

THESIS FOR THE DEGREE OF LICENCIATE OF ENGINEERING

## **Sweet Stability**

An Investigation of Disaccharides as Protein Stabilisers  
and Amyloid Fibril Inhibitors

KAJSA AHLGREN



**CHALMERS**  
UNIVERSITY OF TECHNOLOGY

Department of Physics  
*Division of Nano- Biophysics*  
CHALMERS UNIVERSITY OF TECHNOLOGY  
Gothenburg, Sweden 2024

Sweet Stability  
An Investigation of Disaccharides as Protein Stabilisers  
and Amyloid Fibril Inhibitors

KAJSA AHLGREN

© KAJSA AHLGREN, 2024.

Department of Physics  
Chalmers University of Technology  
SE-412 96 Gothenburg  
Sweden Telephone +46 31 772 1000

Chalmers Digitaltryck  
Gothenburg, Sweden 2021

Cover: A three-dimensional image showing preferential hydration of myoglobin in a trehalose solution. The surface of the protein is depicted in blue, surrounded by water molecules shown in white and red. The trehalose molecules are represented by grey stick figures, illustrating their distribution around the protein.

Sweet Stability  
An Investigation of Disaccharides as Protein Stabilisers  
and Amyloid Fibril Inhibitors

KAJSA AHLGREN  
Department of Physics  
Chalmers University of Technology  
Gothenburg, Sweden 2024

## Abstract

The disaccharide trehalose is known for its ability to stabilise proteins by preventing aggregation and elevating their denaturation temperature, consequently avoiding unfolding. The folding and function of proteins are vital for many biological processes. However, the underlying mechanisms of the stabilising effect of trehalose are not yet fully elucidated. A group of diseases known as amyloidosis is associated with the formation of amyloid fibrils, highly structured protein aggregates that can form in various tissues in the body. Enhancing our understanding of protein misfolding and the role of sugars in preventing amyloid fibril formation is therefore of great importance.

In this thesis, the stabilising effect of trehalose is compared with that of sucrose, a structurally similar disaccharide. This comparison aims to solve unanswered questions regarding their interactions with proteins and surrounding water molecules. X-ray and neutron scattering confirmed that myoglobin is preferentially hydrated by water and revealed that trehalose slows down the dynamics of the protein to a greater extent than sucrose, with minimal direct interaction. Differential scanning calorimetry showed that both disaccharides increase the denaturation temperature of the protein lysozyme, sucrose slightly more than trehalose. Additionally, the glass transition temperature of trehalose is marginally higher. Small and wide-angle x-ray scattering demonstrated that both sugars inhibit amyloid fibril formation.

The findings suggest that the effectiveness of both disaccharides in stabilising proteins varies with temperature; trehalose is more effective at lower temperatures around the glass transition, whereas sucrose may be slightly more efficient at higher temperatures around protein denaturation.

**Keywords:** trehalose, sucrose, myoglobin, lysozyme, amyloid fibrils, protein denaturation, glass transition, neutron scattering, DSC.





---

# List of Publications

This thesis is based on the following papers:

## Paper I

**New insights into the protein stabilizing effects of trehalose by comparing with sucrose**

Kajsa Ahlgren, Christoffer Olsson, Inna Ermilova, and Jan Swenson

*Phys. Chem. Chem. Phys.* **2023**, 25, 21215-21226

## Paper II

**Arrangement of trehalose in relation to the hydration layer of myoglobin**

Kajsa Ahlgren, Christoffer Olsson, and Jan Swenson

*In Manuscript*

## Paper III

**The effect of disaccharides on the thermal stability and fibril formation of lysozyme in concentrated conditions**

Kajsa Ahlgren, Fritjof Havemeister, Julia Andersson, Elin K. Esbjörner, and Jan Swenson

*Submitted*

## Paper IV

**A Comparison of Sucrose and Trehalose for Stabilization of Protein using Differential Scanning Calorimetry**

Olivia Jonsson, Agnes Lundell, John Rosell, Sophie You, Kajsa Ahlgren, Jan Swenson

*Submitted*

---

## Contribution to Papers

### **Paper I**

I performed the analysis of the neutron and x-ray scattering data and EPSR modelling together with C.O. All authors contributed to the writing of the paper. Shared main author of the publication.

### **Paper II**

I performed the analysis of the neutron scattering data and EPSR modelling. I was the main author of the manuscript.

### **Paper III**

I prepared all samples and performed all measurements. I performed the AFM imaging together with F.H. I was the main author of the manuscript.

### **Paper IV**

I supervised the project together with J.S. All authors contributed to the writing of the paper.

---

## Publications not included in the thesis

### **Lyotropic liquid crystal elastomers for drug delivery**

Annija Stepulane, Kajsa Ahlgren, Adrian Rodriguez-Palomo, Anand Kumar Rajasekharan, and Martin Andersson

*Colloids and Surfaces B: Biointerfaces* **2023**, 226, 113304



# Acknowledgements

Firstly, I would like express my gratitude to the Swedish research council for their financial support of this project. I would also like to thank Chalmers Materials Analysis Laboratory (CMAL) for providing the SAXS instrument and ISIS Neutron and Muon Source for the allocated beam time.

To my supervisor Jan Swenson, thank you for giving me this opportunity and for guiding me through this journey. I would also like to thank the former and current members of Jan Swenson group, Inna Ermilova, and Achilleas Pipertzis for all your support and knowledge.

Special thanks to Christoffer Olsson, for your patience and for selflessly helping me during the past couple of years. Without you this journey would have been far less smooth.

My acknowledgement extends to my examiner Fredrik Höök, for being such a big scientific inspiration.

I would like to thank Elin Esbjörner, for clarifying the questions I have had concerning amyloid fibrils, thank you for patiently answering all my questions. Special recognition goes to my co-author, Fritjof Havemeiser, for your collaborative spirit and patience during the development of paper III.

To both former and current members of the Biolab. My thanks goes to Julia, for making the long nights in the lab and office more fun. To Simon, Petteri, Nima, Björn, Julie, Erik, Mattias, as well as all former and present students, for the many interesting lunch and fika conversations. I would like to especially thank Julia, Nima, and Achilleas, for proofreading this licentiate thesis. I would also like to thank the women of physics, for creating a safe and relatable community.

Emelie, thank you for always being there, for recognising the difficult times, and for brightening my days with an endless stream of nonsense chitchat. Frida, Johan, Pantea, Amir, Ale, and all my other friends, thank you for all the laughs, late nights and being the most amazing people. Love you guys.

I would like to thank my parents and my brother Måns, for always encouraging and believing in me. In addition, I would like to thank my dear friend Sanna, for always being there. You are the sister I never had.

Lastly, but certainly not least, I would like to thank my boyfriend Govin. For making me laugh every day, for my extended Induchoodan family, for buying me gurr and dip, and for all the love. This would not have been possible without you.

Kajsa Ahlgren



# Abbreviations

AFM	Atomic Force Microscopy
CHARMM	Chemistry at HARvard of Molecular Mechanics
DSC	Differential Scanning Calorimetry
EP	Empirical Potential
EPSR	Empirical Potential Structure Refinement
GAFF	General Amber Force Field
HCl	Hydrochloric acid
HEWL	Hen Egg White Lysozyme
KWW	Kohlrausch-Williams-Watts
LJ	Lennard-Jones
LNCS	Liquid Nitrogen Cooling System
Mb	Myoglobin
MD	Molecular Dynamics
MQ	Milli-Q
NIMROD	Near to Intermediate Range Order Diffractometer
OPLS	Optimised Potential for Liquid Simulations
OPLS-AA	Optimised Parameters for Liquid Simulations All Atom
PDB	Protein Data Bank
QENS	Quasi-Elastic Neutron Scattering
RMC	Reverse Monte Carlo
RP	Reference Potential
RT	Room Temperature
SAXS	Small-Angle X-ray Scattering
$T_c$	Crystallisation temperature
$T_{den}$	Denaturation temperature
$T_g$	Glass transition temperature
$T_m$	Melting temperature
WAXS	Wide-Angle X-ray Scattering





# Contents

<b>Abbreviations</b>	<b>xii</b>
<b>List of Figures</b>	<b>xvii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Theoretical background</b>	<b>3</b>
2.1 Liquids and Amorphous Materials . . . . .	3
2.2 The Transition from Liquid to Glass . . . . .	3
2.3 Water . . . . .	5
2.3.1 The Structure of Water . . . . .	5
2.4 Disaccharides . . . . .	6
2.4.1 The Structure and Chemistry of Trehalose and Sucrose . . . . .	7
2.5 Proteins . . . . .	9
2.6 The Structure and Chemistry of Proteins . . . . .	10
2.7 Myoglobin . . . . .	11
2.8 Lysozyme . . . . .	12
2.9 Protein Folding . . . . .	14
2.10 Protein Misfolding and Aggregation . . . . .	15
2.11 Stabilisation of Proteins . . . . .	18
<b>3 Experimental Methods</b>	<b>21</b>
3.1 X-ray and Neutron Scattering . . . . .	21
3.1.1 Elastic and Inelastic scattering . . . . .	22
3.1.2 The wave-formalism . . . . .	23
3.1.3 Neutron Diffraction . . . . .	25
3.1.4 Quasielastic Neutron Scattering (QENS) . . . . .	28
3.1.4.1 Analysis of QENS . . . . .	31
3.1.5 X-ray Scattering . . . . .	32
3.2 Empirical Potential Structure Refinement (EPSR) Modelling . . . . .	33
3.3 Molecular Dynamics (MD) Simulations . . . . .	35
3.4 Differential Scanning Calorimetry (DSC) . . . . .	37
3.5 Atomic Force Microscopy (AFM) . . . . .	39
<b>4 Experimental Procedures</b>	<b>41</b>
4.1 Sample Preparation - Mb Samples . . . . .	41

4.1.1	Isotope Substitution . . . . .	41
4.2	Sample Preparation - HEWL Samples . . . . .	42
4.3	Scattering Experiments . . . . .	43
4.3.1	Neutron Diffraction . . . . .	43
4.3.2	X-ray Scattering . . . . .	43
4.3.3	Data Corrections . . . . .	43
4.3.4	EPSR Modelling . . . . .	44
4.3.4.1	Coordination numbers . . . . .	45
4.3.5	QENS . . . . .	45
4.3.5.1	QENS Fitting Procedure . . . . .	45
4.4	DSC Experiments . . . . .	46
4.5	AFM . . . . .	47
<b>5</b>	<b>Summary of Results</b>	<b>49</b>
5.1	Paper I . . . . .	49
5.2	Paper II . . . . .	51
5.3	Paper III . . . . .	54
5.4	Paper IV . . . . .	55
<b>6</b>	<b>Conclusion and Future Perspectives</b>	<b>59</b>
	<b>Bibliography</b>	<b>61</b>

# List of Figures

2.1	Illustration of how the volume of a substance varies with temperature during phase transitions from liquid to crystal or glass. . . . .	4
2.2	Chemical structure of sucrose (left) and trehalose (right). The arrows are pointing out the respective glycosidic bond connecting the monosaccharides. In sucrose it is the first carbon of the glucose ring and the second carbon of the fructose ring that takes part in the linkage while in trehalose it is the first carbon in both of the glucose rings. . . . .	8
2.3	Visualisation of (a) an $\alpha$ -helix and (b) a $\beta$ -sheet. Created with BioRender.com. . . . .	11
2.4	The three-dimensional structure of Mb. Obtained from protein data bank (1DWR). . . . .	12
2.5	The three-dimensional structure of HEWL. A-D indicate typical intramolecular distances. Obtained from protein data bank (1LYZ). . . . .	13
2.6	A simplified view of the ragged energy landscape of protein folding referred to as the "folding funnel". Higher up in the "landscape" represent a higher energy level, where the protein is in its unfolded state. Intermediate meta-stable states are located in local minimums and the native state is at the lowest energy state. A vertical step means a shift in conformational entropy. Created with BioRender.com. . . . .	15
2.7	A simplified view of the extended energy landscape of protein folding, which includes misfolding and aggregation and is referred to as the extended folding funnel. Created with BioRender.com. . . . .	16
2.8	Characteristic structure of an amyloid fibril. The distance separating two $\beta$ -strands is 4.7-4.8 Å and the distance between two $\beta$ -sheets is 10 Å. Created with BioRender.com. . . . .	17
3.1	Elastic scattering where the initial (plane) wave, with wave vector $\mathbf{k}$ and wavelength $\lambda_i$ , collide with an atom at position $\mathbf{r}_j$ . The final (spherical) wave, with wave $\mathbf{k}'$ and wavelength $\lambda_f$ , is scattered onto an area, $dA$ . . . . .	24
3.2	The scattering function as a function of energy transfer. The smaller peaks $\hbar\omega \neq 0$ are the inelastic scattering, the sharp peak around $\hbar\omega = 0$ is the elastic scattering, and the broadening of the latter is the quasielastic scattering. . . . .	29
3.3	Schematic illustration of the difference in detector position between SAXS and WAXS. . . . .	33

3.4	Flowchart of procedure of the EPSR simulation method. . . . .	35
3.5	The working principles of DSC displaying the sample and reference pans being sealed and placed on separate thermoelectric disks in the DSC cell. The difference in temperature between the sample and reference is transfered to the computer for analysis. . . . .	38
3.6	Schematic image of a typical DSC curve showing the heat flow as a function of temperature. The crystallisation is seen as an exotherm peak during the cooling process. The glass transition is seen as a step during heating, while both the melting and the protein denaturation are seen as endothermal dips. . . . .	39
5.1	Three-dimensional images of the structural model produced with EPSR, showing the arrangement of water and sugar molecules relative to the protein surface, where (a) is the sucrose containing system and (b) is the trehalose-containing system. The red parts are exposed protein surface, the blue parts are the hydration water and the white stick figures are the trehalose and sucrose molecules. The image was obtained from paper I. . . . .	50
5.2	(a) Intermediate scattering functions as a function of time for both the diluted and concentrated systems at $Q = 1.06 \text{ \AA}^{-1}$ . (b) Inverse average relaxation times as a function of $Q^2$ . The lines marked with circles represent the concentrated system whereas the lines marked with starts represent the diluted system. The inset shows a magnified view of the inverse average relaxation times as a function of $Q^2$ for Mb. The image was obtained from paper II. . . . .	53
5.3	AFM image of the two-component system at pH 2.0 (a) with a magnified image (b). AFM image of the three-component system containing trehalose (c) and sucrose (d) at pH 2.0. All systems were incubated at 57 °C for 24 h. . . . .	55
5.4	Three different characteristic DSC curves displaying the heat flow as a function of temperature. (a) shows the a typical cycle for a system with high water content displaying crystallisation during cooling, (b) shows the cycle for a system where no crystallisation occurs during cooling but during heating (cold crystallisation), and (c) shows a system for which no crystallisation occurs. The image was obtained from paper IV. . . . .	57

# 1

## Introduction

It was 170 years ago that the term amyloid was coined.<sup>1</sup> A term which much later became highly associated with one of today's most common diseases, Alzheimer's disease. In 2021, 100 000 people was expected to suffer from Alzheimer's disease in Sweden.<sup>2</sup> The first documented case was discovered by Alois Alzheimer, from who the disease later got its name, in the early 1900s.<sup>1</sup> Since then, heavy research has been conducted on the subject which has taught us that the disease is related to the formation of long thread-like protein structures, called amyloid fibrils, in the brain. Amyloid fibrils are not only associated with neurodegenerative diseases, such as Alzheimer's, but are implicated in a group of diseases referred to as amyloidoses.<sup>3</sup>

Through protein stabilisation it is possible to inhibit the formation of amyloid fibrils. This has been done in many studies by the use of different methods and additives.<sup>4</sup> One of which is the disaccharide trehalose, whose molecular structure is very similar to that of ordinary table sugar sucrose. The mechanism behind why trehalose possess this stabilising property is still unknown. The aim of this thesis is to broaden our understanding of the stabilising and amyloid inhibitory effect of trehalose.

This licentiate thesis is based on the work of four papers. In paper I, protein-sugar-water systems were analysed with the purpose of understanding how trehalose affects the structure and dynamics of the system. The study was conducted on both trehalose and sucrose in order to use their differences and similarities as a guide to better understand the stabilising property of trehalose. To obtain a better understanding of the structural arrangement of trehalose in relation to the surface of the protein, a more diluted system was prepared in paper II. The inhibitory effect

trehalose has upon the formation of amyloid fibrils was investigated in paper III, where the properties of trehalose were compared to those of sucrose. In paper IV an extensive scan was performed to compare the different ways in which trehalose and sucrose stabilise proteins depending on temperature. Here the possibility of synergistic effects was also addressed.

In papers I and II, neutron scattering together with empirical potential structural refinement modelling was used to study both structural and dynamical properties of the systems. In paper I, these experiments were combined with classical molecular dynamics simulations. In paper III techniques such as small- and wide-angle x-ray scattering, atomic force microscopy were used. Differential scanning calorimetry was used in both paper III and IV.

This licentiate thesis includes a background both of the studied materials as well as the analytical methods used. In chapter 2 the theoretical background concerning the materials is presented, while chapter 3 provides necessary theory concerning the experimental techniques as well as simulation methods used. In chapter 4 the experimental procedures are presented. Chapter 5 presents a summary of the results obtained and is sorted based on papers. Finally, conclusion and future perspectives can be found in chapter 6.

# 2

## Theoretical background

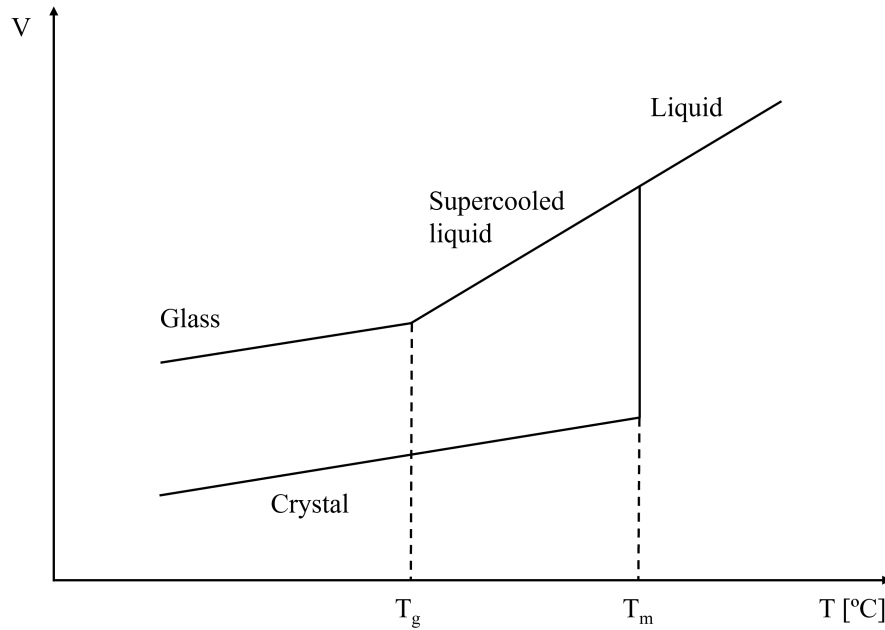
### 2.1 Liquids and Amorphous Materials

Liquids and amorphous materials are examples of physical forms that lack any form of long range order.<sup>5</sup> The only physical form in which the molecules are neatly arranged in a repeating pattern is the crystalline state. If one only considers the ordering of molecules, one could say that liquids and amorphous solids are the same. However, one important difference is the viscosity of the material, where the viscosity of a liquid is low and that of an amorphous solid is high. Despite the lack of long range order, certain amorphous materials contain short-range periodicity.<sup>5</sup> This thesis includes studies of both liquid and amorphous materials as well as the transition from liquid to glass which will be described in the following section.

