter radially with a wave-vector k'

A typical neutron scattering experiment starts by irradiation of a sample with a beam of neutrons. The neutrons are either produced in a reactor – where the fast neutrons are the by-product of radioactive decay from uranium (^{235}U) – or by a spallation source, in which high energetic protons are collided with a block of tungsten which then eject ("spall") neutrons in every direction. From a spallation source the neutrons are guided through e.g. moderators (decreasing the kinetic energy of the neutrons), collimators, monochromators, or other devices designed to select the desired properties of the neutrons destined to hit the sample. Some of the neutrons that hit the sample scatter and are

subsequently detected, and data are collected for further analysis.

There are many types of different neutron scattering techniques, which detect or focus on different properties of the scattered neutrons. In the present work, diffraction experiments, which mainly yield information of the structure of the investigated materials, have been used.

3.1.1 Neutron Diffraction

Neutron diffraction is an experimental technique where one focus on the elastic and coherent part of the scattering. The scattered neutrons mainly contain information about structural correlation lengths within the sample. In the case of a neutron diffraction experiment, such as the one performed in this work (performed at NIMROD, see section 4.1.1 for further details), the sample is hit by a pulse of neutrons with wave-vectors \mathbf{k} , with neutron wavelengths of λ_i (ranging from 0.05Å to 10Å), and scatter with wave-vectors \mathbf{k}' . In a neutron scattering experiment, the sample chamber is surrounded by detectors which essentially counts the number of neutrons scattered at different angles, and also at different arrival times. The scattered neutron wavelengths at a particular angle is calculated from their times of flight. Thus, the raw data produced is an intensity distribution function, $I(2\theta, \lambda)$, that relates the neutron intensity to a specific angle and wavelength. Typically, $I(2\theta, \lambda)$, is typically written as a function of $I(\mathbf{Q})$, where \mathbf{Q} is the scattering vector defined as*:

$$\mathbf{Q} = \mathbf{k} - \mathbf{k}' \quad (3.1)$$

where the relationship between $|\mathbf{Q}| = Q$ and θ can be found by approximating the scattering to just be elastic scattering, such that $|\mathbf{k}| = |\mathbf{k}'| = \frac{2\pi}{\lambda_i}$. From that approximation and simple trigonometry one can obtain:

$$Q = \frac{4\pi}{\lambda} \sin(\theta) \quad (3.2)$$

So how does this scattered intensity relates to the structure of the material? Recall section 2.1.2 where the pair correlation function $g(r)$ was introduced,

*The derivation of the formalism in this section can be found in multiple textbooks. For more details the reader is referred to e.g. Ref. 74–77

which is perhaps the simplest way to represent the structure of an amorphous material. In the remaining part of this section a brief outline is given on how to get from the measured diffracted neutron data to $g(r)$.

First of all the obtained raw data from the diffraction experiment have to be corrected for different background signals, neutron absorption events, etc., which is given an overview of in section 4.1.2. From those corrections, the double differential scattering cross section $\frac{d^2\sigma}{d\Omega dE}(Q)$ is obtained, which is the number of neutrons scattered into a solid angle element $d\Omega$ at an angle corresponding to scattering vector Q , with an energy between E and $E + dE$. When studying elastic scattering, the energy part is integrated out, according to equation 3.3, and the double differential scattering cross section can thus be reduced to the differential scattering cross section $\frac{d\sigma}{d\Omega}(Q)$.

$$\frac{d\sigma}{d\Omega} = \int_{-\infty}^{\infty} \frac{d^2\sigma}{d\Omega dE}(Q) dE \quad (3.3)$$

More about how to account for inelastic effects can be found in e.g. Ref. 78 or 79, however for the present discussion this approximation holds, and the discussion will focus on the static differential scattering cross section: $\frac{d\sigma}{d\Omega}(Q)$

So what is then the origin of this differential scattering cross section? Let us consider the wave formalism for neutrons for a while*; the neutron beam impinges on the sample as a plane wave (collimated beam) with wave-vector \mathbf{k} , and can be written as $\psi_i = \psi_0 \exp(i\mathbf{k} \cdot \mathbf{r})$ at position \mathbf{r} . ψ_0 is the amplitude of the incident wave related to the flux of the beam. After interacting with a single nucleus labeled j at position \mathbf{r}_j , the outgoing wave will propagate radially outwards from the nucleus as:

$$\psi_f = -\psi_0 e^{i\mathbf{k} \cdot \mathbf{r}_j} b_j \frac{e^{i\mathbf{k}' \cdot (\mathbf{r} - \mathbf{r}_j)}}{|\mathbf{r} - \mathbf{r}_j|} \quad (3.4)$$

where b is the so called scattering length of the nucleus, related to the cross section ($\sigma = 4\pi|b|^2$, measured in barns) of the nucleus, and is a measure of how strong the interaction between the incident neutron and the nuclei is. Summing up the contributions from all N atoms in the sample, one obtains the total wave

*For simplicity, the formalism described here applies for single component systems.

function:

$$\psi_f = -\psi_0 e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{j=1}^N b_j \frac{e^{i\mathbf{Q}\cdot\mathbf{r}_j}}{|\mathbf{r} - \mathbf{r}_j|} \quad (3.5)$$

where again $\mathbf{Q} = \mathbf{k} - \mathbf{k}'$. This wave function describes the amplitude of the scattered wave at different positions. The position of each atom could in principle be obtained through a Fourier transform of this function. However, what is measured at the detector is not the wave itself, but the square modulus of the wave function, $|\psi_f|^2$.

$$|\psi_f|^2 \approx \frac{|\psi_0|^2}{r^2} \sum_{i=1}^N b_i e^{i\mathbf{Q}\cdot\mathbf{r}_i} \cdot \sum_{j=1}^N b_j^* e^{-i\mathbf{Q}\cdot\mathbf{r}_j} \quad (3.6)$$

In equation 3.6, it was assumed that the distance between the detector and the sample is much greater than the distances between nuclei in the sample ($|\mathbf{r}| \gg |\mathbf{r}_j|$), and thus $|\mathbf{r} - \mathbf{r}_j| \approx r$.