### 2.2 The Transition from Liquid to Glass

When a liquid is cooled, its viscosity typically increases, a phenomenon largely due to a decrease in the kinetic motion of the individual particles and an increase in density. When a material is cooled below its point of freezing, it becomes more energetically favourable for the particles to form structured lattices; this phenomenon is known as crystals. The crystallisation event is characterised in Figure 2.1 as an abrupt drop in volume at the point of the melting temperature,  $T_m$ . The crystallisation is a first-order phase transition. When a liquid is cooled below its  $T_m$ , there is another possible outcome; the liquid can become supercooled, leading to an increase in viscosity as the temperature decreases, eventually resulting in the formation of a

glass.<sup>6</sup> A glass is an amorphous solid which exhibits glass transition. As oppose to crystallisation, the glass transition is a second-order phase transition. This means that no heat is transferred between the sample and its surroundings. However, there is a change in heat capacity. By monitoring the volume of a liquid as a function of temperature, as in Figure 2.1, the phase transitions can easily be displayed.



**Figure 2.1:** Illustration of how the volume of a substance varies with temperature during phase transitions from liquid to crystal or glass.

For the supercooled liquid, the volume is decreased linearly until a point where the slope changes and a glass is formed. This is commonly referred to as the glass transition and the temperature at which it occurs is called the glass transition temperature,  $T_g$ . The different phases are highly dependent on the cooling rate, where a slower rate enables the liquid to stay supercooled longer, resulting in a lower  $T_g$ . Thus, the  $T_g$  partly depends on the thermal history of the material.<sup>6</sup> A convenient way to measure the glass transition event is via differential scanning calorimetry (DSC), which will be presented in more detail in section 3.4. The thermal events are then detected as changes in enthalpy as the heat is varied.



## 2.3 Water

Water is a fundamental component of our planet, essential to all known forms of life. It is therefore understandable that a great amount of studies have been made to understand the properties of water at a molecular level.<sup>7</sup> In the following section the intra- and intermolecular structure of water will be presented.

### 2.3.1 The Structure of Water

The water molecule consists of one oxygen and two hydrogen atoms, forming a bent structure with an angle of  $105^\circ$ .<sup>8</sup> This specific arrangement is crucial to understanding the properties of water. Oxygen and hydrogen differ significantly in their electronegativity; oxygen has an electronegativity value of 3.5,<sup>8</sup> while that of hydrogen is notably lower at 2.1.<sup>8</sup> Electronegativity refers to an element's ability to attract electrons. Due to the higher electronegativity of oxygen, it tends to draw electrons closer from the less electronegative hydrogen atoms, leading to an uneven distribution of electrical charge within the molecule. As a result, the region around the oxygen atom acquires a slight negative charge, while the regions near the hydrogen atoms become slightly positive. This uneven charge distribution, in combination with the bent geometry of the molecule, makes water a permanent dipole.

The difference in charge makes the water molecule a great subject to form hydrogen bonds. Hydrogen bonds are formed when covalently linked hydrogen atoms are attracted by an electronegative atom of a different molecule, in this case oxygen. In water, each molecule acts both as an acceptor and a donor, enabling intermolecular hydrogen bonding.

At temperatures above the boiling point of water ( $100^\circ\text{C}$ ), water molecules exist as gas (water vapour). In this gaseous state, the molecules move rapidly, preventing the formation of intermolecular hydrogen bonds. As the temperature decreases, the water molecules slow down and transition into a liquid phase, allowing oxygen in one water molecule to form a hydrogen bond with a hydrogen atom in another molecule. However, in this liquid state, due to thermal fluctuation, the water molecules continuously translate and rotate, leading to the breaking of hydrogen bonds approx-

imately once every picosecond at room temperature (RT).<sup>9</sup> Further temperature reduction leads to the crystallisation of water, forming ice. In this solid phase, the molecules move even more slowly. Contrary to what might be expected, the molecules in solid ice are less densely packed compared to their arrangement in the liquid phase, showcasing one of water's many unique properties. On average, each water molecule forms a tetrahedral network by bonding with approximately four other water molecules.<sup>10,11</sup> Two oxygen atoms are generally separated by a distance of 2.8 Å.<sup>12</sup> However, this value is shifted with temperature<sup>13,14</sup> and the density maximum of water can be found at 4 °C.

## 2.4 Disaccharides

Some of the most abundant biomolecules on earth are the carbohydrates. The name "carbohydrates" originates from the belief that such substances were hydrates of carbon.<sup>15</sup> The common molecular formula can be expressed as  $C_n(H_2O)_m$ . Based on their level of polymerisation, carbohydrates can be categorised into four principal groups: monosaccharides, disaccharides, oligosaccharides, and polysaccharides. Monosaccharides are the most simple forms of sugar and are seen as the building block of carbohydrates. One very important monosaccharide is glucose which is the primary source of energy in most organisms.<sup>16</sup> When two monosaccharides come together and form a glycosidic bond through a condensation reaction, a disaccharide is formed. The importance of disaccharides, seen as they are a major source of energy, makes the formation as well as the hydrolysis of the glycosidic bond one of the most crucial reactions in the chemistry of carbohydrates.<sup>17</sup>

The most common disaccharide, commonly known as ordinary table sugar, is sucrose. Sucrose is produced in nature by plants such as sugar beets and sugar cane. The disaccharide is an important part of the human diet and is absorbed into the body after being hydrolysed by the enzyme sucrase. A structurally similar disaccharide is trehalose. Trehalose is synthesised by some bacteria and fungi to be used as a source of energy. In addition, trehalose exists in various desiccation-tolerant organisms, such as the tardigrade, and is used as a means of protection from drying. High

levels of trehalose is accumulated in many of these organisms. However, tardigrades accumulate very small amounts of trehalose yet it is still seen as a key component for the survival of the extremophile.<sup>18</sup> Trehalose is not the only disaccharide that increases the stability of proteins and other biomolecules, disaccharides in general have been seen to possess this ability.<sup>19,20</sup> For example, sucrose has been found to exist in desiccation-tolerant pollen and seeds.<sup>21</sup>

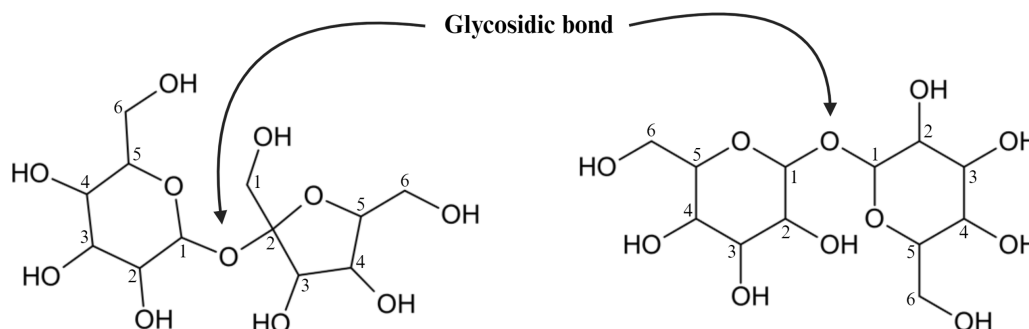
The main purpose of this thesis is to understand the stabilising and anti-aggregating effect trehalose has upon proteins, which partly includes a comparative analysis with sucrose. To properly understand their differences, their chemistry and structure will be scrutinised in the following subsection.

### 2.4.1 The Structure and Chemistry of Trehalose and Sucrose

As mentioned in section 2.4, sucrose and trehalose are two structurally similar molecules. They have exactly the same chemical formula ( $C_{12}H_{22}O_{11}$ ) and a molecular weight of 342.3 g/mol. However, there is a major difference in their respective intramolecular structures, which gives rise to potentially crucial differences. The fact that one of the glucose rings in trehalose is replaced by a fructose ring in sucrose makes the latter less likely to form intermolecular hydrogen bonds, whereas the former is less likely to form intramolecular hydrogen bonds.<sup>22</sup> As mentioned in section 2.4, the monosaccharides are connected via a glycosidic bond, which is pointed out in Figure 2.2. The linkage in sucrose is a  $\alpha$ -1,  $\beta$ -2-glycosidic linkage,<sup>23</sup> meaning that it is the first carbon in the glucose ring and the second carbon in the fructose ring that participate in the glycosidic linkage. In contrast, in trehalose it is a  $\alpha$ -(1-1)-glycosidic linkage,<sup>23</sup> which implies that it is the first carbon of the two glucose rings that participate in the bond. It has been confirmed that the strength of this bond is not the same for trehalose (1 kcal/mol) and sucrose (27 kcal/mol).<sup>24</sup> The bond energy for sucrose is quite high compared to trehalose, which makes the former a less stable structure compared to the latter.<sup>24,25</sup> In addition, this also makes sucrose more susceptible to acidic hydrolysis than trehalose, which is only

## 2. Theoretical background

hydrolysed under severe hydrolysis conditions.<sup>24</sup> The reaction is exotherm and is catalysed by elevated temperatures.<sup>22, 26</sup>



**Figure 2.2:** Chemical structure of sucrose (left) and trehalose (right). The arrows are pointing out the respective glycosidic bond connecting the monosaccharides. In sucrose it is the first carbon of the glucose ring and the second carbon of the fructose ring that takes part in the linkage while in trehalose it is the first carbon in both of the glucose rings.

Disaccharides can be divided into reducing and non-reducing sugars depending on the chemical structure of the molecule. Reducing sugars, unlike non-reducing ones, have free aldehyde or ketone groups, which makes them more chemically reactive. Both sucrose and trehalose are non-reducing disaccharides<sup>23</sup> since the glycosidic bond is between their hemiacetal carbon atoms. This reduces the chemical reactivity of both sucrose and trehalose, which might be of great importance when it comes to their use as stabilising agents. However, it should be mentioned that upon hydrolysis, both disaccharides are reduced into two reducing monosaccharides, a process more likely for sucrose.<sup>24</sup>

As mentioned in section 2.4, both trehalose and sucrose have been found to act as stabilising agents in various desiccation-tolerant organisms and it is widely believed that the stabilising effect of trehalose is superior compared to other disaccharides. The  $T_g$  is higher for trehalose compared to other disaccharides including sucrose,<sup>22, 27</sup> which is also something we are showing in paper IV. Therefore, trehalose is widely used in pharmaceutical products to make it possible to store and ship them without cooling.<sup>28–30</sup>

These structural and chemical differences may seem small. However, they might be crucial when it comes to stabilising proteins, and other biomolecules, as it affects both how the disaccharides interact with the biomolecule itself and the surrounding water molecules. As mentioned in section 2.3.1, water forms a tetrahedral network via hydrogen bonding to other water molecules. The addition of disaccharides to water has been seen, in a vast number of studies, to disrupt the tetrahedral network.<sup>31–35</sup> These studies have also shown that, via extensive hydrogen bonding, trehalose possesses the ability to disrupt the water structure more efficiently compared to other disaccharides. This has been seen as a contribution to the superior stabilising effect of trehalose as it breaks up the structure of water which would give rise to ice formation upon cooling. Furthermore, this could be an explanation to the higher  $T_g$  of trehalose compared to other disaccharides such as sucrose.

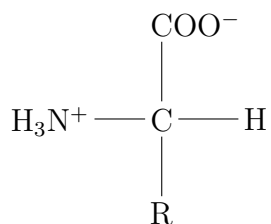
Moreover, it has previously been shown that the dynamics of the disaccharides and water is strongly coupled and that the coupling strength in trehalose-water solutions is higher compared to other disaccharides, such as sucrose.<sup>35</sup> Furthermore, trehalose slows down the dynamics of the water to a greater extent compared to other disaccharides,<sup>35</sup> which also was observed in paper I.

## 2.5 Proteins

In 1838, the Swedish chemist Jöns Jakob Berzelius suggested to the Dutch chemist Gerardus Mulder that an unknown substance, vital for both animals and plants, should be named protein.<sup>36</sup> It originates from the Greek word *proteios* which means *of primary importance*.<sup>36,37</sup> This term captures the importance of proteins and the role they play. Their discovery has revolutionised our understanding of biological processes and led to significant advancements in health and medicine. The subsequent sections will emphasise the significance of functional proteins and explore the implications of their malfunction.

## 2.6 The Structure and Chemistry of Proteins

Amino acids linked together, by the means of covalent peptide bonds, form a long peptide chain which constitutes the macromolecules commonly referred to as proteins. The peptide bond is formed when two amino acids come together via the carboxylic and amino group, respectively, and release a water molecule. *In vivo* proteins are synthesised in the ribosome, which is an organelle in the cell consisting of both RNA and protein. It is the responsibility of the ribosome to translate the genetic code transcribed in mRNA into the amino acid sequence of each protein.<sup>38</sup> The amino acid sequence is unique for every protein and it defines the function of the protein. Almost all amino acids that have been found in proteins have the structure



where the central carbon is referred to as the  $\alpha$ -carbon and has a hydrogen atom, a carboxyl group, and an amino group attached to it.<sup>39,40</sup> The R-group is referred to as an amino acid side chain specific for each amino acid.

The structure of protein molecules is organised in a structural hierarchy. It was the Danish biochemist Kai Linderstrøm-Lang who coined the terms *primary*, *secondary*, and *tertiary* structure,<sup>41</sup> which refer to the different structure levels that describe the structural hierarchy in proteins. The primary protein structure contains information about the number and sequence of amino acids. The secondary protein structure describes the local three-dimensional structure of proteins, where the two most common structures are  $\alpha$ -helix and  $\beta$ -sheet. Figure 2.3(a) and (b) shows a visualisation of an  $\alpha$ -helix and a  $\beta$ -sheet. In the former, a single polypeptide chain twists around itself to form an  $\alpha$ -helical structure with 3.6 residues per turn.<sup>41</sup> Both  $\alpha$ -helices and  $\beta$ -sheets are stabilised by hydrogen bonds between the amine groups and the carbonyl oxygen.<sup>41</sup> However, for  $\alpha$ -helices the hydrogen bonds are formed within the polypeptide chain, whereas in  $\beta$ -sheets they are formed between adjacent

polypeptide chains.



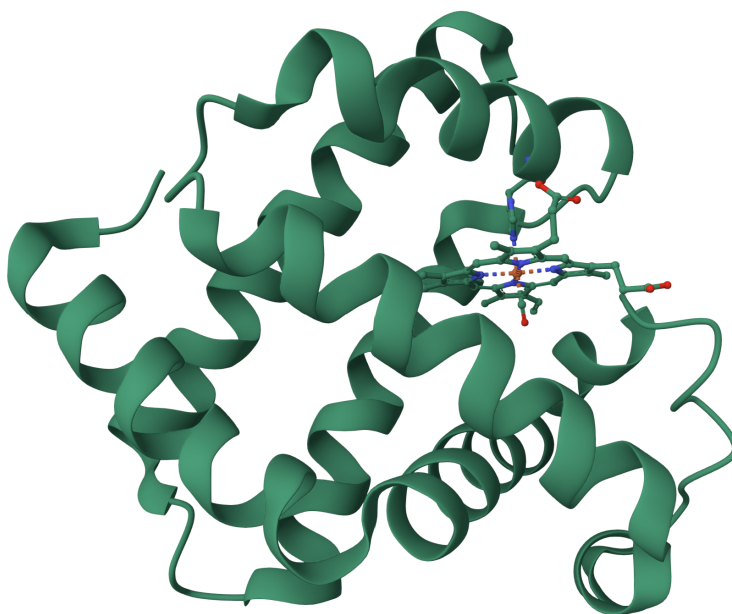
**Figure 2.3:** Visualisation of (a) an  $\alpha$ -helix and (b) a  $\beta$ -sheet. Created with BioRender.com.

The three-dimensional fold is referred to as the tertiary protein structure and occurs primarily due to side chain interactions such as hydrogen bonding, ionic bonding, dipole-dipole interactions, and London dispersion forces. Hydrophobic amino acids are generally buried inside the interior of the protein while hydrophilic side chains are located at the surface of the protein. This is referred to as the hydrophobic effect and is also of great importance for the tertiary structure. If a protein consists of more than one polypeptide chain it can form a fourth hierarchical structure referred to as the *quaternary* protein structure. The two proteins that have been studied in this thesis are myoglobin (Mb) and hen egg white lysozyme (HEWL), both of which consist of a single polypeptide chain. The following sections will provide a brief introduction to the two proteins.

## 2.7 Myoglobin

The first protein to have its three-dimensional structure revealed was Mb. This discovery was made in 1958 by the British biochemist and crystallographer, John

Kendrew, by the use of x-ray crystallography.<sup>42</sup> It is a significant milestone in the field of molecular biology and biochemistry as it provides deep insights into protein structure and function. The main function of the small globular protein is to supply oxygen to the cells in our muscles. The Mb found in humans consists of 153 amino acids and has a molecular weight of 17.7 kDa. The Mb used in papers I and II comes from equine heart which has the same amount of amino acids and a similar molecular weight as Mb found in humans. Figure 2.4 shows the three-dimensional structure of Mb, from which it can be seen that the secondary structure of Mb consists of eight  $\alpha$ -helices and a heme-group.



**Figure 2.4:** The three-dimensional structure of Mb. Obtained from protein data bank (1DWR).<sup>43</sup>

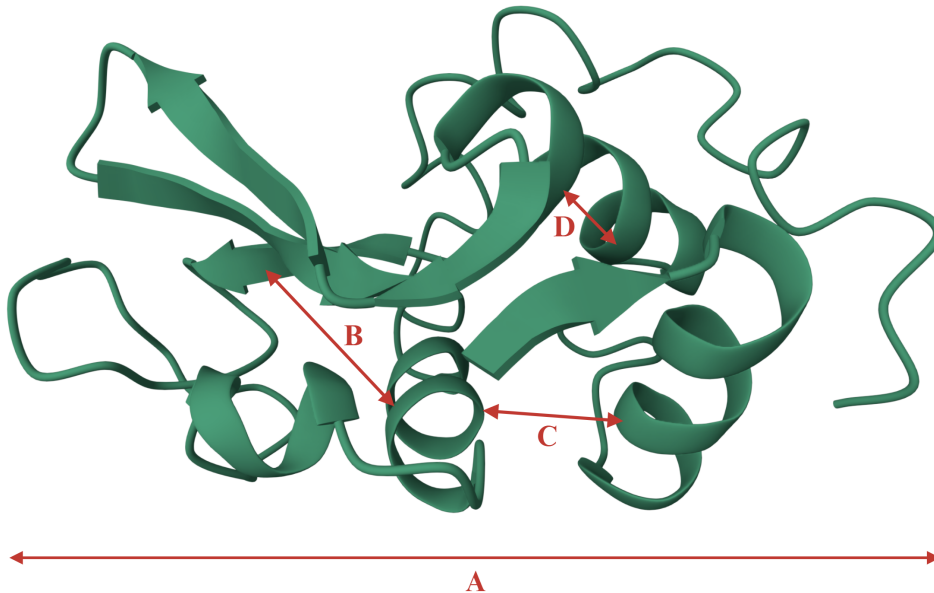
## 2.8 Lysozyme

In 1922, Alexander Fleming discovered the protein lysozyme.<sup>44</sup> It is a small (14.3 kDa)<sup>45</sup> globular enzyme, consisting of 130 amino acids,<sup>45</sup> and is commonly found in human tears, saliva<sup>46</sup> and nasal mucous.<sup>47</sup> The protein is essential as its main function is to hydrolyze the hydrocarbon chains in bacterial cell walls. However, human lysozyme is also related to the very rare disease hereditary lysozyme amyloidosis.<sup>48</sup>



Hereditary lysozyme amyloidosis is a condition in which the protein lysozyme misfolds, aggregates and forms insoluble amyloid fibrils (Described in section 2.10). Such fibrils accumulate in various tissues in the body, disrupting their normal structure and function.<sup>49,50</sup>

Lysozyme from hen egg white (HEWL), commonly used as a model protein for amyloid fibril analysis,<sup>51,52</sup> is an ortholog of human lysozyme. Despite similar function, their amino acid sequences differ slightly; the amino acid sequence of HEWL consists of 129 amino acids and has a molecular weight of 13.9 kDa.<sup>53</sup> The three-dimensional structure of HEWL is shown in Figure 2.5, where it can be seen that its secondary structure consists of a mixture of five to seven  $\alpha$ -helices and a three-stranded anti-parallel  $\beta$ -sheet. The protein has four stabilising disulfide bridges. *A-D* refer to different hierarchical structure levels of the protein. Firstly, *A* refers to the tertiary structure of the protein and *B* to the inter-domain correlation (the distance between the  $\beta$ -sheet and the  $\alpha$ -helix). *C* refers to the intra-domain correlation, for example the distance separating two  $\alpha$ -helices. Lastly, *D* refers to secondary structure, including the correlation between side-chains.<sup>54</sup>



**Figure 2.5:** The three-dimensional structure of HEWL. A-D indicate typical intramolecular distances. Obtained from protein data bank (1LYZ).<sup>55</sup>

## 2.9 Protein Folding

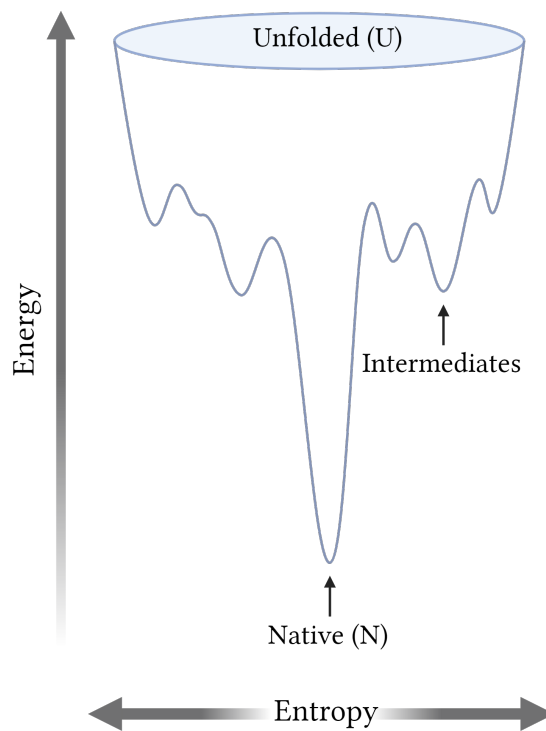
Protein folding can, simplified, be explained as a process driven by the hydrophobic effect, simultaneously packing interactions and hydrogen bonds are optimised.<sup>56</sup> The stability of a protein is given by the free energy difference between its folded, or native (N), and unfolded (U) state

$$\Delta G_{\text{folding}} = -RT \ln K_{\text{eq}} = \Delta H - T \cdot \Delta S, \quad (2.1)$$

where  $R$  is the universal gas constant,  $T$  the absolute temperature,  $K_{\text{eq}}$  the equilibrium constant,  $\Delta H$  the change in enthalpy, and  $\Delta S$  is the conformational entropy change. To be able to function properly proteins need to be folded into their respective native state. A correctly folded protein is one where the amino acid chain forms the secondary structure, which then folds into the tertiary structure, resulting in the protein achieving its lowest energy state.

The folding process of proteins is a complicated subject that has been studied for many years. Due to the complexity of proteins and given their ability to assume their N-state from denatured (U) in matter of seconds, advocates for that proteins do not fold at random. This theory, later called "Levinthal's paradox", was coined by the American scientist Cyrus Levinthal in the 1960s.<sup>57</sup> According to Levinthal, protein folding should occur under kinetic control rather than thermodynamic control, which suggests that a protein must fold into the local energy minimum via the fastest folding pathway.<sup>56,58</sup> This observation has led scientists to explore alternative hypotheses that could explain the folding of proteins.

Different hypotheses for explaining the mechanism of protein folding have been proposed, one of which is called the *free energy landscape theory*. A visualisation of the protein folding funnel is shown in Figure 2.6. It visualises the free energy of each conformation in the folding process as a function of the degrees of freedom. The theory states that the protein can fold into intermediate states, which are metastable and from which the protein then quickly can fold into its native state.<sup>58,59</sup>

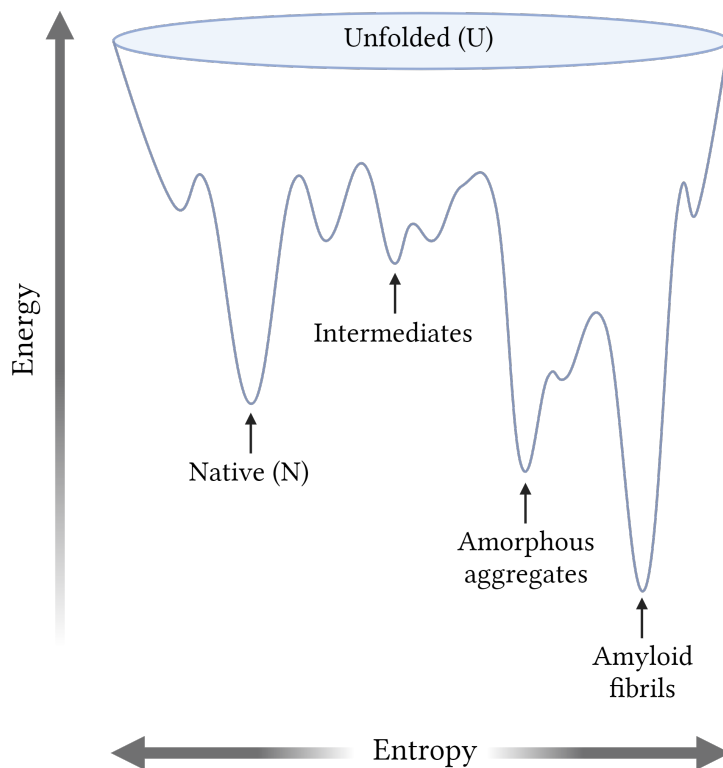


**Figure 2.6:** A simplified view of the ragged energy landscape of protein folding referred to as the "folding funnel". Higher up in the "landscape" represent a higher energy level, where the protein is in its unfolded state. Intermediate meta-stable states are located in local minimums and the native state is at the lowest energy state. A vertical step means a shift in conformational entropy. Created with BioRender.com.

## 2.10 Protein Misfolding and Aggregation

When a protein is unable to fold correctly or maintain its proper structure, it can misfold or assume an alternative form. This is called misfolding and can lead to a destabilised protein and aggregation. Faulty protein folding and aggregation can in some cases lead to the formation of amyloid fibrils. Amyloid fibrils are highly structured protein aggregates, which are associated with two of the most common neurodegenerative diseases: Alzheimer's and Parkinson's,<sup>60-62</sup> but also with non-neurodegenerative diseases such as hereditary lysozyme amyloidosis, as mentioned in section 2.8. In Figure 2.7 it can be seen that the stability of the amorphous aggregates and the amyloid fibrils is higher than for the native state of the protein.<sup>63,64</sup>

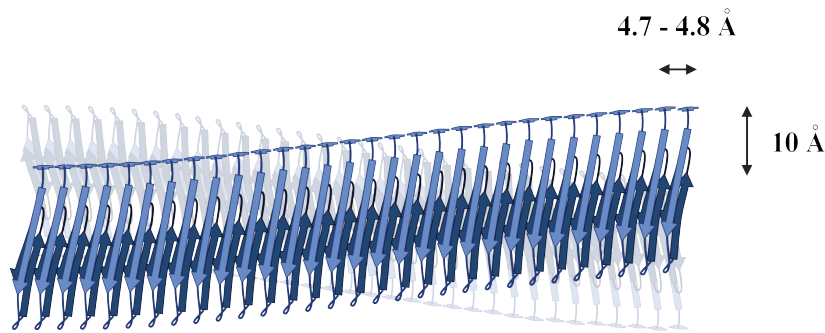
Amorphous protein aggregates are protein aggregates that lack specific high-order structure.<sup>65</sup>



**Figure 2.7:** A simplified view of the extended energy landscape of protein folding, which includes misfolding and aggregation and is referred to as the extended folding funnel. Created with BioRender.com.

It was the German physician Rudolph Virchow who, in 1854, first introduced the term "amyloid" to describe a macroscopic tissue abnormality that exhibited a positive iodine staining reaction.<sup>1</sup> In 1859, Friedrich and Kekulé uncovered the high nitrogen content in fibrils,<sup>66</sup> leading to the classification of amyloid fibrils as proteins.

Amyloid fibrils are highly structured, linear, protein aggregates, which consist of  $\beta$ -sheet structure oriented perpendicular to the long axis of the fibril. The structure of an amyloid fibril can be seen in Figure 2.8. Amyloid fibrils are characterised by the cross- $\beta$  fold, which contains two main distances: 4.7–4.8 Å and 10 Å. The former is the intrastrand distance separating two adjacent monomers, and the latter is the intersheet spacing between two  $\beta$ -sheet layers.<sup>67</sup> This gives rise to characteristic anisotropic patterns within amyloid fibrils, which were discovered using x-ray diffraction.<sup>68,69</sup>



**Figure 2.8:** Characteristic structure of an amyloid fibril. The distance separating two  $\beta$ -strands is 4.7-4.8 Å and the distance between two  $\beta$ -sheets is 10 Å. Created with BioRender.com.

To be able to form amyloid fibrils, the protein needs to undergo partial unfolding or misfolding. This can be triggered by various factors such as mutations, changes in pH, temperature, ionic strength, or interaction with other molecules and is unique for each protein. The misfolded state exposes hydrophobic regions that are usually buried in the native state, making the protein more prone to aggregate. The formation of amyloid fibrils from lysozyme is suggested to start with nucleation, a critical step, which involves the formation of a small cluster of misfolded proteins. This step is a rate-limiting step referred to as primary nucleation and involves protein monomers only. Monomer-dependent secondary nucleation is when a new nuclei is formed on the surface of an already formed fibril, which accelerates the aggregation process.<sup>70</sup> Once a nucleus is formed, the structure is elongated by the addition of monomers to the ends of the fibrils. The elongation is a fast process compared to the nucleation. Mature lysozyme fibrils can undergo further association, leading to larger aggregates. This can result in the formation of insoluble amyloid deposits, which can be found in various tissues and organs in the case of amyloidoses.

### 2.11 Stabilisation of Proteins

Most proteins need to be able to fold properly into its native state and remain there<sup>a</sup>, under physiological conditions, in order to function. The focus of this thesis has been to investigate the stabilising effect the two disaccharides trehalose and sucrose have upon proteins. The stability of a protein refers to its ability to keep the structure intact in the face of external stresses such as heat or denaturants.<sup>71</sup> The biological activity of a protein can be lost if the protein undergoes unfolding, which is also known as denaturation<sup>b</sup>.

Additionally, the protein needs to be able to move in order to function properly. However, if the dynamics are too fast there is a great risk of destabilisation of the protein. The dynamics increase with increasing temperature, which in turn destabilise the protein. One way to avoid this is by simply decrease the temperature, which is referred to as cryopreservation. This is a common method to enable storing and transportation of biological materials. In addition, the protein dynamics can be slowed down by reducing the concentration of water in the system, which can be done through, for example, freeze-drying. However, these techniques do not come without difficulties, one of which is the formation of ice crystals.

To avoid the formation of ice crystals the protein solution can be frozen in such a way that it first enters the supercooled phase and subsequently the glassy state. This is commonly done by the use of cryoprotectants, such as trehalose and sucrose. Although the stabilising effect of both sugars has been widely studied, the mechanism behind this property is not fully understood. However, there have been some debated hypotheses, such as preferential hydration and water replacement, which are briefly described in the following paragraph.

The preferential hydration model includes little to no interaction between the sugar and the surface of the protein. Rather, the water molecules interact directly with the hydrophilic parts of the protein surface and the sugar molecules are preferentially excluded from the surface of the protein. However, this means that the sugar

---

<sup>a</sup>This does not apply to the intrinsically disordered proteins which fail to form a stable structure. However, they still exhibit biological activities.

<sup>b</sup>Denaturation may be reversible if the destabilising environment is restored.

molecules stabilise the native state of the protein without a direct sugar-protein interaction. The model was first proposed by Arakawa and Timasheff<sup>72</sup> and has since then been extensively confirmed both experimentally and through simulations<sup>73-76</sup> and is confirmed also in papers I, II and IV.

An alternative model, known as the water replacement model, proposes how disaccharides stabilise proteins. This model suggests that disaccharide molecules replace water molecules at the surface of the protein, thereby interacting directly with it. However, this model lacks empirical validation, as it has not been corroborated by neither experiments nor computer simulations.





# 3

## Experimental Methods

In this thesis a combination of different experimental methods have been used to investigate the stabilising property of disaccharides. Among them is neutron diffraction, which was used in combination with empirical potential structure refinement (EPSR) modelling to analyse the structural properties of trehalose and sucrose in paper I and II. Quasi-elastic neutron scattering (QENS) combined with molecular dynamics (MD) simulations were also used in paper I to investigate the dynamics of the disaccharides and water molecules in the system. In addition, QENS was used to investigate the dynamics of trehalose and water in paper II. In paper III, inter- and intra-molecular distances were investigated by small-angle x-ray scattering (SAXS) and wide-angle x-ray scattering (WAXS). Furthermore, differential scanning calorimetry (DSC) was used in both paper III and IV to study the heat induced denaturation of HEWL. In paper III, DSC was also used to study the crystallisation of the water in the systems, and in paper IV DSC was used to study the glass transition temperature of trehalose and sucrose. Finally, the presence of fibrillar structures was confirmed through atomic force microscopy (AFM).

### 3.1 X-ray and Neutron Scattering

X-ray and neutron scattering are fundamentally similar non-destructive techniques,<sup>77</sup> which can be used to investigate both amorphous and crystalline materials on a nanometre scale. Even though the scattering of x-rays and neutrons are based on the same principle, there are some differences separating the two techniques. As the names suggest, the particles of the incoming wave in neutron scattering are neutrons

whereas in x-ray scattering they are photons. The photons interact with the electrons surrounding the atomic nuclei. Therefore, the scattering cross section scales with the atomic number of the periodic table. However, in neutron scattering no such correlation can be made as the neutrons instead interact with the nucleus of the atom. As a consequence, heavy elements are easier to detect using x-ray scattering whereas in neutron scattering there is no systematic difference. The difference in contrast enables the use of x-ray and neutron scattering as complementary techniques.

In neutron scattering there are two different sources to harvest neutrons to perform neutron scattering experiments, reactor-source and spallation-source.<sup>78</sup> A reactor-source is usually continuous and the neutrons are a by-product of radioactive decay from uranium ( $^{235}\text{U}$ ), whereas a spallation-source is pulsed and the neutrons are released from a heavy material, such as tungsten, after the collision with high energy protons.<sup>79</sup>

#### 3.1.1 Elastic and Inelastic scattering

The principle of particle scattering, whether the particle is a photon or a neutron, is the change of momentum,  $P$ , and energy,  $E$ , of the incident particle. The momentum transfer can be expressed as

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

where  $\hbar$  is the reduced Planck constant and  $\mathbf{Q} = \mathbf{k} - \mathbf{k}'$  is the scattering vector defined as the difference between the wave vector of the incident and final wave. The energy transfer is expressed as

$$\Delta E = \hbar\Delta\omega, \quad (3.2)$$

where  $\Delta\omega$  is the difference between the frequency of the incident,  $\omega_i$ , and final,  $\omega_f$ , particle. When there is no exchange of energy,  $\Delta E = 0 \Leftrightarrow \Delta\omega = 0$ , the scattering is referred to as elastic. In such a case,  $|\mathbf{k}| = |\mathbf{k}'| = 2\pi/\lambda$ , where  $\lambda$  is the particle wavelength. In contrast, inelastic scattering is when the energy is not conserved,

$\Delta E \neq 0$ . One difference between neutrons and x-rays is that the energy of thermal neutrons is on the same order of magnitude as the vibrations of atoms. This implies that neutrons can be used to measure the transfer of energy. However, the energy of x-rays is much higher compared to neutrons and it is therefore difficult to measure inelastic scattering using x-rays. The elastic neutron scattering is used to gain information regarding the atomic distribution in a system, whereas the inelastic neutron scattering provides information about the dynamics of the atoms. In the following sections, elastic scattering will be assumed if nothing else is stated.

The geometry involved in the scattering process is the same for both x-rays and neutrons. The incoming wave collides and scatters off the sample through an angle,  $2\theta$ , which leads to Bragg's law

$$|\mathbf{Q}| = Q = \frac{4\pi}{\lambda} \sin\left(\frac{\theta}{2}\right). \quad (3.3)$$

The stream of incident particles is referred to as the particle flux,  $\Psi$ , which is defined as the number of those particles impinging on a surface per second per surface area perpendicular to the direction of the beam.<sup>79</sup> In neutron scattering this value is typically on the magnitude of  $10^4$ – $10^9$  neutrons/cm<sup>2</sup>/s. In contrast to x-rays, which typically have 6–9 order of magnitude more, neutrons will never be used to study single nano-sized objects.