This equation can furthermore be rewritten in terms of the differential scattering cross section by noting that the number of neutrons impinging on a small area, dA , is $|\psi_f|^2 dA$. The differential scattering cross section can thereby be written as the fraction of incident neutrons per solid angle $d\Omega$:

$$\frac{d\sigma}{d\Omega} = \frac{|\psi_f|^2 dA}{|\psi_0|^2 d\Omega} = \left\{ d\Omega = \frac{dA}{r^2} \right\} = \sum_{i=1}^N b_i e^{i\mathbf{Q}\cdot\mathbf{r}_i} \cdot \sum_{j=1}^N b_j^* e^{-i\mathbf{Q}\cdot\mathbf{r}_j} \quad (3.7)$$

Rewriting equation 3.7 and averaging the nuclear scattering length, b , over different spin orientations and isotopes the following equation can be obtained:

$$\frac{d\sigma}{d\Omega} = \sum_{i,j}^N \langle b_i b_j^* \rangle e^{i\mathbf{Q}\cdot(\mathbf{r}_i - \mathbf{r}_j)} \quad (3.8)$$

It is useful to separate the term $\langle b_i b_j^* \rangle$ into two different cases as well, either $i \neq j$ or $i = j$. If $i \neq j$ then the average is taken over two different atoms which do not have any correlation in their scattering lengths. Hence for $i \neq j$ then $\langle b_i b_j^* \rangle = \langle b_i \rangle \langle b_j^* \rangle = \langle b \rangle^2$. In the case when $i = j$ however, the scattering refers to the "self-scattering" of individual nuclei. The term $\langle b_i b_j^* \rangle$ then becomes: $\langle b^2 \rangle = \langle b \rangle^2 + \langle (b - \langle b \rangle)^2 \rangle$, i.e. a measure of how much the scattering length deviates from the mean value.⁷⁴

Using these scattering lengths and that for $i = j$ then $r_i = r_j$, and equation 3.8 can thus be written as:

$$\frac{d\sigma}{d\Omega} = \underbrace{\langle b \rangle^2 \sum_{i,j}^N e^{i\mathbf{Q}\cdot(\mathbf{r}_i - \mathbf{r}_j)}}_{\text{Coherent scattering}} + \underbrace{N(\langle b^2 \rangle - \langle b \rangle^2)}_{\text{Incoherent scattering}} \quad (3.9)$$

The incoherent scattering part is clearly not dependent on any structural parameters concerning the sample, and is thus merely added to the total scattering cross section as an (often unwanted) background in the case of a diffraction study. The coherent part however is typically written as:

$$\frac{d\sigma}{d\Omega_{coh}} = \langle b \rangle^2 N S_{coh}(\mathbf{Q}) \quad (3.10)$$

Where $S_{coh}(\mathbf{Q}) = \frac{1}{N} \sum_{i,j}^N e^{i\mathbf{Q}\cdot(\mathbf{r}_i - \mathbf{r}_j)}$ is the coherent structure factor (from now on just called the structure factor, or $S(\mathbf{Q})$), which is a function describing the system that only depends on the investigated material. Using a general property of the Dirac delta function that $\sum_i e^{i\mathbf{Q}\cdot\mathbf{r}_i} = \int_V e^{i\mathbf{Q}\cdot\mathbf{r}_i} \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$, $S(\mathbf{Q})$ can be rewritten as:

$$S(\mathbf{Q}) = 1 + \frac{1}{N} \int e^{-i\mathbf{Q}\cdot\mathbf{r}} \sum_{i,j \neq i}^N \delta(\mathbf{r} - (\mathbf{r}_i - \mathbf{r}_j)) d\mathbf{r} \quad (3.11)$$

The summation of Dirac delta functions in equation 3.11 is familiar from section 2.1.2. By recalling the definition of $g(r)$ through equation 2.3 and inserting this into equation 3.11, the following equation is obtained:⁸⁰

$$S(\mathbf{Q}) = 1 + \rho \int e^{-i\mathbf{Q}\cdot\mathbf{r}} g(\mathbf{r}) d\mathbf{r} \quad (3.12)$$

which, for an isotropic fluid can be simplified to:

$$S(Q) = 1 + 4\pi\rho \int_0^\infty r^2 g(r) \frac{\sin(Qr)}{Qr} dr \quad (3.13)$$

It can thus be concluded that there exist a relationship between the obtained neutron diffraction data and the pair correlation function.

Partial structure factors

The structure factor $S(Q)$ describes the total coherent elastic scattering of the sample. However, it is sometimes useful to divide $S(Q)$ into a sum of contributions arising from the correlations between different atom types. Similar to what was done for the total pair correlation function into partial pair correlation functions in section 2.1.2. The individual terms in $S(Q)$ is called partial structure factors and are usually written as $S_{\alpha\beta}(Q)$, where α and β represents different atoms. They are defined through:

$$S(Q) = \frac{\sum_{\alpha\beta} \langle b_\alpha \rangle \langle b_\beta \rangle (S_{\alpha\beta}(Q) - 1)}{\sum_\alpha c_\alpha \langle b_\alpha \rangle^2} + 1 \quad (3.14)$$

The partial structure factors in turn relate to the correlations between atom types α and β via the partial pair correlation functions $g_{\alpha\beta}$ through:

$$S_{\alpha\beta}(Q) = c_\alpha \delta_{\alpha\beta} + c_\alpha c_\beta \rho \int e^{-i\mathbf{Q}\cdot\mathbf{r}} g_{\alpha\beta}(r) d\mathbf{r} \quad (3.15)$$

3.2 Empirical Potential Structure Refinement Modeling

In principle, a full set of partial pair correlation functions of a material can be obtained by isotopically labelling every unique atom type in the material, exchanging each atom type with a different isotope at the time. However, this method would be very time consuming, and such accurate partial isotopical marking is rarely available. Instead, several methods of computer modeling have been developed in order to obtain a complete model of a sample from the available diffraction data.

One such method, used in the work of this thesis, is the EPSR method. To describe this method it is useful to start with a previous method which EPSR is based on, namely the Reverse Monte Carlo (RMC) method.

3.2.1 RMC

In the RMC method (described in more detail in Ref. 81) the researchers typically set up a system where the structure of the molecules is initially de-

finer, and then a hard-sphere model is applied for the interactions between the molecules. A set of constraints are typically also set up, such as minimum intermolecular distances, or specific bond angles, as determined from other experimental methods (e.g. NMR, MD). Such constraints are necessary for avoiding unreliable outcomes, by for example preventing atomic overlaps.

The RMC program then randomly moves around an atom or a molecule and calculates e.g. the structure factors, and compares them with the experimental structure factors. If the move decreased the difference between the two corresponding structure factors, the move is accepted. If it increased the difference, the move is only accepted with a probability of:

$$\exp\left(-(\chi_{before\ move}^2 - \chi_{after\ move}^2)/2\right) \quad (3.16)$$

where χ^2 is a quantity of the difference between the simulated and experimental data, and defined as:

$$\chi^2 = \sum_{i=1} [S_{calc}(Q_i) - S_{exp}(Q_i)]^2 / \sigma^2(Q_i) \quad (3.17)$$

where $\sigma(Q)$ is the standard deviation of experimental error for any measured value of Q .