### 3.1.2 The wave-formalism

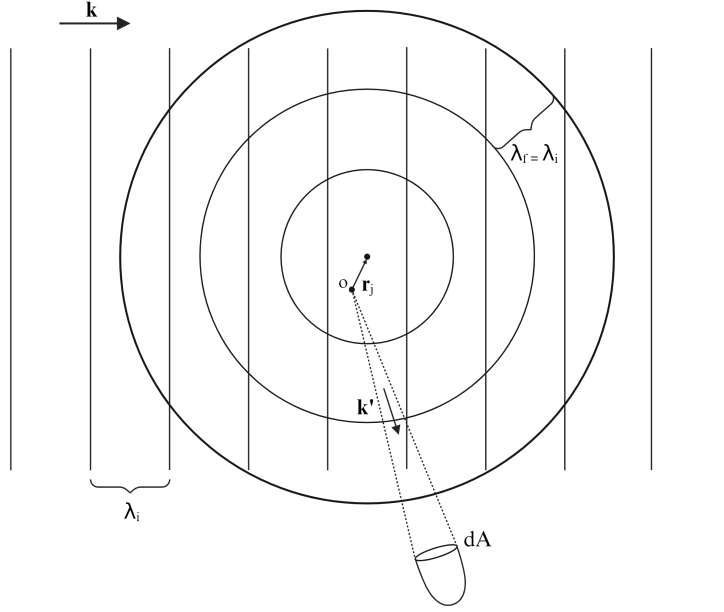
For both x-ray and neutron scattering, the incoming particles can, before impinging the sample, be described as a plane wave given by<sup>79</sup>

$$\psi_i(\mathbf{r}) = \psi_0 e^{i\mathbf{k}\mathbf{r}}, \quad (3.4)$$

where  $\psi_0$  is the amplitude of the incident wave related to the incident flux as  $\Phi = |\psi_0|^2$ , and  $\mathbf{k}$  is the wave-vector of the incoming wave at position  $\mathbf{r}$ . The outgoing scattered wave is spherical and the wave function (from a single nucleus) can hence be described as

$$\psi_f(\mathbf{r}) = -\psi_0 f(\lambda, \theta) \frac{e^{i\mathbf{k}' \cdot (\mathbf{r} - \mathbf{r}_j)}}{|\mathbf{r} - \mathbf{r}_j|}, \quad (3.5)$$

where  $\mathbf{k}'$  is the wave-vector of the outgoing wave at position  $\mathbf{r}_j$ . Figure 3.1 displays an illustration of elastic scattering of an incoming plane to a final spherical wave.



**Figure 3.1:** Elastic scattering where the initial (plane) wave, with wave vector  $\mathbf{k}$  and wavelength  $\lambda_i$ , collide with an atom at position  $\mathbf{r}_j$ . The final (spherical) wave, with wave  $\mathbf{k}'$  and wavelength  $\lambda_f$ , is scattered onto an area,  $dA$ .

The function  $f(\lambda, \theta)$  given in Equation (3.5) is different depending on whether the incoming particles are neutrons or photons.<sup>79</sup> For neutrons,  $f(\lambda, \theta) = b$ , where  $b$  is the scattering length of the nucleus, which is a measure of how strong the scattering from the nucleus is.<sup>79</sup> The scattering length depends on the nucleus as well as its spin relative to the neutron, which means that the scattering length changes depending on the isotope and that there are two different values of the scattering length if the nucleus has a non-zero spin.<sup>79,80</sup> Thus, Equation (3.5) can be written as

$$\psi_f(\mathbf{r}) = -\psi_0 e^{i\mathbf{k} \cdot \mathbf{r}_j} b_j \frac{e^{i\mathbf{k}' \cdot (\mathbf{r} - \mathbf{r}_j)}}{|\mathbf{r} - \mathbf{r}_j|}, \quad (3.6)$$

for neutron scattering. In contrast, the scattering length in x-ray scattering depends on the atomic number,  $Z$ , as the x-rays do not interact with the nucleus but rather the electron cloud surrounding the atom. Therefore,  $f(\lambda, \theta)$  is characterised differently

using x-ray scattering. The function decays with increasing  $\theta$  and decreasing  $\lambda$  and is a function of  $\sin \theta/\lambda$ , which leads to a decrease in  $f(\lambda, \theta)$  with  $Q$ . This decrease is known as the atomic form factor.<sup>79</sup>

For a system consisting of  $N$  atoms, one needs to consider multiple spherical waves giving rise to interference. The outgoing wave function is then described as the sum of  $N$  individual wave functions,

$$\psi_f = -\psi_0 e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{j=1}^N \frac{b_j}{|\mathbf{r} - \mathbf{r}_j|} e^{i\mathbf{Q}\cdot\mathbf{r}_j}. \quad (3.7)$$

The interaction of a neutron with a nucleus can either lead to it being absorbed or scattered. The cross section,  $\sigma$  can therefore be divided into

$$\sigma = \sigma_a + \sigma_s, \quad (3.8)$$

where  $\sigma_a$  and  $\sigma_s$  are the absorption and scattering cross section, respectively. While the absorption cross section is related to  $\mathbf{k}$ , the total scattering cross section is related to the scattering length through<sup>79</sup>

$$\sigma_s = 4\pi\langle b^2 \rangle, \quad (3.9)$$

where  $\langle b^2 \rangle = \frac{1}{N} \sum_{i=1}^N b_i^2$ . In neutron scattering, the total cross section contains a coherent,  $\sigma^{\text{coh}}$ , and an incoherent,  $\sigma^{\text{inc}}$ , part.<sup>79,81</sup> The difference between the respective parts is that the former contains both the scattering due to different nuclei and contributions due to self-scattering, whereas the incoherent part only contains self-scattering and is a result of the fact that different nuclei of the same atom type can have different spin states. The coherent and incoherent scattering cross sections are given by  $\sigma^{\text{coh}} = 4\pi\langle b \rangle^2$  and  $\sigma^{\text{inc}} = 4\pi(\langle b^2 \rangle - \langle b \rangle^2)$ , respectively.<sup>79,81</sup>

### 3.1.3 Neutron Diffraction

Elastic coherent neutron scattering is known as neutron diffraction which is an experimental technique used to gain information regarding the structural correlation lengths in the system of interest. The atomic structure of a liquid or an amorphous

material is often described by the pair correlation function,  $g(r)$ , which can be seen as the probability of finding another atom within a spherical shell with inner and outer radius of  $r$  and  $r + dr$ , respectively. The following section will describe how the pair correlation function relates to the data of a neutron scattering experiment.

The output from neutron diffraction is the double differential scattering cross section,  $\frac{d^2\sigma}{d\Omega dE_f}(Q)$ , which is defined as the fraction of neutrons, of final energy,  $E_f$ , scattered into solid angle  $d\Omega$  with energies  $\in [E_f; E_f + dE_f]$ . As mentioned in section 3.1.1, one does not have to account for the energy dependence of the scattered neutrons in elastic scattering. Therefore, the differential scattering cross section,  $\frac{d\sigma}{d\Omega}$ , can be obtained by integrating over all final energies,

$$\frac{d\sigma}{d\Omega} = \int_{-\infty}^{\infty} \frac{d^2\sigma}{d\Omega dE_f} dE_f. \quad (3.10)$$

In a neutron diffraction experiment the detectors measure the square modulus of the wave function presented in section 3.1.2. By making the assumption that the sample-to-detector distance,  $\mathbf{r}$ , is much larger than the distance between the nuclei,  $\mathbf{r}_j$ , one can approximate that  $|\mathbf{r} - \mathbf{r}_j| \approx r$ . The square modulus of Equation (3.7) then becomes

$$|\psi_f|^2 = \psi_f \psi_f^* = \frac{|\psi_0|^2}{r^2} \sum_{i=1}^N b_i e^{i\mathbf{Q} \cdot \mathbf{r}_i} \cdot \sum_{j=1}^N b_j^* e^{-i\mathbf{Q} \cdot \mathbf{r}_j}, \quad (3.11)$$

where  $i$  and  $j$  are indices representing two different scattering sites and  $|\psi_0|^2 = \Psi$  is the flux of the incident wave. By defining the fraction incident neutrons scattered onto the area,  $dA$ , as  $d\sigma = \frac{|\psi_f|^2}{|\psi_0|^2} dA$ , and the solid angle as  $d\Omega = dA/r^2$ , the differential scattering cross section can be written as

$$\frac{d\sigma}{d\Omega} = \frac{|\psi_f|^2}{|\psi_0|^2} \frac{dA}{d\Omega} = \sum_{i=1}^N b_i e^{i\mathbf{Q} \cdot \mathbf{r}_i} \cdot \sum_{j=1}^N b_j^* e^{-i\mathbf{Q} \cdot \mathbf{r}_j}. \quad (3.12)$$

As mentioned in section 3.1.2, the scattering length in neutron scattering can vary based on element or isotope and on different spin orientations. For a nucleus with non-zero spin the scattering length has one of two different values,  $b^+$  and  $b^-$ .<sup>79</sup> By averaging the nuclear scattering length,  $b$ , Equation (3.12) can be rewritten as

$$\frac{d\sigma}{d\Omega} = \sum_{i,j}^N \langle b_i b_j^* \rangle e^{i\mathbf{Q} \cdot (\mathbf{r}_i - \mathbf{r}_j)}, \quad (3.13)$$

where the term  $\langle b_i b_j^* \rangle$  can be separated into two different cases,  $i = j$  and  $i \neq j$ .  $i = j$  refers to the "self-scattering" of a nucleus whereas  $i \neq j$  refers to the scattering of two different atoms. The two cases can be written as:<sup>80</sup>

$$\begin{aligned} \text{for } i = j : \quad & \langle b^2 \rangle = \langle b \rangle^2 + \langle (b - \langle b \rangle)^2 \rangle \\ \text{for } i \neq j : \quad & \langle b_i b_j^* \rangle = \langle b_i^* \rangle \langle b_j \rangle = \langle b \rangle^2. \end{aligned}$$

The scattering cross section can hence be divided into coherent and incoherent scattering, enabling Equation (3.13) to be rewritten as

$$\frac{d\sigma}{d\Omega} = \underbrace{\langle b \rangle^2 \sum_{i,j}^N e^{i\mathbf{Q} \cdot (\mathbf{r}_i - \mathbf{r}_j)}}_{\text{coherent scattering}} + \underbrace{N(\langle b^2 \rangle - \langle b \rangle^2)}_{\text{incoherent scattering}}. \quad (3.14)$$

The coherent part is of most interest when it comes to diffraction studies since, as seen in Equation (3.14), the incoherent part does not include any structural parameters. The coherent part is usually written as

$$\frac{d\sigma}{d\Omega}^{\text{coh}} = N \langle b \rangle^2 S^{\text{coh}}(\mathbf{Q}), \quad (3.15)$$

where  $S^{\text{coh}}(\mathbf{Q})$  is referred to as the coherent structure factor. To understand how the structure factor can be used to describe the structure of a material one needs to understand the definition of the pair correlation function,  $g(r)$ . In a crystalline material the atoms are highly ordered. In contrast, the atoms of an amorphous material are more randomly distributed. The pair correlation function comes from the autocorrelation function

$$G(\mathbf{r}) = \frac{1}{N} \int n(\mathbf{r}') n(\mathbf{r}' + \mathbf{r}) d\mathbf{r} = \frac{1}{N} \sum_{ij}^N \delta(\mathbf{r} + \mathbf{r}_j - \mathbf{r}_i) \quad (3.16)$$

of the local atomic density,  $n(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$ , where  $\delta$  is the Dirac delta function. By distinguishing the cases of  $i = j$  and  $i \neq j$  and using the definition of the local atomic density the autocorrelation function can be rewritten as

$$G(\mathbf{r}) = \delta(\mathbf{r}) + \frac{1}{N} \sum_{i \neq j}^N \delta(\mathbf{r} + \mathbf{r}_j - \mathbf{r}_i) = \delta(\mathbf{r}) + \rho g(\mathbf{r}), \quad (3.17)$$

where  $\delta(\mathbf{r})$  describes how an atom correlates with itself and is commonly referred to as the self-correlation and  $\rho$  is the average number density. In addition,  $\rho g(\mathbf{r}) = \frac{1}{N} \sum_{i \neq j}^N \delta(\mathbf{r} + \mathbf{r}_j - \mathbf{r}_i)$  gives the definition of the pair correlation function,  $g(\mathbf{r})$ . By using one of the main properties of the Dirac delta function,  $\sum_i e^{i\mathbf{Q} \cdot \mathbf{r}_i} = \int_V e^{i\mathbf{Q} \cdot \mathbf{r}_i} \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$ , Equation (3.15) can be rewritten as

$$S^{\text{coh}}(\mathbf{Q}) = 1 + \frac{1}{N} \int e^{-i\mathbf{Q} \cdot \mathbf{r}} \sum_{i,j \neq i}^N \delta(\mathbf{r} - (\mathbf{r}_i - \mathbf{r}_j)) d\mathbf{r}. \quad (3.18)$$

Subsequently, the definition of  $g(\mathbf{r})$  can be inserted in Equation (3.18) to obtain

$$S^{\text{coh}}(\mathbf{Q}) = 1 + \rho \int e^{-i\mathbf{Q} \cdot \mathbf{r}} g(\mathbf{r}) d\mathbf{r}, \quad (3.19)$$

which describes the relation between the coherent elastic structure factor and the pair correlation function. For an isotropic liquid the expression of the structure factor can be simplified to

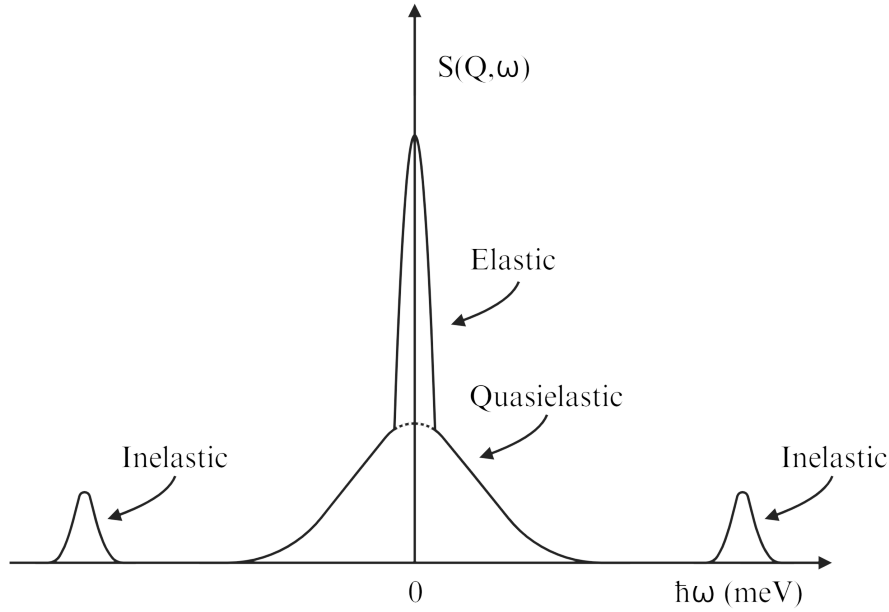
$$S^{\text{coh}}(Q) = 1 + 4\pi\rho \int_0^\infty r^2 g(r) \frac{\sin(Qr)}{Qr} dr. \quad (3.20)$$

#### 3.1.4 Quasielastic Neutron Scattering (QENS)

In this thesis, the main focus has been to investigate the stabilising and anti-aggregating effect disaccharides have on proteins in aqueous solutions. In such systems the presence of water is extremely important as it induces translational and rotational motions within the biological macromolecules. One of the advantages with neutrons is that they can be used to study the dynamics of matter on an atomic scale. As mentioned in section 3.1.1, the inelastic part ( $\Delta E \neq 0$ ) of the scattering provides information about the dynamics of the atoms. In Figure 3.2, the scattering function as a function of energy transfer is shown. The difference between elastic and inelastic scattering is clear as elastic scattering is seen as a very sharp peak around  $\hbar\omega = 0$  while inelastic scattering is seen as peaks at the



respective energy transfer. The broadening of the peak around  $\hbar\omega = 0$  is referred to as quasielastic scattering. The low energy transfers that are measured in a QENS experiment are related to relatively slow atomic motions such as rotations and translations. The resulting energy transfer is related to a frequency shift according to de Broglie equation expressed in Equation (3.2).



**Figure 3.2:** The scattering function as a function of energy transfer. The smaller peaks  $\hbar\omega \neq 0$  are the inelastic scattering, the sharp peak around  $\hbar\omega = 0$  is the elastic scattering, and the broadening of the latter is the quasielastic scattering.

The theory behind QENS is very similar to the case of elastic neutron scattering described in section 3.1.3. However, the positions of the atoms are now time-dependent and the energy transfer is not zero, resulting in a double differential scattering cross section given by<sup>82</sup>

$$\frac{d^2\sigma}{d\Omega dE} = \frac{1}{\hbar} \frac{d^2\sigma}{d\Omega d\omega} = \frac{k'}{k} \frac{1}{2\pi} \int_0^\infty e^{-i\omega t} \sum_{i,j}^N \langle b_i b_j^* \rangle e^{i\mathbf{Q} \cdot (\mathbf{r}_i(0) - \mathbf{r}_j(t))} dt, \quad (3.21)$$

which can be divided into coherent and incoherent scattering

$$\begin{aligned}
\frac{d^2\sigma}{d\Omega dE} = & \underbrace{\frac{k'}{k} \frac{1}{2\pi} \langle b \rangle^2 \int_0^\infty e^{-i\omega t} \sum_{i,j}^N e^{i\mathbf{Q} \cdot (\mathbf{r}_i(0) - \mathbf{r}_j(t))} dt}_{\text{Coherent scattering}} \\
& + \underbrace{\frac{k'}{k} \frac{1}{2\pi} (\langle b^2 \rangle - \langle b \rangle^2) \int_0^\infty e^{-i\omega t} \sum_i^N e^{i\mathbf{Q} \cdot (\mathbf{r}_i(0) - \mathbf{r}_i(t))} dt}_{\text{Incoherent scattering}},
\end{aligned} \tag{3.22}$$

where the latter corresponds to a pure self-scattering part. Similarly to the elastic case, described in section 3.1.3,  $S^{\text{coh}}(\mathbf{Q}, \omega) = \frac{1}{N} \int_0^\infty e^{-i\omega t} \sum_{i,j}^N e^{i\mathbf{Q} \cdot (\mathbf{r}_i(0) - \mathbf{r}_j(t))} dt$  and  $S^{\text{inc}}(\mathbf{Q}, \omega) = \frac{1}{N} \int_0^\infty e^{-i\omega t} \sum_i^N e^{i\mathbf{Q} \cdot (\mathbf{r}_i(0) - \mathbf{r}_i(t))} dt$ , which enables division of the double differential scattering cross section into

$$\frac{d^2\sigma^{\text{coh}}}{d\Omega dE} = \frac{k'}{k} \frac{N}{2\pi} \langle b \rangle^2 S^{\text{coh}}(\mathbf{Q}, \omega) \tag{3.23}$$

and

$$\frac{d^2\sigma^{\text{inc}}}{d\Omega dE} = \frac{k'}{k} \frac{N}{2\pi} (\langle b^2 \rangle - \langle b \rangle^2) S^{\text{inc}}(\mathbf{Q}, \omega). \tag{3.24}$$

In section 3.1.3 the particles were assumed to be static when discussing the correlation function. However, to understand the dynamics of particles, it is often useful to include the displacements over time, which is done in the dynamic pair correlation function, referred to as the van Hove function:<sup>83</sup>

$$G(\mathbf{r}, t) = \frac{1}{N} \sum_{i,j}^N \delta(\mathbf{r} + \mathbf{r}_i(0) - \mathbf{r}_j(t)). \tag{3.25}$$

Similarly to the static case,  $G(\mathbf{r}, t)$  can be interpreted as the probability density to find a particle  $i$  at position  $\mathbf{r}$  at time  $t$ , and it is assumed that particle  $j$  is at the origin at  $t = 0$ .<sup>84</sup> By distinguishing between the cases  $i = j$  and  $i \neq j$ , Equation (3.25) can be separated into  $G(\mathbf{r}, t) = G_s(\mathbf{r}, t) + G_d(\mathbf{r}, t)$ , where

$$G_s(\mathbf{r}, t) = \frac{1}{N} \sum_i^N \delta(\mathbf{r} + \mathbf{r}_i(0) - \mathbf{r}_i(t)) \tag{3.26}$$

and

$$G_d(\mathbf{r}, t) = \frac{1}{N} \sum_{i, j \neq i}^N \delta(\mathbf{r} + \mathbf{r}_i(0) - \mathbf{r}_j(t)), \quad (3.27)$$

correspond to the self-correlation, and the distinct part, respectively. The self-correlation and distinct-correlation part can be interpreted as the probability density of finding a particle  $i$  and  $j \neq i$ , respectively, at time  $t$  with the assumption that the particle  $i$  was at the origin at  $t = 0$ .<sup>84</sup> Finally, by using a general property of the Dirac delta function ( $\sum_i e^{i\mathbf{Q} \cdot \mathbf{r}_i} = \int_V e^{i\mathbf{Q} \cdot \mathbf{r}_i} \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$ ) and Equations (3.26) and (3.27), the dynamic structure factors can be divided into

$$\begin{aligned} S^{\text{coh}}(\mathbf{Q}, \omega) &= \int_0^\infty \sum_{i, j}^N e^{i(\mathbf{Q} \cdot (\mathbf{r}_i(0) - \mathbf{r}_j(t)) - \omega t)} dt \\ &= \int_0^\infty \int_{-\infty}^\infty e^{i(\mathbf{Q} \cdot (\mathbf{r}_i(0) - \mathbf{r}_j(t)) - \omega t)} G(\mathbf{r}, t) d\mathbf{r} dt \end{aligned} \quad (3.28)$$

and

$$\begin{aligned} S^{\text{inc}}(\mathbf{Q}, \omega) &= \int_0^\infty \sum_i^N e^{i(\mathbf{Q} \cdot (\mathbf{r}_i(0) - \mathbf{r}_i(t)) - \omega t)} dt \\ &= \int_0^\infty \int_{-\infty}^\infty e^{i(\mathbf{Q} \cdot (\mathbf{r}_i(0) - \mathbf{r}_i(t)) - \omega t)} G_s(\mathbf{r}, t) d\mathbf{r} dt. \end{aligned} \quad (3.29)$$

$S^{\text{coh}}(\mathbf{Q}, \omega)$  and  $S^{\text{inc}}(\mathbf{Q}, \omega)$  are the coherent and incoherent dynamic structure factors related to the total dynamic pair correlation function<sup>a</sup> and the dynamic self-correlation function, respectively, through Fourier transformation.

#### 3.1.4.1 Analysis of QENS

As mentioned in section 3.1.4, QENS is obtained from the broadening of the scattering function as a function of the energy transfer. Therefore, in the analysis of QENS one looks at the broadening and how it depends on  $Q$ . By Fourier transforming the scattering function in the time domain, the intermediate scattering signal  $I(Q, t)$ , can be obtained, which is what has mainly been analysed in this work.

---

<sup>a</sup>The total pair correlation function ( $G(\mathbf{r}, t)$ ) includes both the self-correlation and the distinct part.

The dynamics of atoms in a system can be described by a number of different models, where one example is the use of a stretched exponential function

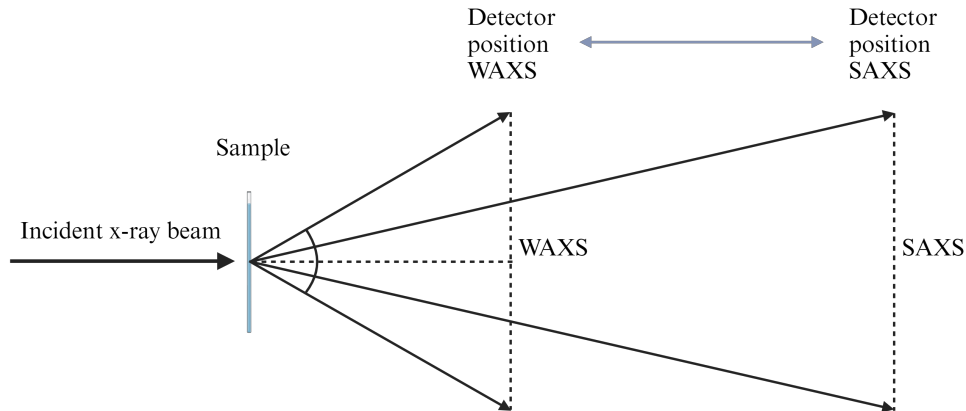
$$I_i(Q,t) = A_i \cdot e^{((-\frac{t}{\tau_i})^{\beta_i})}, \quad (3.30)$$

also referred to as the Kohlrausch-William-Watts (KWW) function,<sup>85</sup> where  $A^i$  is the scattering contribution from atom type  $i$  to the total scattering,  $\tau_i$  is the typical relaxation time, and  $\beta_i$  ( $0 < \beta_i < 1$ ) is the stretching parameter. Low values of  $\beta_i$  means that the atoms have a large distribution of relaxation times. In contrast,  $\beta = 1$  means no stretching which implies that all atoms of type  $i$  have exactly the same relaxation time.

Which relaxation times correspond to which atom type can be determined by the use of isotope substitution of hydrogen for deuterium. As previously mentioned, hydrogen has a significantly larger scattering cross section compared to deuterium, which makes it possible to "filter out" for example the fast dynamics of water by replacing the H<sub>2</sub>O by D<sub>2</sub>O.

#### 3.1.5 X-ray Scattering

To probe different Q-ranges the scattering angle to the detector and/or the detector distance relative to the sample can be adjusted. Placing the detector far away results in a smaller scattering vector and information on relatively large structures can be gained. This also results in scattering at small angles (typically 0.1–10°) and is referred to as small-angle x-ray scattering (SAXS). By the use of SAXS, structural information of materials on a nanometre scale can be studied. By moving the detector closer to the sample, scattering at wider angles (typically > 10°) can be analysed. This method is commonly referred to as wide-angle x-ray scattering (WAXS) and can be used to gain information regarding distances as small as those within amyloid fibrils. Figure 3.3 shows an illustration of the difference between SAXS and WAXS.



**Figure 3.3:** Schematic illustration of the difference in detector position between SAXS and WAXS.

As mentioned in section 3.1.1, the scattering behaviour of x-rays can be explained by Bragg's law (see Equation (3.3)). Like neutron diffraction, the diffraction peaks correspond to repeated distances in the material. The wave vector can be related to real-space distances by combining  $\mathbf{Q} = \mathbf{k}_i - \mathbf{k}_f$  and Equation (3.3), which results in  $Q = 2\pi/d$ , where  $d$  is the distance between repeated structures in the material.<sup>81</sup> SAXS can also be used to obtain information regarding the shape and size of particles in a system.<sup>86</sup>

## 3.2 Empirical Potential Structure Refinement (EPSR) Modelling

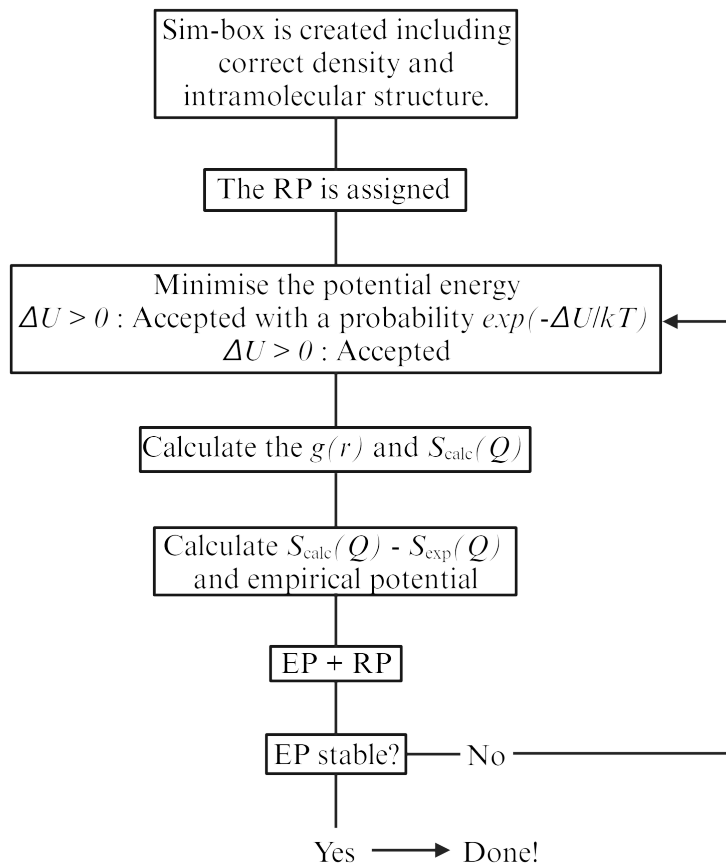
As mentioned in sections 3.1.3 and 3.1.5, neutron diffraction and x-ray scattering data are used to gain information regarding the structure within the system of interest. These data can be used as an experimental basis for an atomic model. Empirical Potential Structure Refinement (EPSR)<sup>87</sup> is an iterative modelling method based on the Monte Carlo method derived from Reverse Monte Carlo (RMC),<sup>88</sup> which is an atomistic simulation technique where the model is improved until its parameters have the greatest consistency with experimental data. To be able to use the EPSR method one typically needs to know quite a lot about the system of interest. Correct

assumptions are needed to avoid non-physical results.

A flowchart to visualise the working procedure of EPSR is presented in Figure 3.4. To start an EPSR simulation a simulation box of correct density and intramolecular structure is initially created. The density is assigned in atoms/Å<sup>3</sup> to ensure the same value independent of the isotope composition. The molecules are defined through harmonic force constants between each atom type in combination with bond angle forces and dihedral angle forces. Moreover, inter- and intramolecular distances are controlled by a Lennard-Jones (LJ) potential combined with a Coulomb potential. This is defined as the reference potential (RP). After the RP is assigned, the potential energy of the system is minimised by the use of the Metropolis Monte Carlo method. This includes random translation, rotation and bending of the molecule. The difference in potential energy before ( $U_{\text{before}}$ ) and after ( $U_{\text{after}}$ ) the move is recorded as

$$\Delta U = U_{\text{after}} - U_{\text{before}}, \quad (3.31)$$

where if  $\Delta U < 0$  the move will be accepted, else the move is accepted with a probability  $\exp(-\frac{\Delta U}{kT})$ . The  $g(r)$  is then calculated and by Fourier transforming it, the simulated  $S_{\text{calc}}(Q)$  is obtained. The difference between the experimentally obtained structure factor,  $S_{\text{exp}}(Q)$ , and the simulated  $S_{\text{calc}}(Q)$  is then used to calculate the empirical potential (EP), which is added to the RP. The process is then iterated, starting from the step of minimising the potential energy until the EP is stable. More information about how EPSR works can be found in the EPSR manual.<sup>89</sup>



**Figure 3.4:** Flowchart of procedure of the EPSR simulation method.

### 3.3 Molecular Dynamics (MD) Simulations

The simulation method EPSR, described in section 3.2, is an excellent tool for analysing the structural arrangement of experimentally measured systems. However, the primary focus of EPSR is structural refinement and may not provide insights into the dynamic behaviours of systems over time. Additionally, the high-quality experimental data required for EPSR may not always be feasible to obtain. To complement EPSR and gain a deeper understanding of the dynamics of a molecular system, theoretical models are indispensable. One such powerful method is classical MD simulations. The first molecular dynamics simulation was performed in 1957 to study the properties of hard sphere fluids.<sup>90</sup> However, MD simulations can be used in a number of different areas. In this thesis, MD simulations have been used to study

the dynamics of aqueous solutions containing sucrose and trehalose. This section offers a concise introduction to the basic principles of classical MD simulations.

To perform MD simulations one first needs to create an initial state of the system, the simulation box. In the initial state of the system one needs to define the initial atomic positions and velocities as well as the types of particles in the system. For a sufficiently simple system this step can be done manually. Once the initial state is defined, interaction potentials between the particles in the system need to be introduced. This is done by the use of existing force fields. There are many different force fields that are used for different types of particles and purposes. Examples of such force fields are the general amber force field (GAFF)<sup>91</sup> and the chemistry at harvard macromolecular mechanics force field (CHARMM36)<sup>92–94</sup> which were utilised in paper I to perform MD simulations on the carbohydrates and the protein, respectively. The force fields are commonly based on a sum of interatomic potentials, where the total potential is a sum of the intra- and intermolecular potentials and can be expressed as

$$\begin{aligned}
 U_i = & \sum k_i^{\text{bond}}(r_i - r_0)^2 + \sum k_i^{\text{angle}}(\theta_i - \theta_0)^2 + \sum k_i^{\text{dihedral}}(1 + \cos(n_i\phi_i + \delta_i)) \\
 & + \sum_i \sum_{j \neq i} 4\varepsilon_{i,j} \left[ \left( \frac{\sigma_{i,j}}{r_{i,j}} \right)^{12} + \left( \frac{\sigma_{i,j}}{r_{i,j}} \right)^6 \right] + \sum_i \sum_{j \neq i} \frac{q_i q_j}{4\pi\varepsilon_0 r_{i,j}},
 \end{aligned}
 \tag{3.32}$$

where the first term provides the potential energy corresponding to the variation in bond length between two bonded atoms in which  $k_i^{\text{bond}}$  is a specific bond strength and  $r_0$  is the atom position at equilibrium. The second term corresponds to the variation of the bond angle within a molecule, where  $\theta_0$  is preferred with the strength  $k_i^{\text{angle}}$ . For larger molecules (more than four atoms) the third term, in Equation (3.32), is included. It corresponds to the torsion potential where  $\phi_i$  is the torsion angle for atom  $i$ ,  $n_i$  is the number of potential minima the torsion angle can reside in, and  $\delta_i$  is the phase. These three terms correspond to the potential energy of the bonded atoms.<sup>95</sup> The two final terms correspond to the potential energy of the non-bonded atoms, where the fourth term, known as the Lennard-Jones (LJ)



potential, approximates the forces between non-bonded atoms. This term combines the repulsive effect originating from the Pauli exclusion principle, with the attractive van der Waals force. In this context,  $\sigma_{i,j}$  represents the equilibrium distance between atoms  $i$  and  $j$ , and  $\epsilon_{i,j}$  is the depth of the potential energy-well at that distance. The fifth and final term relates to the Coloumb-interactions between two atoms with charges  $q_i$  and  $q_j$  separated by a distance of  $r_{i,j}$ . Finally, by solving Newton's equations for motion

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = \mathbf{F}_i = -\nabla U, \quad (3.33)$$

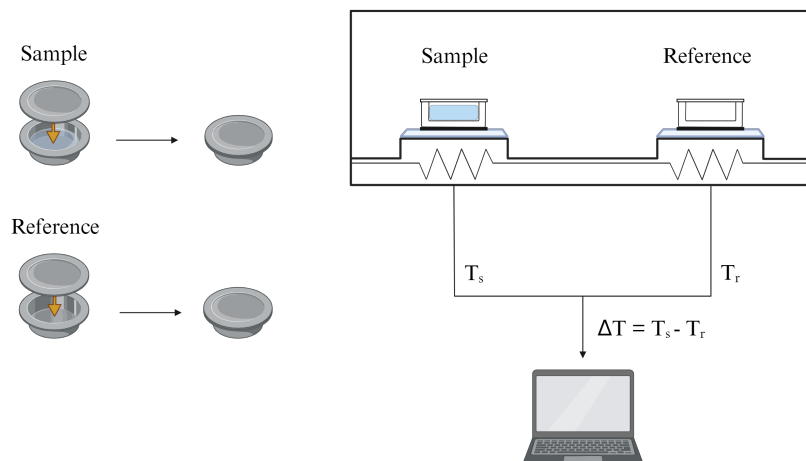
where  $m_i$  and  $r_i$  are the mass and position of atom  $i$ ,  $\mathbf{F}_i$  is the force on atom  $i$ , and  $U$  is the potential energy, the positions of each atom as a function of time can be determined. The movement of the particles can be simulated by the use of a number of algorithms, where the most simple one is the Verlet algorithm.

### 3.4 Differential Scanning Calorimetry (DSC)

When the temperature of a material changes, it undergoes both physical and chemical transformations, which may be either endothermic (absorption of heat) or exothermic (emission of heat). Processes like crystallisation, glass transition, melting, and protein denaturation are examples of such changes. Differential scanning calorimetry (DSC) is a widely-used technique to study these thermal behaviours, as demonstrated in papers III and IV.

During DSC measurements, the sample and a reference (typically an empty sample holder) is heated or cooled at a constant rate ( $^{\circ}\text{C}/\text{min}$ ). Both are placed in identical, hermetically sealed aluminium pans on separate platforms within the DSC cell. The sample and reference are heated or cooled via their respective thermoelectric disk, which would result in a difference in temperature due to the heat capacity of the sample ( $C_p$ ). This difference is detected by thermocouples and is related to the heat flow via Ohm's law,  $q = \Delta T/R$ , where  $q$  is the sample heat flow,  $\Delta T$  the difference in temperature between the sample and reference, and  $R$  is the resistance in the thermoelectric disk.<sup>96</sup> The difference in temperature corresponds

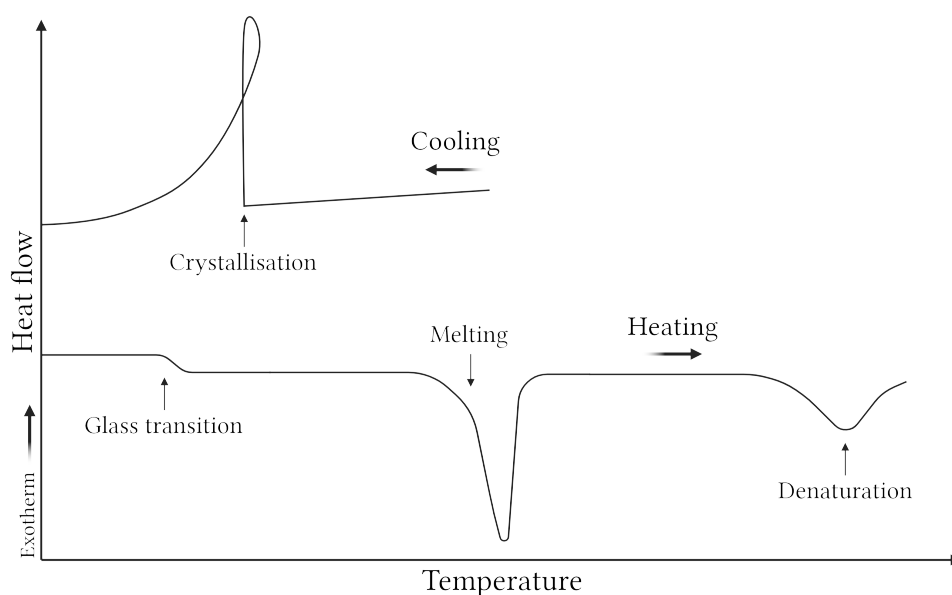
to the energy required to change the temperature of the sample. The working principles of DSC are shown in Figure 3.5.