RMC modeling is a widely used method which has been used for decades, and has helped to solve a significant number of molecular structures of disordered materials.⁸² However, RMC fails to take into account several physical aspects of a material when just searching for a structure which fits the experimental data, such as the fact that the obtained structure could in principle have a molecular potential which is energetically unfavorable.⁸¹

3.2.2 EPSR

The EPSR method used in this work is derived from the RMC method, and just like RMC, EPSR seeks to obtain a minimum difference between the derived data and a set of experimental data. This is opposed to the Metropolis Monte Carlo (MMC) method which is an efficient method of obtaining the minimum intermolecular potential energy of a system. However, there are some important differences; in EPSR, the goal is to obtain a correct intermolecular potential

which leads to convergence with the experimental data, rather than fitting the simulated structure directly to the diffraction data.

The procedure works in principle according to the following scheme:⁸³

1. Create a simulation box with the correct density and intramolecular structure.
2. Assign a reference potential (RP) to the system, i.e. find Lennard-Jones parameters for each atom type.
3. Run a MMC simulation on the system to minimise the potential energy. This is done by randomly moving (translation, rotation, or bending) an atom or molecule and subsequently measuring the potential difference (ΔU) due to that move. The move is always accepted if $\Delta U < 0$, and only accepted with a probability of $\exp[-\frac{\Delta U}{k_B T}]$ if $\Delta U > 0$.
4. From the energetically minimized structure, the pair correlation function is calculated. This is subsequently Fourier transformed to give the simulated structure factor, $S_{calc}(Q)$.
5. The difference, $S_{calc}(Q) - S_{exp}(Q)$ is then calculated and used to calculate the empirical potential (EP).
6. The empirical potential is then added to the reference potential.
7. Steps 3 - 6 are then iterated until the EP becomes stable, or until its absolute energy exceeds a predefined value.

To start an EPSR simulation, first one needs to find a reference potential. This includes different Lennard-Jones parameters but also intramolecular structures, intermolecular starting configurations and minimal distances. Typically, one has a lot of knowledge about a material determined through previous experiments which can be effectively included in the RP. The Lennard-Jones (LJ) parameters sets up an often used potential (Lennard-Jones potential) between two atoms and is the basis of the reference potential combined with an added Coloumb potential:

$$U_{\alpha,\beta}^{LJ} = 4\epsilon_{\alpha\beta} \left[\left(\frac{\sigma_{\alpha\beta}}{r} \right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{r} \right)^6 \right] + \frac{q_{\alpha}q_{\beta}}{4\pi\epsilon_0 r} \quad (3.18)$$

Where r is the distance between atom α and atom β , q is the electric charge of a specific atom, $\sigma_{\alpha\beta}$ is the distance where the potential is zero, and ϵ is the depth of the potential well. $\sigma_{\alpha\beta}$ and $\epsilon_{\alpha\beta}$ are calculated according to the Lorentz-Berthelot mixing rules, based on the LJ parameters of the individual atoms.

$$\epsilon_{\alpha\beta} = \sqrt{\epsilon_{\alpha}\epsilon_{\beta}} \quad \sigma_{\alpha\beta} = \frac{\sigma_{\alpha} + \sigma_{\beta}}{2} \quad (3.19)$$

The individual LJ parameters in turn can be determined through various force-field calculations such as the OPLS-AA (Optimized Potentials for Liquid Simulations - All Atoms) force field which was used in paper I.

Further additions to the reference potential is well described in the EPSR manual (Ref. 83) and Ref. 82.

Below is given an overview of how the empirical potential (EP) is defined. For a more detailed description about these steps, the reader is again referred to Ref. 83 and Ref. 82. The empirical potential in real space is defined as a sum of Poisson functions:

$$U(r)^{EP} = kT \sum_i C_i p_{n_i}(r, \sigma) \quad (3.20)$$

where

$$p_{n_i}(r, \sigma) = \frac{1}{4\pi\sigma^3(n+2)} \left(\frac{r}{\sigma}\right)^n e^{-\frac{r}{\sigma}} \quad (3.21)$$

and $n_i = \frac{r_i}{\sigma} - 3$ (r_i and σ is set by the user). C_i :s are weights which are fitted through comparison with real data. $p_{n_i}(r, \sigma)$ can be Fourier-transformed to:

$$P_n(Q, \sigma) = \frac{1}{\sqrt{1+Q^2\sigma^2}^{(n+4)}(n+2)} \left[2\cos(n\alpha) + \frac{(1-Q^2\sigma^2)}{Q\sigma} \sin(n\alpha) \right] \quad (3.22)$$

where $\alpha = \arctan(Q\sigma)$. The empirical potential can then be written in Q-space accordingly:

$$U(Q)^{EP} = \sum_i C_i P_{n_i}(r, \sigma_Q) \quad (3.23)$$

The C_i -weights are determined by fitting $U(Q)^{EP}$ to the difference between $S_{exp}(Q)$ and $S_{calc}(Q)$. Once these go to small values (ideally to zero), the EP

converges and thus a final total potential is obtained. From this stage it is possible to start analyzing the obtained model and extract useful structural information.

3.3 Differential Scanning Calorimetry

Differential Scanning Calorimetry (DSC) is a method for determining different types of thermal events in a material. The principle behind this method is to change the temperature of a sample and measuring the involved heat (enthalpy). This process enables the experimenter to detect e.g. glass transitions, crystallization, denaturation, etc.

Inside the DSC cell there are two sample platforms; one for the sample and one for a reference sample. The reference sample is typically an empty sample holder, identical to the one holding the sample under investigation. When the temperature is set to change a certain rate ΔT , the sensors on the sample and on the reference sample registers the difference in heat flow to and from the two samples ΔQ . The difference ΔQ corresponds to the amount of energy required to change the temperature of the sample by ΔT . Another important investigated property is the specific heat capacity, obtained after normalization with respect to e.g. sample mass and baseline. The heat flow to and from a sample is typically constant when no physical or chemical changes occur. However, when such events do occur, they show up in the signal in different ways depending on the nature of the event. In Figure 3.2 some events are pointed out that are relevant for the present thesis and will be briefly discussed here.

1. **Crystallization.** When the sample crystallizes it transforms from a highly disordered state to a more ordered one, the entropy (S) decreases and thus the process is exothermic, which shows up as a positive peak as indicated in Figure 3.2.
2. **Melting.** Opposite of crystallization, the disorder of the sample is increased more energy is required to break the crystal structure, which shows up as an endothermic dip in the DSC curve.