**Figure 3.5:** The working principles of DSC displaying the sample and reference pans being sealed and placed on separate thermoelectric disks in the DSC cell. The difference in temperature between the sample and reference is transferred to the computer for analysis.

The heat flow to and from the sample is typically constant in the absence of physical or chemical changes. However, exothermic and endothermic processes cause deviations in this heat flow, resulting in distinct features on the DSC curve, which can be seen in Figure 3.6.

In biological systems, which usually contain a significant amount of water, the crystallisation of the water during cooling appears as a dramatic exothermal peak, which can be seen in Figure 3.6. The crystallisation temperature  $T_c$  can be extrapolated from the onset point of this peak. The intensity, as well as the  $T_c$ , may vary with the cooling rate. Conversely, melting of a crystallised sample, an endothermic process requiring energy to disrupt the crystal structure, is indicated as a dip in the DSC curve upon heating. As shown in Figure 3.6, the glass transition can be observed as a step in the DSC curve during heating of the sample and the  $T_g$  can be obtained from the midpoint of the step. In addition, heat-induced protein denaturation is seen as an endothermal dip and the  $T_{den}$  is obtained from the position of the dip minimum.



**Figure 3.6:** Schematic image of a typical DSC curve showing the heat flow as a function of temperature. The crystallisation is seen as an exotherm peak during the cooling process. The glass transition is seen as a step during heating, while both the melting and the protein denaturation are seen as endothermal dips.

### 3.5 Atomic Force Microscopy (AFM)

Atomic force microscopy (AFM) is one of the most commonly used scanning probe microscopy techniques, offering nanometer resolution for examining sample topography.<sup>97</sup> The main principle of AFM involves a resonating cantilever with a sharp tip which raster scans a selected area. It measures surface contact forces between the sample and the probe, generating an image of the sample's topography. This specific mode is referred to as tapping mode and is what is used in paper III. Besides tapping mode, there are two other standard imaging modes: contact mode and non-contact mode. In contact mode, the tip remains in constant contact with the sample and is "dragged" across the surface, generating an image from the feedback signal required to keep the cantilever at a constant height. Non-contact mode, on the other hand, involves positioning the tip in close proximity to the surface, typically within a few Ångströms, without direct contact. The topography image in this mode is derived from the force interactions during the scanning process.



# 4

## Experimental Procedures

### 4.1 Sample Preparation - Mb Samples

To gain new insights regarding the stabilising and anti-aggregating properties of trehalose, in paper I a system consisting of 25 wt% myoglobin, 25 wt% trehalose, and 50 wt% water was compared with a system where trehalose was replaced with sucrose. The samples were prepared by dissolving the protein in sugar solutions. In paper II, similar but more diluted systems containing trehalose were prepared. The weight ratio of the systems were 10 wt% myoglobin, 20 wt% trehalose, and 70 wt% water. As a reference, a two-component system consisting of 33 wt% myoglobin and 67 wt% water was prepared in paper I.

#### 4.1.1 Isotope Substitution

As mentioned in section 3.1.2, the scattering length depends on the isotope. If the molecular structure of the system is unaffected by isotope substitution, then the changes in the structure factor are exclusively due to differences in the scattering lengths. This characteristic is advantageous for creating contrast in neutron scattering experiments. In papers I and II, neutron scattering was used to analyse the role of the disaccharides in stabilising the native state of Mb and the following six isotopically different samples were prepared:

1. Mb D-sugar D<sub>2</sub>O
2. Mb D-sugar H<sub>2</sub>O
3. Mb H-sugar D<sub>2</sub>O
4. Mb H-sugar H<sub>2</sub>O

5. Mb HD-sugar D<sub>2</sub>O
6. Mb HD-sugar HDO

The protonated and deuterated sugars are denoted H-sugar and D-sugar, respectively. The isotope substitution in these sugars only apply to their carbon-bound hydrogens. The sugar in samples 1 and 3 were repeatedly dissolved in D<sub>2</sub>O and dried under vacuum at 70 °C to exchange the hydroxyl hydrogens for deuterium. Mb was freeze dried in either H<sub>2</sub>O (for samples 2 and 4) or D<sub>2</sub>O (for samples 1 and 3) before use, to remove residual water molecules, and to deuterate the exchangeable hydrogens of the protein. Samples 5 and 6 were created from a 50/50 mol% mixture of samples 1 and 3 and samples 1 and 4, respectively.

### 4.2 Sample Preparation - HEWL Samples

In paper III the stabilising and anti-aggregating properties of trehalose and sucrose on HEWL were analysed. Respective disaccharide was dissolved in MQ water during continuous stirring at ambient conditions. For the three-component systems HEWL was dissolved in the sugar solutions, adding small amounts at a time under gentle stirring. The systems contain 20 wt% HEWL, 27 wt% sugar, and 53 wt% MQ water. For the two-component systems HEWL was dissolved in pure MQ water, small amounts at a time during gentle stirring. The system contained 27 wt% HEWL and 73 wt% water. The pH value of the samples was recorded using both a Mettler Toledo Seven2go pH-meter and a pH stick. The pH was altered by adding small amounts of 2 M HCl to the solution until pH 2.0 was reached. The samples were then incubated at different temperatures ( $T_{\text{inc}}$ ). A table containing the different samples with their respective pH value and  $T_{\text{inc}}$  can be found in paper III.

## 4.3 Scattering Experiments

### 4.3.1 Neutron Diffraction

The neutron diffraction experiments in paper I and II were performed on the NIMROD diffractometer<sup>98</sup> at the ISIS Pulsed Neutron and Muon Source, STFC Rutherford Appleton Laboratory, UK. The diffractometer has a broad Q-range of 0.01–50  $\text{\AA}^{-1}$  and the neutrons arrive to the sample with wavelengths between 0.05–10  $\text{\AA}$ . This is what makes NIMROD unique as it can probe structures at different length scale ranging from the interatomic ( $< 1 \text{ \AA}$ ) to the mesoscopic ( $> 300 \text{ \AA}$ ).<sup>98</sup> The samples were measured at 300 K inside 1 mm thick  $\text{Ti}_{0.676}\text{Zr}_{0.324}$  containers, which were sealed with a PTFE O-ring.

### 4.3.2 X-ray Scattering

The fully protonated Mb samples (sample 4 in section 4.1.1) were analysed using x-ray diffraction in paper I, which were performed at the ISIS Pulsed Neutron and Muon Source, STFC Rutherford Appleton Laboratory, UK, using Ag anode x-ray diffractometer (Empyrean),<sup>99</sup> with a wavelength of  $\lambda = 0.5609 \text{ \AA}$ . The samples were placed in silica glass capillaries with a thickness of 2 mm.

The SAXS and WAXS measurements in paper III took place at Chalmers Materials Analysis Laboratory (CMAL) using the Mat: Nordic SAXS/WAXS/GISAXS instrument (SAXSLAB) at RT. The exposure time was 15 min/sample over a Q-range of 0.00138 - 0.30659  $\text{\AA}^{-1}$  and 5 min/sample over a Q-range of 0.00386 - 2.2827  $\text{\AA}^{-1}$  for SAXS and WAXS, respectively, with an x-ray wavelength of  $\lambda = 1.54 \text{ \AA}$ . The samples were placed in quartz capillaries with a thickness of 1.5 mm.

### 4.3.3 Data Corrections

The raw neutron and x-ray diffraction data in papers I and II was "corrected" using the GUDRUN and GUDRUN-X data correction software,<sup>100</sup> respectively, in order to obtain useful structure factors. Apart from the scattering from the sample holder ( $\text{Ti}_{0.676}\text{Zr}_{0.324}$  containers and quartz/glass capillaries for neutron and x-ray scatter-

ing, respectively), the measured scattering data needs to be corrected for multiple scattering events, attenuation, and detector *dead time*.

To provide an absolute scale for the neutron scattering data, it was normalised to a vanadium<sup>a</sup> standard of suitable geometry, which enables the number of detected neutrons at the detector to be divided with the number of theoretically detected neutrons. This provides a normalisation factor for each detector. Unfortunately, there is nothing equivalent for x-rays. However, the method used in GUDRUN-X works rather well.<sup>b</sup>

In section 3.1, it was assumed that multiple scattering events did not occur and that there is no absorption of neutrons in the sample. However, in reality, this assumption is only accurate if the attenuation and multiple scattering cross sections are not too large. By knowing the atom composition and the geometry of the sample, the attenuation and probability of multiple scattering can be calculated and subsequently corrected for. This is then calculated as the background signal and is subtracted from the final structure factor.

Finally, in neutron scattering, corrections for the detector *dead time* need to be performed. The *dead time* refers to the interval between the detection of one neutron and the registration of the next by the detector. It results in some neutrons being neglected. Therefore, this is compensated for using GUDRUN.

### 4.3.4 EPSR Modelling

The EPSR simulation box used to perform the EPSR modelling was created using the Mb structure (PDB ID: 1DWR) from the protein data bank (PDB) and the water and sugar molecules were created using Jmol.<sup>101</sup> The SPC/E force field was used for the water molecules,<sup>102</sup> the OPLS-AA force field for the sugar molecules,<sup>103</sup> and the reference potential for the protein was loosely based on the OPLS-AA force field.<sup>43</sup> The EPSR modelling was performed following the scheme described in section 3.2.

---

<sup>a</sup>Vanadium is used due to its small coherent scattering length.

<sup>b</sup>This method is described in more detail in Ref. 100



#### 4.3.4.1 Coordination numbers

The coordination number,  $N_{\alpha\beta}(r_1, r_2)$ , provides information regarding how many atoms of type  $\beta$  surround the atom type  $\alpha$  within an interval of  $r_1$  and  $r_2$ . It is typically given by<sup>104</sup>

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

where  $c_\beta$  is the atomic number density of atom  $\beta$  and  $g_{\alpha\beta}(r)$  is the partial pair correlation function. Coordination numbers can be divided into coordination shells, which is based on distance separating the two atom types  $\alpha$  and  $\beta$ . The first coordination shell is from  $r_1 = 0$  Å to  $r_2$ , where  $r_2$  is defined as the position of the first minimum point after the first peak.

The coordination numbers presented in papers I and II were calculated from the EPSR produced models using the software VMD.<sup>105</sup>

#### 4.3.5 QENS

The QENS measurements in papers I and II were performed at 300 K using the IRIS spectrometer<sup>106</sup> at the neutron spallation source, ISIS. To be able to extract the dynamics of the water, sugar, and protein molecules separately, isotope substitution was utilised. Samples 2, 3, and 4, listed in section 4.1.1, were used for the QENS measurements. In addition, the following four reference systems were prepared: Mb in D<sub>2</sub>O, Mb in H<sub>2</sub>O, H-sugar in D<sub>2</sub>O, and D-sugar in H<sub>2</sub>O. The samples were placed in annular aluminium alloy cans providing a sample thickness of 0.1 or 0.25 mm<sup>c</sup>.

##### 4.3.5.1 QENS Fitting Procedure

To calculate the diffusion constants,  $D_s$ , presented in papers I and II, the following fitting procedure was performed. The three isotope compositions ( $I^1(Q, t)$ ,  $I^2(Q, t)$ , and  $I^3(Q, t)$ ) were fitted simultaneously by solving the non-linear equation system

---

<sup>c</sup>The thickness depends on the hydrogen content of the sample to ensure less than 10% scattering from the sample

$$\begin{bmatrix} A_P^1 & A_S^1 & A_W^1 \\ A_P^2 & A_S^2 & A_W^2 \\ A_P^3 & A_S^3 & A_W^3 \end{bmatrix} \cdot \begin{bmatrix} e^{-(t/\tau_P)^{\beta_P}} \\ e^{-(t/\tau_S)^{\beta_S}} \\ e^{-(t/\tau_W)^{\beta_W}} \end{bmatrix} = \begin{bmatrix} I^1(Q,t) \\ I^2(Q,t) \\ I^3(Q,t) \end{bmatrix}, \quad (4.2)$$

where P, S, and W stand for protein, sugar, and water, respectively. The superscripts (1, 2, and 3) represent the three isotope compositions. The  $A^i$  of each component was fixed to its relative contribution to the total incoherent scattering cross section. The  $A$  values can be found in Table S3 in the SI of paper I.  $\tau$  and  $\beta$  are the relaxation time and stretching parameter of each molecular component. The average relaxation times were calculated using

$$\langle \tau \rangle = \frac{\tau}{\beta} \Gamma \left( \frac{1}{\beta} \right), \quad (4.3)$$

where  $\Gamma$  is the gamma function. At low  $Q$  values the translational dynamics exhibit a  $Q^2$ -dependence (i.e.  $1/\langle \tau \rangle \propto Q^2$ ). However, sometimes, at higher  $Q$ -values a considerable deviation from the linearity can be observed. Hence, the  $Q^2$ -dependence of  $\langle \tau \rangle$  was fitted using a Gaussian jump-diffusion model<sup>107</sup>

$$\frac{1}{\tau} = \frac{1}{\tau_{\text{res}}} \left[ 1 - \exp \left( -\frac{Q^2 \langle r^2 \rangle}{6} \right) \right], \quad (4.4)$$

where  $\tau_{\text{res}}$  is the average residence time between two consecutive jumps of the moving atoms and  $\langle r^2 \rangle$  is the mean square jump length. These two parameters can be used to calculate the self-diffusion constant,<sup>107</sup>  $D_s = \langle r^2 \rangle / 6\tau_{\text{res}}$ .

## 4.4 DSC Experiments

The calorimetric measurements were performed using a DSC Q2000 (TA Instruments). The instrument is equipped with a liquid nitrogen cooling system (LNCS) and a thermal element, enabling cooling to -180 °C and heating to 550 °C. The sample was placed in a hermetic aluminium pan, sealed using a sealing press. The weights of both the sample and pan were recorded using a balance (Precisica 262SMA-FR) with a precision of 0.01 mg. The measurement of the crystallisation,

glass transition and protein denaturation was performed as described in papers III and IV.

The instrument was calibrated to optimise its performance for a specific temperature range and the utilised heating and cooling rates. The primary Tzero calibration sequence involved setting a baseline for calculating the time constants and capacitances of both the sample and reference sensors, utilising a sapphire<sup>d</sup> standard for accuracy. Following this, a three-point calibration was executed using indium, mercury, and MQ water. Finally, the calibration of heat capacity was conducted using a Temperature Modulated Differential Scanning Calorimetry (TM-DSC) approach, again employing a sapphire standard reference.

## 4.5 AFM

To image the amyloid fibrils in paper III, atomic force microscopy (AFM) was used. The samples were diluted and pipetted onto mica plates which were rinsed 10 times with 200  $\mu$ L MQ water following incubation for 10 min at RT. The surface was then carefully dried using filter paper and nitrogen. The AFM images were recorded on an NT-MDT NTEGRA Prima atomic force microscope with a NSG01 gold-coated single crystal silicon probe (Resonant frequency  $\sim$  150 kHz, force constant  $\sim$  5.1 N/m) using tapping mode (0.5 Hz scan frequency). The images were processed using the software Gwyddion.<sup>108</sup>

---

<sup>d</sup>The heat capacity of sapphire is well-known and stable.



# 5

## Summary of Results

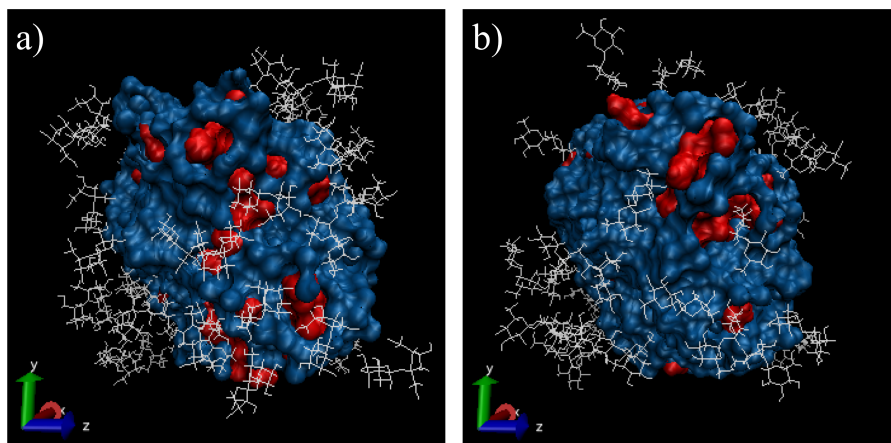
This thesis is based on the work in four different papers. In paper I, trehalose was compared to the structurally similar disaccharides sucrose to provide new insights on the structural and dynamical properties of such a system. The work presented in paper II partly stems from the results obtained from paper I. The focus of paper II is the arrangement of the trehalose molecules in relation to the surface of Mb and the hydration water. The aim of paper III was to investigate inhibition of amyloid fibrils by the use of trehalose and sucrose. Finally, in paper IV the stabilising effect of the two disaccharides was compared at different temperatures and levels of hydration. The following sections includes a summary of the results presented in the four papers.

### 5.1 Paper I

Paper I aims to address the structural and dynamical properties of trehalose and sucrose solutions and how they, together with water molecules, interact with Mb. The purpose was to increase the knowledge regarding the mechanism behind protein stabilisation by the addition of disaccharides and why trehalose has been regarded as superior compared to other disaccharides, such as sucrose.

The structural properties were studied using neutron and x-ray diffraction in combination with EPSR modelling, from which the partial pair correlation functions,  $g(r)$ s, were calculated. The results show a strong peak at around 2.8 Å for both disaccharide solutions. However, the observed peak is slightly more pronounced for the trehalose-containing system. Additionally, the interaction between the oxygen

on the surface of the protein and any oxygen of the sugar molecules was seen to be slightly more pronounced for the sucrose containing system compared to the trehalose. However, it should be mentioned that the interaction is quite small in both cases. These results suggest a more pronounced preferential hydration in the trehalose-containing systems, which can also be seen via the three-dimensional image of the EPSR produced structural model in Figure 5.1(a) for sucrose and Figure 5.1(b) for trehalose.



**Figure 5.1:** Three-dimensional images of the structural model produced with EPSR, showing the arrangement of water and sugar molecules relative to the protein surface, where (a) is the sucrose containing system and (b) is the trehalose-containing system. The red parts are exposed protein surface, the blue parts are the hydration water and the white stick figures are the trehalose and sucrose molecules. The image was obtained from paper I.

The dynamical properties were investigated using both QENS and MD simulations. The experimentally obtained  $I(Q,t)$ s were fitted according to the method mentioned in section 4.3.5.1, enabling calculation of the average relaxation times. By plotting the inverse of the average relaxation times as a function of  $Q^2$ , it was possible to observe that the data could be described rather well using the Gaussian jump-length diffusion model. This was used to calculate the diffusion constants. The results show that the dynamics of both water and sugar in both the two- and three-component systems is faster in the systems with sucrose. The results show that the diffusion constant of water in the three-component system containing sucrose is significantly higher compared to that in the presence of trehalose, indicating faster dynamics in the former. As expected the water dynamics was observed to

be much faster than the dynamics of the two disaccharides. However, there is not considerable difference between the diffusion constant of sucrose and trehalose in the three-component system, suggesting that the sugar dynamics are quite similar.

By the use of classical MD simulations it was possible to investigate the rotational motion around the dihedrals between the two glucose rings of trehalose and the glucose and fructose rings of sucrose. The results showed that the rotational motion in trehalose was slower than the same motion between the rings of sucrose, suggesting that sucrose has a higher rotational mobility. As previously mentioned, trehalose was found to not interact directly with the protein. This result in combination with the slower rotational motion could be the reason for the superior stabilising effect compared to other disaccharides. To further investigate this hypothesis, rotational free energy calculations were performed. The results from these calculations show that the trehalose molecule has a similar mobility around both dihedrals, whereas the mobility around the dihedral  $\psi$  (dihedral closer to the glucose ring) was slower than that around the dihedral  $\varphi$  (dihedral closer to the fructose ring). It could be concluded that the values of rotational potential of mean force for simulations with trehalose are higher, indicating that sucrose has a higher rotational mobility.

In addition, trehalose was observed to bind to fewer oxygen atoms of amino-acids in the backbone of Mb than sucrose. This suggest that, by binding more directly to the protein, sucrose destabilises the native state of the protein by the more direct binding in combination with the faster rotational motions. Hence, it can be concluded that trehalose possesses the ability to stabilise the native state of the protein without almost any direct contact with the surface of the protein.

## 5.2 Paper II

As mentioned in section 5.1, paper I addressed differences in the stabilising effect of the two disaccharides, where it could be concluded that trehalose gives rise to a more pronounced preferential hydration compared to sucrose. Moreover, it was concluded that trehalose is able to stabilise the native state of Mb by almost no interaction with the protein surface. However, due to the high protein concentration in the

systems of paper I, it was not possible to investigate how the disaccharides are arranged in the surroundings of the protein. Therefore, the aim of paper II is to examine the structural arrangement of the trehalose molecules relative the protein surface. In contrast to the high protein concentration (25 wt%) used in paper I, the protein concentration used for paper II was 10 wt%, leaving 20 wt% trehalose and 70 wt% water. Neutron diffraction in combination with EPSR modelling was used to investigate the structural properties, whereas QENS was used to study the dynamical properties. The results from this study are compared with those obtained in paper I.

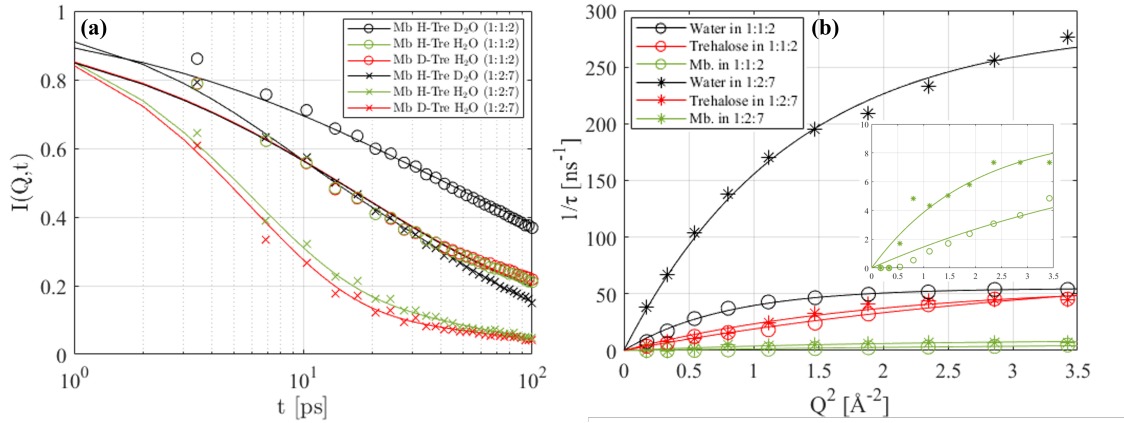
By the use of EPSR modelling, an initial model of the system could be created and the  $g(r)$ s calculated. The results show that the preferential hydration of Mb in the diluted system is even more pronounced compared to the more concentrated system. By studying the interaction between the oxygen atoms on the surface of the protein and any oxygen of trehalose, the difference between the more diluted and the concentrated system is very small. However, by integrating the pair correlation function from 0–3.4 Å the coordination numbers, between the oxygen atoms on the surface of the protein and those of trehalose, was obtained. The coordination numbers tell us that there are slightly more trehalose molecules closer to the protein surface in the more concentrated system. In the more diluted system the peak is shifted slightly towards higher  $r$ -values, which is expected since the system is not as crowded as the more concentrated system leaving more space for the trehalose molecules. Since the peak is shifted the first minimum after the first peak is also shifted and upon integration only part of the peak is included.

An additional model was created where the force field was initially disabled and attractive forces between the oxygen atoms of Mb and trehalose were applied. As the EPSR simulation was initiated the force field was once again enabled and the system was iterated until the model could not be further improved. This was done to investigate whether alternative start configurations would converge with the experimental data. The results showed that both models converged with experimental data. However, the necessary amplitude of the empirical potential for the alternative model was higher than that of the randomly distributed one. The obtained  $g(r)$ s



from the alternative configuration differed considerably from those obtained from the randomly distributed system. The results still suggest that Mb is preferentially hydrated by water molecules but that the trehalose molecules cluster to a great extent close to the water hydration layer.

The dynamical properties were investigated by the use of QENS and were compared to those observed for the three-component system with trehalose presented in paper I. As expected, the dynamics are faster in the more diluted system due to the higher water content. The diffusion constant of the water molecules is very close to that of bulk water, which is to be expected in a system consisting of 70 wt% water. The diffusion constant of the trehalose and Mb molecules was also observed to increase in the more diluted system. A comparison of the average relaxation times in the concentrated and diluted system showed that the dynamics of water and myoglobin are both  $\sim 4$  times faster in the former. Figure 5.2(a) shows the intermediate scattering functions as a function of time at  $Q = 1.06 \text{ \AA}^{-1}$  and Figure 5.2(b) shows the inverse average relaxation times as a function of  $Q^2$ .



**Figure 5.2:** (a) Intermediate scattering functions as a function of time for both the diluted and concentrated systems at  $Q = 1.06 \text{ \AA}^{-1}$ . (b) Inverse average relaxation times as a function of  $Q^2$ . The lines marked with circles represent the concentrated system whereas the lines marked with stars represent the diluted system. The inset shows a magnified view of the inverse average relaxation times as a function of  $Q^2$  for Mb. The image was obtained from paper II.

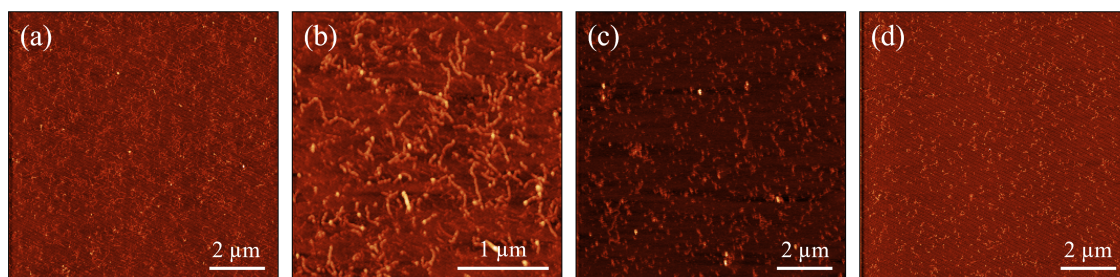
### 5.3 Paper III

In paper III, the protein Mb was exchanged for the protein HEWL, partly to investigate whether the observed effects of the two disaccharides are protein specific but also to investigate their ability to inhibit amyloid fibril formation. The concentrations used in paper III were similar to those in paper I. The two-component system consisted of 27 wt% protein and 73 wt% water, while the three-component systems consisted of 20 wt% protein, 27 wt% sugar, and 53 wt% water. The sugar:water ratio was 1:2 and the protein:water ratio of the two-component system was kept equal to that of the three-component system. As mentioned in section 2.8, HEWL forms amyloid fibrils at low pH values. Therefore, all systems were analysed at both pH 2.0 and 3.5. The initial pH value of the systems was 3.5, which was lowered to pH 2.0 by adding HCl. The systems were incubated at 57 °C for 24 h.

As the point of denaturation can be used as a measure of the conformational stability of the protein, DSC was used to monitor the  $T_{\text{den}}$ . From the results, it was clear that both trehalose and sucrose increase the  $T_{\text{den}}$  of the protein, indicating that both disaccharides stabilise the native fold of the protein. However, it was seen that the latter was able to increase  $T_{\text{den}}$  slightly more than the former. In addition, the results showed that the protein was less stable at pH 2.0, compared to at pH 3.5, and that the presence of the disaccharides was able to increase the stability of the protein also at this pH value.

The systems were studied using the scattering techniques SAXS and WAXS to investigate the structural properties. The presence of an additional peak in the lower Q-range revealed that the protein was observed to aggregate more in the absence of trehalose and sucrose. Furthermore, the protein was observed to aggregate to a greater extent at pH 2.0 compared to at pH 3.5. However, signs of protein aggregation was observed in all systems most likely due to the high protein concentration. To determine whether the larger aggregates formed in the two-component system at pH 2.0 are amyloid fibrils, AFM was used. From the AFM images, it was clear that fully formed amyloid fibrils were only present in this system. However, thinner and shorter fibrillar structures were observed also in the sugar containing systems at pH

2.0. Figure 5.3 shows the AFM images of the different systems at pH 2.0 incubated at 57 °C, where Figure 5.3(a) and (b) show the two-component system and Figure 5.3(c) and (d) the trehalose and sucrose containing systems, respectively.



**Figure 5.3:** AFM image of the two-component system at pH 2.0 (a) with a magnified image (b). AFM image of the three-component system containing trehalose (c) and sucrose (d) at pH 2.0. All systems were incubated at 57 °C for 24 h.

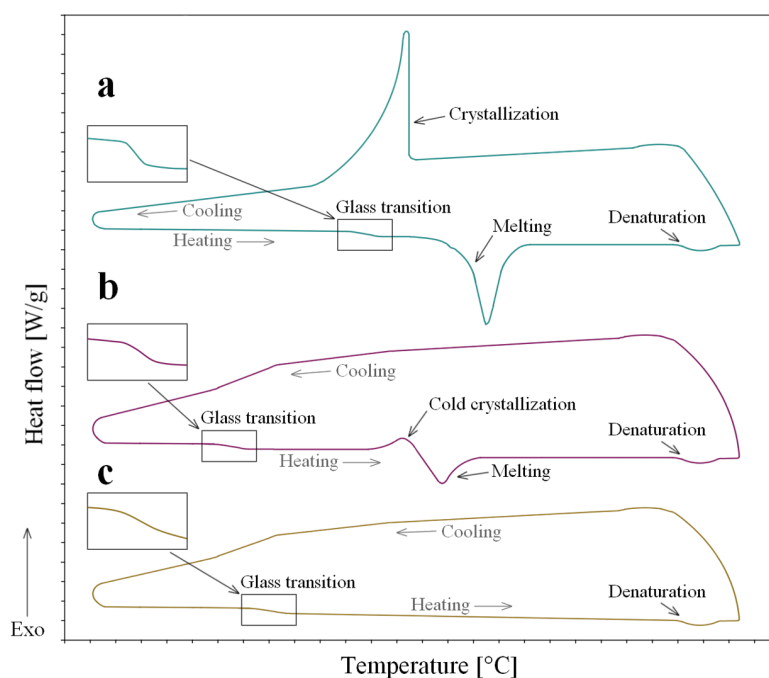
As previously mentioned, the addition of disaccharides significantly increased the  $T_{\text{den}}$ , to such an extent that HEWL molecules remained inadequately unfolded, preventing the formation of fibrillar structures during incubation at 57°C. This suggests that the two sugars inhibit the formation of amyloid fibrils by increasing the  $T_{\text{den}}$ . To further investigate this hypothesis, the sugar containing systems, at pH 2.0, were incubated at higher temperatures than the previously used 57 °C. Since the sucrose containing system increased the  $T_{\text{den}}$  slightly more than trehalose, it was also subjected to a slightly higher  $T_{\text{inc}}$ . After incubation at the new temperatures, the results showed that the system containing sugar aggregated more compared to the corresponding systems incubated at 57 °C. However, interestingly enough, they do not show the same level of amyloid fibril formation as previously observed for the two-component system at pH 2.0. These results indicate that there might be additional factors, other than the increase in  $T_{\text{den}}$ , that inhibit the amyloid fibril formation.

## 5.4 Paper IV

The results from paper I indicated that trehalose is more effective compared to sucrose in stabilising the native state of the protein Mb. However, the results from paper III showed that sucrose is superior in regards to both increasing the  $T_{\text{den}}$  and

inhibiting the formation of amyloid fibrils from HEWL. In paper IV, it was therefore interesting to compare the stabilising effect of the two sugars at different conditions such as temperature and concentration. In this paper, the stability of both Mb and HEWL was evaluated via the  $T_g$  and the  $T_{den}$ , at different sample compositions, by the use of DSC.

When a sample has a sufficiently low water content, crystallisation can be avoided. By adding sugar, the temperature at which crystallisation occurs can be lowered while the sample still contains a considerable amount of water. By investigating the maximal amount of water possible in a sample before it crystallises, and comparing the samples with trehalose to those with sucrose, the amount of amorphous hydration water present in these systems can be deduced. Figure 5.4(a) shows a typical DSC curve of a system with a high water content. Such a system undergoes crystallisation during cooling. In Figure 5.4(b) a system with a lower water content is displayed. In such a system it is typical that the system avoids crystallisation during cooling but undergoes crystallisation during heating (cold crystallisation). Finally, Figure 5.4(c) shows a characteristic DSC curve for system with a sufficiently low water content so that crystallisation is avoided during both cooling and heating. During heating the three aforementioned systems undergo glass transition and protein denaturation, melting will occur during heating in the case of crystallisation. All these thermal processes are marked with arrows in Figure 5.4. In paper IV, these measurements were performed on the sucrose containing samples, which were then compared with the findings for trehalose-containing samples found in Ref. 76.



**Figure 5.4:** Three different characteristic DSC curves displaying the heat flow as a function of temperature. (a) shows the a typical cycle for a system with high water content displaying crystallisation during cooling, (b) shows the cycle for a system where no crystallisation occurs during cooling but during heating (cold crystallisation), and (c) shows a system for which no crystallisation occurs. The image was obtained from paper IV.

The results for the trehalose-containing systems show a linear relation between the maximum water content of the five different trehalose-to-protein ratios, indicating that the maximum amount amorphous water in the systems should be a weighted average of the binary systems. This suggests that basically no trehalose molecules replace the water at the protein surface. However, the results obtained for the sucrose containing systems show that the relation between the maximum water content of the five different sucrose-to-protein ratios is non-linear and that the water content at intermediate sucrose-to-protein ratios is considerably lower compared to the trehalose-containing samples. This suggests that the protein and sucrose molecules are not fully hydrated at these concentrations and that some sucrose molecules replace water at the protein surface. These results confirm that both systems are in agreement with the preferential hydration model. This agreement is more pronounced for trehalose, which was also shown in paper I.

Furthermore, in paper IV it was seen that the  $T_g$  rapidly decreases with an increasing water content at low water concentrations for both the trehalose and sucrose containing systems. However, at higher water concentrations partial crystallisation occurs. As a consequence, the  $T_g$  of such samples is higher. Increasing the water concentration further only results in an increase of the amount of ice in the system and not an increase of the amount hydration water. Therefore, the  $T_g$  is seen to be relatively unaffected by such increments. Since the protein denaturation occurs at higher temperatures, the  $T_{den}$  is unaffected by crystallisation. Therefore, the  $T_{den}$  is seen to decrease linearly with increasing water content across the entire concentration range. Similar to the results obtained in paper III, both sucrose and trehalose increase the  $T_{den}$  of the protein, the former slightly more than the latter. As expected, a higher sugar-to-protein ratio results in a higher  $T_{den}$ .

In addition, the occurrence of possible synergistic effects was explored through different combinations of the two sugars. It was seen that both the  $T_g$  and the  $T_{den}$  change linearly when one disaccharide is exchanged by the other, indicating that the samples containing both trehalose and sucrose behave as a linear combination of the samples containing either sucrose or trehalose. This suggests no synergistic effects when combining the two disaccharides.

# 6

## Conclusion and Future Perspectives

By comparing trehalose with the structurally similar disaccharide sucrose, the knowledge of the stabilising effect of the former could be increased. In paper I, we showed that trehalose exhibits a more pronounced preferential hydration and stabilises the native state of the protein with close to no direct interaction with the protein, as opposed to sucrose. Furthermore, it was shown that the dynamics were slowed down more by the presence of trehalose compared to sucrose.

In order to explore the possibility of a sugar shell surrounding the hydration water, the arrangement of the sugar molecules was in focus in paper II. As expected the preferential hydration was more pronounced and the dynamics faster for this more diluted system. However, no clear peak was observed to prove the presence of a hypothesised sugar shell. An alternative start configuration was compared to the randomly distributed one and the results suggest that both models converge with experimental data, which indicates a possibility that both structures could explain the structure of the system. Future studies will focus on creating a completely homogeneously distributed system and calculate the  $g(r)$ s to compared with the already presented start configurations in order to gain deeper insights regarding the arrangement of the sugar molecules.

The two disaccharides were compared against each other again in papers III and IV. In paper IV we confirm the more pronounced preferential hydration caused by the presence of trehalose by analysing the protein and sugars ability to remain fully hydrated at different protein:sugar ratios. We also showed that the  $T_g$  of trehalose is

higher than that of sucrose, which implies that trehalose is able to keep the protein in that glassy state at a higher temperature compared to sucrose. In both papers III and IV, the  $T_{\text{den}}$  of the protein was observed to increase in the presence of both trehalose and sucrose. Surprisingly, this effect was seen to be slightly more pronounced for the latter. These results suggest that trehalose and sucrose both stabilise the native state of the protein but more or less effectively depending on temperature and composition of the system.

The inhibitory effect the two disaccharides have upon amyloid fibril formation, which was analysed in paper III, was seen to be slightly more effective for sucrose compared to trehalose. One reason for this can be the slightly higher  $T_{\text{den}}$ , hindering the unfolding of the protein. By increasing the incubation temperature of the sugar-containing systems, it could be concluded that the inhibitory effect of the disaccharides is not only due to them shifting the point of denaturation. Answers to the mechanism behind the inhibitory effect of the sugars might be found by studying the dynamics of the system. The plan for future studies is to investigate this using dielectric spectroscopy, which is an ideal tool for this purpose as it probes within the timescale in which the fibrils are expected to be within. In addition, the dynamics of the HEWL amyloid fibrils can be analysed using the neutron spin echo (NSE) technique. NSE is a powerful technique to study slow relaxation and diffusion phenomenon. It is the neutron scattering technique that gives access to the largest lengthscales and the longest timescales, which makes it a good technique to study the dynamics of the fibrils. NSE is a special type of QENS experiment based on the change in velocity of the neutrons in the scattering process.

In addition to the dynamics measurements, the plan is to analyse the already obtained data from neutron diffraction experiments on the systems presented in paper III. The plan is to create model systems using the software Dissolve.<sup>109</sup> The data will be analysed to increase our understanding of the structural properties of trehalose and sucrose as amyloid fibril inhibitors.