3. **Glass transition.** The glass transition, is a second-order process (explained in section 2.1), and it is thus associated with a change in the heat capacity of the sample. It can be seen as a step in the baseline of the heat flow.
4. **Denaturation.** During the denaturation process the hydrogen bonds keeping the protein in its functional state are broken. When this happens the protein can easily unfold and aggregate. This is an irreversible process and shows up as an endothermic peak in the DSC scan.

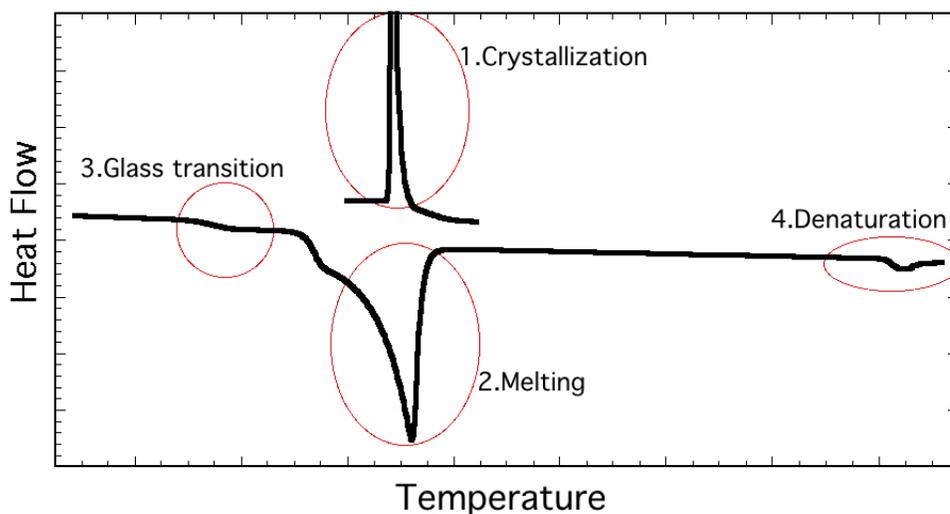


Figure 3.2: Typical DSC scan of a sample containing protein and trehalose. (1) Sample partially crystallize during cooling. (2) During heating, the crystalline part of the sample melts. (3) A part of the sample exhibits a glass transition. (4) The protein denatures.

4

Experimental Procedures

4.1 Neutron Diffraction Experiments

4.1.1 Diffractometer

The diffraction data presented in paper I were obtained using the NIMROD (Near and InterMediate Range Order Diffractometer) diffractometer at the spallation source ISIS, Rutherford Appleton Laboratory, UK. The incident neutrons at NIMROD arrive to the sample in wavelengths between 0.05–10Å, and it has a broad Q-range of 0.01–50 Å⁻¹. This broad Q-range makes NIMROD specialized in measuring a very wide range of length-scales, from less than 1Å, up to more than 300Å. It is thus possible to combine structural data from the microscopic range (through the large scattering angles) to the mesoscopic range (via the small scattering angles). This makes NIMROD suitable for probing large-scale structures inside a disordered medium, such as liquids or macromolecules in solutions.

4.1.2 Data Corrections

To obtain useful structure factors from the raw neutron diffraction data, the GUDRUN data correction software has been used in the work presented in this thesis. More details about the discussed corrections can be found in Ref. 11.

Neutron diffraction data is typically normalized using a vanadium plate since it exhibits almost only incoherent scattering. Therefore, the scattering

from the vanadium plate does not depend on Q , and thus show up as a uniform background signal for a given incident neutron flux. This background signal is used to give the data an absolute scale, since the theoretical scattering of the vanadium plate is relatively easy to calculate. The number of detected neutrons at each detector can thus be divided by the number of theoretically detected neutrons to give the normalization factor for that particular detector.

Measurements are normally also made on an empty sample holder, in order to subtract the signal from the sample holder from the measured data.

- **Multiple scattering corrections.**

In the theory presented about neutron scattering above, it is assumed that a scattered neutron travels directly from the sample to the detector. This is however merely an approximation; multiple scattering events are possible, and sometimes this has to be taken into account. The probability that a neutron is scattered multiple times can be calculated from the atomic composition of the sample and the sample geometry. From these parameters, a background signal can be calculated and subsequently removed from the final structure factor.

- **Absorption corrections.** Another approximation in the theory presented in section 3.1.1 is that there is no absorption of neutrons in the sample. The cross section of a sample was presented as a *scattering* cross section. However in reality, the total cross section should include the absorption cross section ($\sigma_{tot} = \sigma_{scattering} + \sigma_{absorption}$). For the energies of the incident neutrons in the present experiments, the absorption cross section is assumed to have a linear dependency to the wavelength of the incident neutrons, which is a good approximation.¹¹ Similar to the situation with the multiple scattering corrections, the absorption is estimated from the atomic composition of the sample plus sample holder, and their combined geometry. The atoms are furthermore approximated to be isotropically distributed within the sample (as well as in the sample holder).¹¹

- **Deadtime corrections.**

Deadtime corrections are necessary due to the "deadtime" of the detectors. When a detector is hit by a neutron, it requires some time before it is able to count the next one. Thus, they may count multiple hits as just one if they are unable to resolve these in time. GUDRUN takes this into account and compensates for such effects.

4.1.3 Isotope substitution

As previously discussed in section 3.1, the structure factor, $S(Q)$, can be expressed as a linear combination of the different partial structure factors, $S_{\alpha\beta}(Q)$, as in equation 3.14. In theory, this equation can be inverted, with the help of substitution of isotopes. Given that a molecular system has the same structure for different isotopes, the only thing differing in the obtained structure factors is the scattering lengths. By choosing appropriate substitutions different partial structure factors can therefore be emphasized or hidden, thus in total yielding a full set of partial structure factors (which subsequently generates the partial pair correlation functions through Fourier transformation). The number of substitutions required to fully solve this in a sample of n different chemical species is the same as the number of different partial structure factors that can be generated, i.e. $\frac{n(n+1)}{2}$.

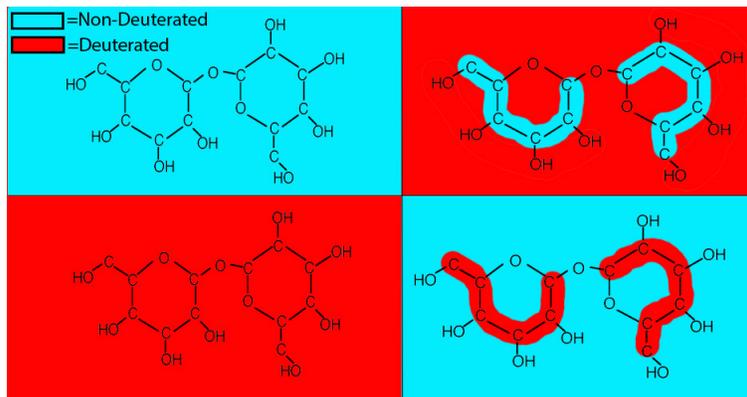


Figure 4.1: Sketch showing how isotope substitution gives contrast to particular sets of atoms in a sample.