# Bibliography

- [1] Sipe JD, Cohen AS. Review: History of the Amyloid Fibril. *J Struct Biol.* 2000 6;130(2-3):88-98.
- [2] Hjärnfonden. Vad är Alzheimers sjukdom? [Internet]; 2021. Available from: <https://www.hjarnfonden.se/om-hjarnan/diagnoser/alzheimers/>.
- [3] Iadanza MG, Jackson MP, Hewitt EW, Ranson NA, Radford SE. A new era for understanding amyloid structures and disease. *Nat Rev Mol Cell Biol.* 2018 12;19(12):755-73.
- [4] Härd T, Lendel C. Inhibition of Amyloid Formation. *J Mol Biol.* 2012 8;421(4-5):441-65.
- [5] Elliott SR. Preparation. In: *Physics of amorphous materials*. 2nd ed. John Wiley & Sons; 1990. p. 1-27.
- [6] Elliott SR. Glasses. In: *Physics of amorphous materials*. 2nd ed. John Wiley & Sons; 1990. p. 29-68.
- [7] Soper AK. Water and ice structure in the range 220 - 365K from radiation total scattering experiments. 2014 11.
- [8] Stenesh J. Frameworks of Biochemistry. In: *Biochemistry*; 1998. p. 3-23.
- [9] Stanley HE. Unsolved Mysteries of Water in Its Liquid and Glass States. *MRS Bulletin.* 1999 5;24(5):22-30.
- [10] Ben-Amotz D. Hydration-Shell Vibrational Spectroscopy. *JACS.* 2019 7;141(27):10569-80.
- [11] Zhang Y, Lewis NHC, Mars J, Wan G, Weadock NJ, Takacs CJ, et al. Water-in-Salt LiTFSI Aqueous Electrolytes. 1. Liquid Structure from Combined Molecular Dynamics Simulation and Experimental Studies. *J Phys Chem B.* 2021 5;125(17):4501-13.
- [12] Petkov V, Ren Y, Suchomel M. Molecular arrangement in water: random but not quite. *Journal of Physics: Condensed Matter.* 2012 4;24(15):155102.
- [13] Narten AH, Danford MD, Levy HA. X-ray diffraction study of liquid water in the temperature range 4–200°C. *Discuss Faraday Soc.* 1967;43(0):97-107.
- [14] Bergmann U, Di Cicco A, Wernet P, Principi E, Glatzel P, Nilsson A. Nearest-neighbor oxygen distances in liquid water and ice observed by x-ray Raman based extended x-ray absorption fine structure. *J Chem Phys.* 2007 11;127(17).
- [15] Powar C, Chatwal G. Carbohydrates. In: *Biochemistry*. Global media; 2007. p. 117-55.
- [16] Stenesh J. Carbohydrates. In: *Biochemistry*; 1998. p. 117-40.
- [17] Miljković M. Chemistry of the Glycosidic Bond. In: *Carbohydrates*. New York, NY: Springer New York; 2010. p. 323-421.

- [18] Nguyen K, KC S, Gonzalez T, Tapia H, Boothby TC. Trehalose and tardigrade CAHS proteins work synergistically to promote desiccation tolerance. *Commun Biol.* 2022 10;5(1):1046.
- [19] Hoekstra FA, Golovina EA, Buitink J. Mechanisms of plant desiccation tolerance. *Trends Plant Sci.* 2001 9;6(9):431-8.
- [20] Viner RI, Clegg JS. Influence of trehalose on the molecular chaperone activity of p26, a small heat shock/ $\alpha$ -crystallin protein. *Cell Stress Chaperones.* 2001;6(2):126.
- [21] Buitink J, Hemminga MA, Hoekstra FA. Is There a Role for Oligosaccharides in Seed Longevity? An Assessment of Intracellular Glass Stability. *Plant Physiol.* 2000 4;122(4):1217-24.
- [22] Thakral S, Sonje J, Munjal B, Suryanarayanan R. Stabilizers and their interaction with formulation components in frozen and freeze-dried protein formulations. *Adv Drug Deliv Rev.* 2021 6;173:1-19.
- [23] Stick RV, Williams SJ. Disaccharides, Oligosaccharides and Polysaccharides. In: *Carbohydrates: The Essential Molecules of Life.* Elsevier; 2009. p. 321-41.
- [24] Chen A, Gibney PA. Dietary Trehalose as a Bioactive Nutrient. *Nutrients.* 2023 3;15(6):1393.
- [25] Iturriaga G, Suárez R, Nova-Franco B. Trehalose Metabolism: From Osmoprotection to Signaling. *Int J Mol Sci.* 2009 9;10(9):3793-810.
- [26] Tombari E, Salvetti G, Ferrari C, Johari GP. Kinetics and Thermodynamics of Sucrose Hydrolysis from Real-Time Enthalpy and Heat Capacity Measurements. *J Phys Chem B.* 2007 1;111(3):496-501.
- [27] Green JL, Angell CA. Phase relations and vitrification in saccharide-water solutions and the trehalose anomaly. *J Phys Chem.* 1989 4;93(8):2880-2.
- [28] Hancock BC, Shamblin SL. Water vapour sorption by pharmaceutical sugars. *Pharmaceutical Science & Technology Today.* 1998 11;1(8):345-51.
- [29] Crowe JH, Carpenter JF, Crowe LM. The role of vitrification in anhydrobiosis. *Annu Rev Physiol.* 1998 10;60(1):73-103.
- [30] Surana R, Pyne A, Suryanarayanan R. Effect of Aging on the Physical Properties of Amorphous Trehalose. *Pharm Res.* 2004 5;21(5):867-74.
- [31] Branca C, Magazù S, Maisano G, Migliardo P. Anomalous cryoprotective effectiveness of trehalose: Raman scattering evidences. *J Chem Phys.* 1999 7;111(1):281-7.
- [32] Branca C, Magazù S, Migliardo F, Migliardo P. Destructuring effect of trehalose on the tetrahedral network of water: a Raman and neutron diffraction comparison. *Phys A: Stat Mech Appl.* 2002 2;304(1-2):314-8.
- [33] Branca C, Magazù V, Maisano G, Migliardo F, Soper AK. Study on Destructuring effect of trehalose on water by neutron diffraction. *Appl Phys A.* 2002 12;74(0):s450-1.
- [34] Branca C, Maccarrone S, Magazù S, Maisano G, Bennington SM, Taylor J. Tetrahedral order in homologous disaccharide-water mixtures. *J Chem Phys.* 2005 5;122(17).
- [35] Magazù S, Migliardo F, Telling MTF. Structural and dynamical properties of water in sugar mixtures. *Food Chemistry.* 2008 2;106(4):1460-6.
- [36] Stenesh J. Proteins. In: *Biochemistry*; 1998. p. 47-82.

- 
- [37] Gomes CM, Faísca PFN. Protein Structure. In: Protein Folding: An Introduction; 2018. p. 1-10. Available from: <http://www.springer.com/series/11958>.
- [38] Vanzi F, Vladimirov S, Knudsen CR, Goldman YE, Cooperman BS. Protein synthesis by single ribosomes. *RNA*. 2003 10;9(10):1174-9.
- [39] Branden C, Tooze J. The building blocks. In: Introduction to protein structure. 2nd ed. New York, NY: Garland Publishing, Inc.; 1999. p. 3-12.
- [40] Walsh G. Protein structure and engineering. In: Proteins: Biochemistry and Biotechnology. 2nd ed. John Wiley & Sons; 2014. p. 25-65.
- [41] Branden C, Tooze J. Motifs of protein structure. In: Introduction to protein structure; 1999. p. 13-34.
- [42] Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature*. 1958 3;181(4610):662-6.
- [43] Chu K, Vojtchovský J, McMahon BH, Sweet RM, Berendzen J, Schlichting I. Structure of a ligand-binding intermediate in wild-type carbonmonoxy myoglobin. *Nature*. 2000 2;403(6772):921-3.
- [44] Tan S, Tatsumura Y. Alexander Fleming (1881–1955): Discoverer of penicillin. *Singapore Medical Journal*. 2015 7;56(07):366-7.
- [45] Jafari M, Mehrnejad F. Molecular Insight into Human Lysozyme and Its Ability to Form Amyloid Fibrils in High Concentrations of Sodium Dodecyl Sulfate: A View from Molecular Dynamics Simulations. *PLoS One*. 2016 10;11(10):e0165213.
- [46] Hankiewicz J, Swierczek E. Lysozyme in human body fluids. *Clin Chim Acta*. 1974 12;57(3):205-9.
- [47] Woods CM, Hooper DN, Ooi EH, Tan LW, Carney AS. Human Lysozyme has Fungicidal Activity against Nasal Fungi. *Am J Rhinol Allergy*. 2011 7;25(4):236-40.
- [48] Pleyer C, Flesche J, Saeed F. Lysozyme amyloidosis – a case report and review of the literature. *Clinical Nephrology Case Studies*. 2015 12.
- [49] Granel B, Valleix S, Serratrice J, Chérin P, Texeira A, Disdier P, et al. Lysozyme Amyloidosis. *Medicine*. 2006 1;85(1):66-73.
- [50] Sipe JD, Benson MD, Buxbaum JN, Ikeda Si, Merlini G, Saraiva MJM, et al. Amyloid fibril proteins and amyloidosis: chemical identification and clinical classification International Society of Amyloidosis 2016 Nomenclature Guidelines. *Amyloid*. 2016 10;23(4):209-13.
- [51] Tokunaga Y, Sakakibara Y, Kamada Y, Watanabe Ki, Sugimoto Y. Analysis of Core Region from Egg White Lysozyme Forming Amyloid Fibrils. *International Journal of Biological Sciences*. 2013;9(2):219-27.
- [52] Ow SY, Dunstan DE. The effect of concentration, temperature and stirring on hen egg white lysozyme amyloid formation. *Soft Matter*. 2013;9(40):9692.
- [53] Walsh G. Proteins and proteomics. In: Proteins: Biochemistry and Biotechnology. 2nd ed. John Wiley & Sons; 2014. p. 1-25.
- [54] Hirai M, Hagiwara Y, Takeuchi K, Kimura R, Onai T, Kawai-Hirai R, et al. Thermal unfolding and refolding of protein under osmotic pressure clarified by wide-angle X-ray scattering. *Thermochim Acta*. 2012 3;532:15-21.

- [55] Diamond R. Real-space refinement of the structure of hen egg-white lysozyme. *J Mol Biol.* 1974 1;82(3):371-91.
- [56] Gomes CM, Faísca PFN. Protein Folding - Why is Structure Acquired? In: *Protein Folding: An Introduction*; 2018. p. 10-20.
- [57] Levinthal C. ARE THERE PATHWAYS FOR PROTEIN FOLDING ?; 1968. 1.
- [58] Dill KA, Chan HS. From Levinthal to pathways to funnels. *Nat Struct Mol Biol.* 1997 1;4(1):10-9.
- [59] Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins: Structure, Function, and Bioinformatics.* 1995 3;21(3):167-95.
- [60] Koo EH, Lansbury PT, Kelly JW. Amyloid diseases: Abnormal protein aggregation in neurodegeneration. *PNAS.* 1999 8;96(18):9989-90.
- [61] Alzheimer's Disease and Related Dementias;. Available from: <https://cdc.gov/aging/aginginfo/alzheimers.htm>.
- [62] Who has Parkinson's?;. Available from: <https://www.parkinson.org/understanding-parkinsons/statistics>.
- [63] Gazit E. The "Correctly Folded" State of Proteins: Is It a Metastable State? *Angew Chem Int Ed.* 2002 1;41(2):257.
- [64] Meersman F, Dobson CM. Probing the pressure-temperature stability of amyloid fibrils provides new insights into their molecular properties. *Biochim Biophys Acta, Proteins Proteomics.* 2006 3;1764(3):452-60.
- [65] Kuroda Y. Biophysical studies of amorphous protein aggregation and in vivo immunogenicity. *Biophys Rev.* 2022 12;14(6):1495-501.
- [66] Friedreich N, Kekulé A. Zur Amyloidfrage. *Archiv für Pathologische Anatomie und Physiologie und für Klinische Medicin.* 1859 1;16(1-2):50-65.
- [67] Geddes AJ, Parker KD, Atkins EDT, Beighton E. "Cross- $\beta$ " conformation in proteins. *J Mol Biol.* 1968 3;32(2):343-58.
- [68] Eanes ED, Glenner GG. X-ray diffraction studies on amyloid filaments. *Journal of Histochemistry & Cytochemistry.* 1968 11;16(11):673-7.
- [69] Sunde M, Serpell LC, Bartlam M, Fraser PE, Pepys MB, Blake CC. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J Mol Biol.* 1997 10;273(3):729-39.
- [70] Linse S. Mechanism of amyloid protein aggregation and the role of inhibitors. *Pure Appl Chem.* 2019 2;91(2):211-29.
- [71] Ciarán Ó'Fágáin. Engineering Protein Stability. In: Walls D, Loughran ST, editors. *Protein Chromatography: Methods and Protocols.* Totowa, NJ: Humana Press; 2011. p. 103-36.
- [72] Arakawa T, Timasheff SN. Stabilization of protein structure by sugars. *Biochem.* 1982 12;21(25):6536-44.
- [73] Belton PS, Gil AM. IR and Raman spectroscopic studies of the interaction of trehalose with hen egg white lysozyme. *Biopolymers.* 1994 7;34(7):957-61.
- [74] Lin T, Timasheff SN. On the role of surface tension in the stabilization of globular proteins. *Protein Science.* 1996 2;5(2):372-81.

- 
- [75] Lins RD, Pereira CS, Hünenberger PH. Trehalose–protein interaction in aqueous solution. *Proteins: Structure, Function, and Bioinformatics*. 2004 4;55(1):177-86.
- [76] Olsson C, Jansson H, Swenson J. The Role of Trehalose for the Stabilization of Proteins. *J Phys Chem B*. 2016 5;120(20):4723-31.
- [77] Bauwens CM. X-Ray Scattering; 2011.
- [78] Fisher Z, Jackson A, Kovalevsky A, Oksanen E, Wacklin H. Biological Structures. In: *Exp. Methods Phys. Sci.*. vol. 49. Academic Press; 2017. p. 1-75.
- [79] D S Sivia. The basics of x-ray and neutron scattering. In: *Elementary scattering theory: for x-ray and neutron users*. Oxford University Press; 2011. p. 63-92.
- [80] Jülich F, Brückel T, Heger G, Richter D, Zorn R. Neutron Scattering: Lectures of the JCNS Laboratory Course Held at Forschungszentrum Jülich and the Research Reactor FRM II of TU Munich. Forschungszentrum Jülich; 2010.
- [81] Fitter J, Gutberlet T, Katsaras J. Neutron Scattering for Biology. In: *Neutron Scattering in Biology: Techniques and Applications*; 2006. p. 1-17.
- [82] Bée M. General Aspects of Neutron Scattering. In: *Quasielastic Neutron Scattering*. Bristol, England: Adam Hilger; 1988. p. 9-71.
- [83] Van Hove L. Correlations in Space and Time and Born Approximation Scattering in Systems of Interacting Particles. *Physical Review*. 1954 7;95(1):249-62.
- [84] Micoulaut M. Correlation functions and linear response theory;. Available from: [https://www.lehigh.edu/imi/teched/AtModel/Lecture\\_6\\_Micoulaut\\_Atomistics\\_Glass\\_Course.pdf](https://www.lehigh.edu/imi/teched/AtModel/Lecture_6_Micoulaut_Atomistics_Glass_Course.pdf).
- [85] Williams G, Watts DC. Non-symmetrical dielectric relaxation behaviour arising from a simple empirical decay function. *Trans Faraday Soc*. 1970;66:80.
- [86] Kikhney AG, Svergun DI. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Letters*. 2015 9;589(19PartA):2570-7.
- [87] Soper AK. Partial structure factors from disordered materials diffraction data: An approach using empirical potential structure refinement. *Phys Rev B*. 2005 9;72(10):104204.
- [88] McGreevy RL, Pusztai L. Reverse Monte Carlo Simulation: A New Technique for the Determination of Disordered Structures. *Molecular Simulation*. 1988 12;1(6):359-67.
- [89] Soper AK. Empirical Potential Structure Refinement A User's Guide; 2017.
- [90] Alder BJ, Wainwright TE. Phase Transition for a Hard Sphere System. *J Chem Phys*. 1957 11;27(5):1208-9.
- [91] Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. *J Comput Chem*. 2004 7;25(9):1157-74.
- [92] Lim JB, Rogaski B, Klauda JB. Update of the Cholesterol Force Field Parameters in CHARMM. *J Phys Chem B*. 2012 1;116(1):203-10.

- [93] Huang J, MacKerell AD. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J Comput Chem.* 2013 9;34(25):2135-45.
- [94] Lin FY, MacKerell AD. Improved Modeling of Halogenated Ligand–Protein Interactions Using the Drude Polarizable and CHARMM Additive Empirical Force Fields. *J Chem Inf Model.* 2019 1;59(1):215-28.
- [95] Bitar M, Offman M. *Molecular Dynamics*;
- [96] Gill P, Moghadam TT, Ranjbar B. Differential scanning calorimetry techniques: applications in biology and nanoscience. *J Biomol Tech.* 2010 12;21(4):167-93.
- [97] Haugstad G. Overview of AFM. In: *Atomic Force Microscopy: Understanding Basic Modes and Advanced Applications*. 1st ed. John Wiley & Sons; 2012. p. 1-33.
- [98] Bowron DT, Soper AK, Jones K, Ansell S, Birch S, Norris J, et al. NIMROD: The Near and InterMediate Range Order Diffractometer of the ISIS second target station. *Review of Scientific Instruments.* 2010 3;81(3).
- [99] Disordered Materials XRD;. Available from: <https://www.isis.stfc.ac.uk/Pages/XRD.aspx>.
- [100] Soper AK. GudrunN and GudrunX: programs for correcting raw neutron and X-ray diffraction data to differential scattering cross section; 2011. Available from: <http://epubs.stfc.ac.uk>.
- [101] Jmol: an open-source Java viewer for chemical structures in 3D;. Available from: <http://www.jmol.org/>.
- [102] Berendsen HJC, Grigera JR, Straatsma TP. The missing term in effective pair potentials. *J Phys Chem.* 1987 11;91(24):6269-71.
- [103] Kony D, Damm W, Stoll S, Van Gunsteren WF. An improved OPLS-AA force field for carbohydrates. *J Comput Chem.* 2002 11;23(15):1416-29.
- [104] Soper AK. The Structure of Water and Aqueous Systems. In: *Exp. Methods Phys. Sci.*. vol. 49. Academic Press; 2017. p. 135-211.
- [105] Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph.* 1996 2;14(1):33-8.
- [106] Carlile CJ, Adams MA. The design of the IRIS inelastic neutron spectrometer and improvements to its analysers. *Phys B.* 1992;182:431-40.
- [107] Bée M. Quasielastic Neutron Scattering for Continuous or Random Jump Diffusion of Molecules in Bounded Media. In: *Quasielastic Neutron Scattering*; 1988. p. 357-98.
- [108] Nečas D, Klapetek P. Gwyddion: An open-source software for SPM data analysis; 2012.
- [109] Youngs T. Dissolve: next generation software for the interrogation of total scattering data by empirical potential generation. *Mol Phys.* 2019 11;117(22):3464-77.