4.2 EPSR analysis

4.2.1 Coordination numbers

The coordination number, $n_{\beta}^{\alpha}(r_1, r_2)$, which gives the number of atoms of type β surrounding a central atom type α within a radial distance between r_1 and r_2 , is calculated using the partial pair correlation functions $g_{\alpha\beta}(r)$:

$$n_{\beta}^{\alpha}(r_1, r_2) = 4\pi c_{\beta} \int_{r_1}^{r_2} g_{\alpha\beta}(r) r^2 dr \quad (4.1)$$

where c_{β} is the atomic number density of atom β , and $g_{\alpha\beta}(r)$ is defined through equation 2.4.

This number is in general reported for different coordination shells (first, second, etc.), where e.g. the first coordination shell is typically defined as the coordination number between $r_1 = 0$ and r_2 . r_2 is the distance where the first minima appear in $g_{\alpha\beta}(r)$ after the first obvious peak (in the O-O correlation for water in Figure 2.4 for example, r_2 would be approximately 3.4Å).

4.2.2 Clusters

In order to calculate the cluster size distribution in the molecular model, the cluster-subroutine in EPSR was used. This subroutine counts the number of molecules in each cluster, ranging from a non-clustered molecule with a cluster size of 1, to a complete clustering, where all specified molecules in the model are connected to the same cluster. In paper I the definition for two trehalose molecules to be considered clustered was that the minimum distance between an oxygen and a hydrogen from different trehalose molecules was 2.5Å. If yet another trehalose bind to one of the already clustered trehalose molecules, the cluster-size for this cluster grow with one unit.

4.2.3 Angle distributions

Another important EPSR subroutine used in this work is one that counts various angle distributions. Within this subroutine the user defines triplets of atoms, and maximum distance criteria for these atoms. The program then measures the angle between every triplet of atoms that satisfies the distance

criteria. This subroutine was used to investigate the geometric configuration of the water structure in paper I (see Figure 4 and 5, paper I).

4.2.4 Hydrogen bonding

The number of intermolecular hydrogen bonds between water and trehalose were calculated for the structures obtained from the EPSR simulation. The calculation was done using different sets of criteria for different definitions of a hydrogen bond. Previous EPSR studies (such as Ref. 84 and 85) in the literature have used a maximum distance criterion, i.e. a hydrogen bond is defined by that the distance between a donor hydrogen and an acceptor oxygen is less than 2.5\AA . However various MD-studies (e.g. Ref. 30 or 86) typically employ a distance criteria combined with an angular criterion. The most common setup for these two criteria is that the minimum distance between an acceptor oxygen and a donor oxygen is less than 3.4\AA and that the O - H - - - O angle is maximum 120° . Other criteria were also used in paper I in order to compare with previous studies.

4.3 DSC Experiments

All DSC experiments for this thesis were performed on a DSC Q1000 (TA Instruments). For the measurements, each of the samples were placed in hermetically sealed aluminum pans, which were placed on the sample podium within the DSC-sample chamber. The DSC-sample chamber is coupled to a liquid nitrogen cooling system, which allows for cooling to temperatures of -180°C . The DSC-sample chamber is also connected to a thermal element, which could heat the samples up to 550°C . The temperature and the heat flow to and from the reference and the sample are measured independently.

Before the measurements, heat capacity calibration was performed using two sapphire disks (one for the reference podium and one for the sample podium). By scanning the heat flow and the heating rate for a large span of temperatures it is possible to find the calibration constant, K , defined as $C_p = K \frac{\text{Heatflow}}{\text{HeatingRate}}$,⁸⁷ since the heat capacity of sapphire (C_p) is well-known and stable. A temperature and enthalpy calibration was also performed by

heating three standard substances (indium, water, and mercury) over their melting temperatures. The melting temperatures and heat of fusion for these substances are well known, and thus any deviations in temperature or enthalpy measurements can be corrected for by using these calibration constants.

5

Summary of Appended Papers

Paper I: Structure of aqueous trehalose solution by neutron diffraction and structural modeling

In paper I, some of the fundamental properties of the structure of aqueous trehalose were investigated. Many of the previous reported studies on these properties have been done by e.g. MD simulations (see for example Ref. 30 and 31). Such studies are sometimes limited by the approximative force fields, which do not necessarily produce correct structures. For the work in this study, the structure was directly investigated by using neutron diffraction. The data were collected on the NIMROD neutron diffractometer, and were thereafter analyzed using EPSR simulations. A similar study has previously been performed,⁸⁸ however with a different isotope substitution scheme than the one performed in this study. In Ref. 88, the authors deuterated the exchangeable trehalose hydrogens, which thereafter were dissolved in H₂O. However, the presence of H₂O makes the deuterated hydrogen groups exchange back to hydrogen, which was not taken into account. In paper I, when deuteration of the exchangeable hydrogens was done, the trehalose was subsequently dissolved in D₂O so that no further exchanges altered the desired isotope configuration. Deuteration was furthermore done on the non-exchangeable hydrogen-groups, which were subsequently dissolved in either D₂O or H₂O, without any further exchanges occurring.

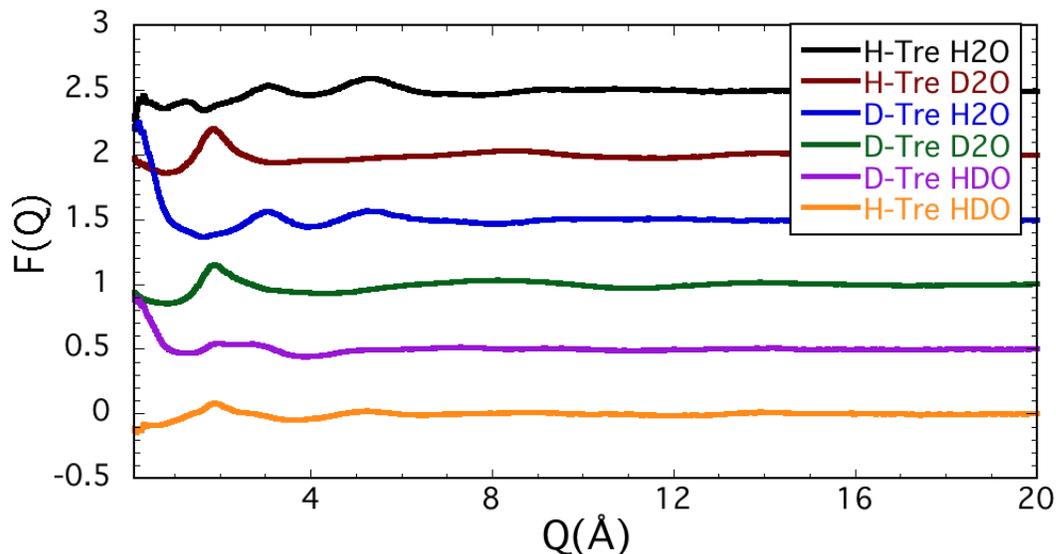


Figure 5.1: Neutron diffraction data

The EPSR modeling was performed by building a simulation box of 2000 water molecules and 52 trehalose molecules. Analysis of the structure of water in the presence of trehalose indicated that the water structure is significantly perturbed, most likely due to that trehalose exposes a large portion of its potential hydrogen-bonding sites for the water molecules to bind to. This finding – which is opposing the work done by Ref. 88 – presumably also correlates with the low degree of clustering of trehalose molecules. Previous studies^{30,31} have reported more trehalose clustering than was found in the present study, and this lack of clustering may play a large role in explaining the peculiarity of trehalose.

Paper II: The Role of Trehalose for the Stabilization of Proteins

In paper II, the role of trehalose during the stabilization process was investigated by differential scanning calorimetry. The samples studied were composed of mixtures of water, trehalose, and myoglobin, at a wide range of concentrations, typically at a water concentration less than 75wt%. Particularly, the glass transition temperatures and the denaturation temperatures were studied to correlate the stability of the glassy matrix (as determined by T_g) with the

stability of the protein (determined by T_{den}). It was shown that more trehalose was correlated with both an increment in T_g and in T_{den} . Samples which did not form ice exhibited a positive correlation between T_g and an increase in protein concentration, but a negative correlation if the sample did form ice. This is explained by that the protein effectively dries out the glassy trehalose–protein matrix when there is no ice formation, which thus raises T_g . If ice is formed however, the amorphous part of the sample becomes freeze-concentrated by expelling any excess water molecules. In this freeze-concentrated part, T_g is determined by the protein:trehalose ratio, and the protein has a lower T_g than trehalose.

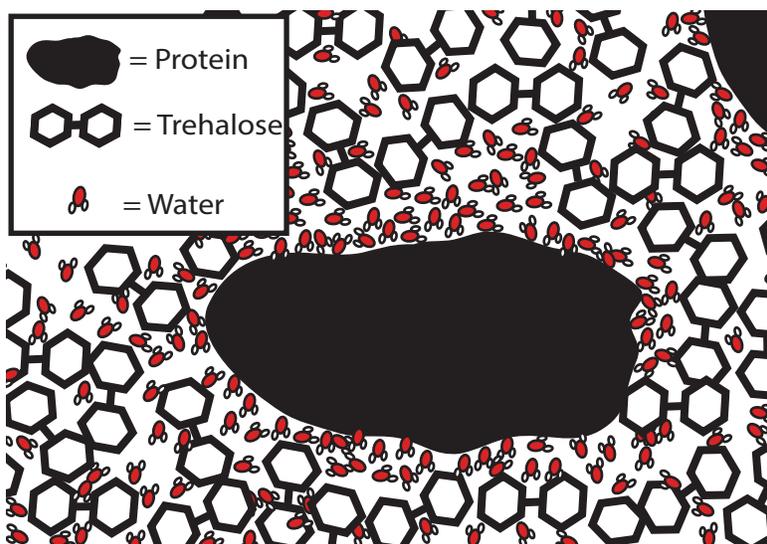


Figure 5.2: Sketch of preferential hydration interpretation

By determining the maximum water concentration before the sample exhibit crystallization, for different trehalose:protein ratios, it was found that water preferentially adsorbs to the protein surface. Thus, the results indicate that the preferential hydration model (as sketched in Figure 5.2 and explained in section 2.5.3) is more likely than the water replacement model, for the investigated concentrations.

6

Conclusions and Outlook

The particular conformation of a trehalose molecule causes the hydroxyl groups of the trehalose molecule to prefer interacting with water molecules (hydrogen bonding), rather than other trehalose hydroxyl groups (both internal and external). This means that trehalose avoids clustering and folding, yielding more interactions with its environment (e.g. Ref. 30 or paper I). Generally, this is not a unique property; many molecules provide a large amount of HB-sites. Trehalose possess chemical properties beyond that. It creates a homogeneous stabilizing environment,^{29,30} and it is flexible enough that it can adapt to complex surfaces, such as proteins,³⁰ thus forming a great scaffold. Furthermore, trehalose tend to not bind directly to protein surfaces, and traps a layer of water at those surfaces instead (see e.g. Ref. 69 or paper II). This means that trehalose does not enforce its own structure on to the protein surface, which – if it occurred – could result in a destabilization of the entire protein.

Although indirect evidence was found for a preferential hydration model – and some previous studies in the literature propose a similar picture as the one given in paper II – it should be noted that a direct observation of such structures has not yet been made. Currently, I am working on such data from small to wide-angle neutron diffraction measurements. That study, in combination with a quasi-elastic neutron scattering experiments, and a collaboration with an MD-study, should provide a complete picture of how trehalose interact with protein in an aqueous solution. Thus, obtaining both the structure and the dynamics of such a system.

A more complete picture of how trehalose acts upon biological material, at preservation conditions, may result in a more general understanding of biomolecular stabilization mechanisms. This can lead to improved cryopreservation and lyophilization procedures, which would be vastly beneficial to a large number of fields.

Glossary

| | |
|---------------|---|
| DSC | D ifferential S canning C alorimetry |
| EPSR | E mpirical P otential S tructure R efinement |
| HB | H ydrogen B ond |
| MMC | M etropolis M onte C arlo |
| NIMROD | N ear to I nter M ediate R ange O rders D iffractometer |
| RMC | R everse M onte C arlo |
| RP | R eference P otential |
| T_g | Glass transition temperature |
| T_d | Denaturation temperature |

Acknowledgements

I would like to start by thanking my supervisor, Jan Swenson, for all his advice and guidance, and also for hiring me to work on this fascinating topic. I also want to thank my co-supervisor, Helén Jansson, for all her support and great discussions. Also, a big thank you to everyone else who contributed to the scientific part.

Furthermore I would like to thank all the people at KMF for making the department a great workplace. Thanks to all the present and former office mates for interesting and fun conversations: Khalid, Fredrik, Manso, Viktor, Rasmus, and Ronny. I would also like to thank Ezio for helping me with basically everything in the lab.

Finally, I want to express my gratitude to my friends and family for all your support, and of course to Anna for being there and adding well-needed silliness to my life.

Christoffer

Bibliography

- [1] A. Liesowska, *Siberian scientists announce they now have a 'high chance' to clone the woolly mammoth*, <http://siberiantimes.com/science/casestudy/news/exclusive-siberian-scientists-announce-they-now-have-a-high-chance-to-clone-the-extinct-woolly-mammoth/>, Accessed on: Jul 15, 2016.
- [2] J. O. Karlsson and M. Toner, *Biomaterials*, 1996, **17**, 243 – 256.
- [3] J. G. Day and G. Stacey, *Cryopreservation and freeze-drying protocols*, Springer Science & Business Media, 2007, vol. 368.
- [4] D. E. Pegg, *Cryopreservation and freeze-drying protocols*, 2007, 39–57.
- [5] G. Adams, *Cryopreservation and Freeze-Drying Protocols*, 2007, 15–38.
- [6] P. Matejtschuk, *Cryopreservation and Freeze-Drying Protocols*, 2007, 59–72.
- [7] A. H. Haines, *Organic & biomolecular chemistry*, 2006, **4**, 702–706.
- [8] N. K. Jain and I. Roy, *Protein Science*, 2009, **18**, 24–36.
- [9] S. R. Elliott, *Physics of amorphous materials*, Longman Scientific and Technical, UK, 1990.
- [10] R. A. Jones, *Soft condensed matter*, Oxford University Press, 2002, vol. 6.

- [11] A. K. Soper, *GudrunN and GudrunX - Programs for correcting raw neutron and x-ray total scattering data to differential cross section*, 2012.
- [12] W. Kauzmann and D. Eisenberg, *The structure and properties of water*, Clarendon Press, 1969.
- [13] A. Narten, M. Danford and H. Levy, *Discussions of the Faraday Society*, 1967, **43**, 97–107.
- [14] H. E. Stanley, *MRS bulletin*, 1999, **24**, 22–30.
- [15] P. Wernet, D. Nordlund, U. Bergmann, M. Cavalleri, M. Odelius, H. Ogasawara, L. Näslund, T. Hirsch, L. Ojamäe and P. Glatzel, *Science*, 2004, **304**, 995–999.
- [16] A. K. Soper, *Water and ice structure in the range 220–365 K from radiation total scattering experiments*, 2014.
- [17] A. K. Soper, *Chemical Physics*, 2000, **258**, 121–137.
- [18] D. Dhabal, M. Singh, K. T. Wikfeldt and C. Chakravarty, *The Journal of Chemical Physics*, 2014, **141**, 174504.
- [19] E. Duboue-Dijon and D. Laage, *The Journal of Physical Chemistry B*, 2015, **119**, 8406–8418.
- [20] *ISIS Disordered Materials Database*, <http://www.isis.stfc.ac.uk/groups/disordered-materials/database/>, Accessed on: Jun 10, 2016.
- [21] N. Møbjerg, K. A. Halberg, A. Jørgensen, D. Persson, M. Bjørn, H. Ramløv and R. M. Kristensen, *Acta Physiologica*, 2011, **202**, 409–420.
- [22] G. G. Birch, *Advances in carbohydrate chemistry*, 1963, **18**, 201–225.
- [23] J. L. Green and C. A. Angell, *The Journal of Physical Chemistry*, 1989, **93**, 2880–2882.

BIBLIOGRAPHY

- [24] F. Sussich, C. Skopec, J. Brady and A. Cesàro, *Carbohydrate Research*, 2001, **334**, 165 – 176.
- [25] D. Kilburn, S. Townrow, V. Meunier, R. Richardson, A. Alam and J. Ubink, *Nat Mater*, 2006, **5**, 632–635.
- [26] G. Bellavia, S. Giuffrida, G. Cottone, A. Cupane and L. Cordone, *Journal of Physical Chemistry B*, 2011, **115**, 6340–6346.
- [27] C. Olsson, H. Jansson and J. Swenson, *The Journal of Physical Chemistry B*, 2016, **120**, 4723–31.
- [28] S. Magazù, P. Migliardo, A. M. Musolino and M. T. Sciortino, *The Journal of Physical Chemistry B*, 1997, **101**, 2348–2351.
- [29] M. Malferrari, A. Nalepa, G. Venturoli, F. Francia, W. Lubitz, K. Mobius and A. Savitsky, *Physical Chemistry Chemical Physics*, 2014, **16**, 9831–9848.
- [30] A. Lerbret, P. Bordat, F. Affouard, M. Descamps and F. Migliardo, *The Journal of Physical Chemistry B*, 2005, **109**, 11046–11057.
- [31] L. Sapir and D. Harries, *The Journal of Physical Chemistry B*, 2011, **115**, 624–634.
- [32] C. Branca, S. Magazù, G. Maisano and P. Migliardo, *The Journal of Chemical Physics*, 1999, **111**, 281–287.
- [33] C. Branca, S. Magazù, G. Maisano, F. Migliardo, P. Migliardo and G. Romeo, *The Journal of Physical Chemistry B*, 2001, **105**, 10140–10145.
- [34] C. Branca, V. Magazù, G. Maisano, F. Migliardo and A. K. Soper, *Applied Physics A*, 2002, **74**, s450–s451.
- [35] C. Branca, S. Magazù, F. Migliardo and P. Migliardo, *Physica A: Statistical Mechanics and its Applications*, 2002, **304**, 314–318.

- [36] S. Magazù, G. Maisano, F. Migliardo and C. Mondelli, *Biophysical Journal*, 2004, **86**, 3241–3249.
- [37] S. Magazù, F. Migliardo and M. Telling, *Food Chemistry*, 2008, **106**, 1460 – 1466.
- [38] C. Mathews, K. Van Holde and K. Ahern, *Biochemistry. 3rd*, San Francisco, Calif.: Benjamin Cummings. xxviii, 2000.
- [39] Image of 1MBN (Watson, H.C. (1969) The Stereochemistry of the Protein Myoglobin. *Prog.Stereochem.* 4: 299) created with Jmol: an open-source Java viewer for chemical structures in 3D. <http://www.jmol.org/>.
- [40] J. C. Kendrew, G. Bodo, H. M. Dintzis, R. Parrish, H. Wyckoff and D. C. Phillips, *Nature*, 1958, **181**, 662–666.
- [41] F. Franks, *Water, a Comprehensive Treatise: The physics and physical chemistry of water*, Plenum Press, 1972.
- [42] H. Frauenfelder, P. Fenimore, G. Chen and B. McMahon, *Proceedings of the National Academy of Sciences*, 2006, **103**, 15469–15472.
- [43] P. W. Fenimore, H. Frauenfelder, B. H. McMahon and F. G. Parak, *Proceedings of the National Academy of Sciences*, 2002, **99**, 16047–16051.
- [44] V. Lubchenko, P. G. Wolynes and H. Frauenfelder, *The Journal of Physical Chemistry B*, 2005, **109**, 7488–7499.
- [45] J. A. Rupley and G. Careri, *Advances in protein chemistry*, 1991, **41**, 37–172.
- [46] C. A. Knight, A. L. De Vries and L. D. Oolman, *Fish antifreeze protein and the freezing and recrystallization of ice*, Nature Publishing Group, 1984.
- [47] J. Lovelock, *Biochemical Journal*, 1954, **56**, 265.

BIBLIOGRAPHY

- [48] D. Pegg, *Seminars in reproductive medicine*, 2002, pp. 005–014.
- [49] G. B. Strambini and E. Gabellieri, *Biophysical journal*, 1996, **70**, 971.
- [50] S. Leibo, J. McGrath and E. Cravalho, *Cryobiology*, 1978, **15**, 257–271.
- [51] Z. W. Yu and P. J. Quinn, *Bioscience Reports*, 1994, **14**, 259–281.
- [52] J. H. Crowe, J. F. Carpenter and L. M. Crowe, *Annual Review of Physiology*, 1998, **60**, 73–103.
- [53] S. J. Prestrelski, N. Tedeschi, T. Arakawa and J. F. Carpenter, *Biophysical Journal*, 1993, **65**, 661–671.
- [54] S. J. Prestrelski, T. Arakawa and J. F. Carpenter, *Archives of Biochemistry and Biophysics*, 1993, **303**, 465–473.
- [55] J. Wolfe, *Functional Plant Biology*, 1987, **14**, 311–318.
- [56] J. H. Crowe, J. F. Carpenter, L. M. Crowe and T. J. Anchordoguy, *Cryobiology*, 1990, **27**, 219 – 231.
- [57] J. H. Crowe, S. B. Leslie and L. M. Crowe, *Cryobiology*, 1994, **31**, 355–366.
- [58] I. Roy and M. N. Gupta, *Biotechnology and applied biochemistry*, 2004, **39**, 165–177.
- [59] J. F. Carpenter and J. H. Crowe, *Biochemistry*, 1989, **28**, 3916–3922.
- [60] N. Jovanović, A. Bouchard, G. W. Hofland, G.-J. Witkamp, D. J. Crommelin and W. Jiskoot, *European journal of pharmaceutical sciences*, 2006, **27**, 336–345.
- [61] S. N. Timasheff, *Biochemistry*, 2002, **41**, 13473–13482.
- [62] S. N. Timasheff, *Biochemistry*, 1992, **31**, 9857–9864.
- [63] T. Arakawa and S. N. Timasheff, *Biochemistry*, 1982, **21**, 6536–6544.

- [64] T. Arakawa and S. N. Timasheff, *Biophysical Journal*, 1985, **47**, 411–414.
- [65] L. M. Crowe, D. S. Reid and J. H. Crowe, *Biophysical Journal*, 1996, **71**, 2087–2093.
- [66] K. Tanaka, T. Takeda and K. Miyajima, *Chemical & Pharmaceutical Bulletin*, 1991, **39**, 1091–1094.
- [67] S. D. Allison, M. C. Manning, T. W. Randolph, K. Middleton, A. Davis and J. F. Carpenter, *Journal of Pharmaceutical Sciences*, 2000, **89**, 199–214.
- [68] J. H. Crowe, L. M. Crowe, J. F. Carpenter and C. Aurell Wistrom, *Biochemical Journal*, 1987, **242**, 1–10.
- [69] P. S. Belton and A. M. Gil, *Biopolymers*, 1994, **34**, 957–961.
- [70] A. Cerami, *The role of the Maillard reaction in vivo*, The Royal Society of Chemistry: Cambridge, England, 1994.
- [71] M. R. Brown, T. J. Keith and H. R. Knull, *Neurochemistry international*, 1992, **21**, 177–183.
- [72] C. Schebor, L. Burin, M. d. P. Buera and J. Chirife, *LWT - Food Science and Technology*, 1999, **32**, 481–485.
- [73] M. Sola-Penna and J. R. Meyer-Fernandes, *Archives of biochemistry and biophysics*, 1998, **360**, 10–14.
- [74] S. W. Lovesey, *Theory of neutron scattering from condensed matter*, Oxford University Press, 1984.
- [75] D. S. Sivia, *Elementary scattering theory: for X-ray and neutron users*, Oxford University Press, 2011.
- [76] G. E. Bacon, *Neutron scattering in chemistry*, Butterworth-Heinemann Limited, 1977.

- [77] T. Brückel, *Neutron Scattering: Lectures of the JCNS Laboratory Course Held at Forschungszentrum Jülich and the Research Reactor FRM II of TU Munich*, Forschungszentrum Jülich, 2010, vol. 27.
- [78] A. K. Soper, *Molecular Physics*, 2009, **107**, 1667–1684.
- [79] G. Placzek, *Physical review*, 1952, **86**, 377.
- [80] J.-P. Hansen and I. R. McDonald, *Theory of simple liquids*, Elsevier, 1990.
- [81] R. L. McGreevy and L. Pusztai, *Molecular Simulation*, 1988, **1**, 359–367.
- [82] A. K. Soper, *Chemical Physics*, 1996, **202**, 295–306.
- [83] A. K. Soper, *Empirical Potential Structure Refinement - EPSRshell - A Users Guide*, 2015.
- [84] S. Pagnotta, M. Ricci, F. Bruni, S. McLain and S. Magazu, *Chemical Physics*, 2008, **345**, 159–163.
- [85] W. B. O'Dell, D. C. Baker and S. E. McLain, *PLoS ONE*, 2012, **7**, e45311.
- [86] P. B. Conrad and J. J. de Pablo, *The Journal of Physical Chemistry A*, 1999, **103**, 4049–4055.
- [87] L. C. Thomas, *Making Accurate DSC and MDSC® Specific Heat Capacity Measurements with the Q1000 Tzero™ DSC*, http://www.tainstruments.co.jp/application/pdf/Thermal_Library/Applications_Briefs/TA310.PDF, Accessed on: Oct 15, 2016.
- [88] S. E. Pagnotta, S. E. McLain, A. K. Soper, F. Bruni and M. A. Ricci, *The Journal of Physical Chemistry B*, 2010, **114**, 4904–4908.