


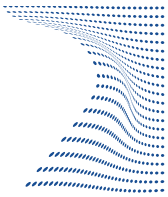

Doctoral Dissertation
Doctoral Program in Chemical Enginering (32.nd cycle)

Freeze Drying of Therapeutic Proteins: A Simulation Approach to Optimize Formulation and Process Conditions

Andrea Arsiccio
* * * * * *

Supervisors
Prof. Roberto Pisano, Supervisor
Prof. Antonello Barresi, Co-supervisor

Doctoral Examination Committee:
Prof. Alessandra Adrover, Università degli Studi di Roma La Sapienza
Prof. Giovanni Maria Pavan, Politecnico di Torino
Prof. Francesco Trotta, Università degli Studi di Torino
Prof. Gerhard Winter, Ludwig-Maximilians-Universität München
Prof. Mire Zloh, University College of London

Politecnico di Torino
2019

I hereby declare that, the contents and organisation of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

. . . . . .

Andrea Arsiccio

Turin, 2019

Summary

Protein-based pharmaceuticals are playing an increasingly important role in the treatment of a wide number of human diseases. This occurs because they are highly effective, and have fewer side effects compared to other therapeutics. For instance, the global sales of biopharmaceutical medicines were as high as US$ 237 thousand million in 2018, and are estimated to be valued at US$ 389 thousand million in 2024. However, a problem which is intrinsic to proteins is their high instability, and their tendency to lose therapeutic activity. The development of technologies capable of preserving their three-dimensional structure, and therefore biological potency, is therefore of utmost importance.

In this context, freeze-drying, or lyophilization, is a commonly used method for preparing solid protein-based pharmaceuticals. In spite of this, both the freezing and the drying steps may result in undesired stresses for the active ingredient, and therefore potential loss of activity. In the worst-case scenario, an immune response may even be generated, which may result in serious consequences for the patient. For this reason, a suitable formulation and a well-designed process should be selected to minimize denaturation phenomena.

However, the choice of the formulation is at present mainly based on experience, lacking real knowledge of the molecular-scale phenomena involved. Also the selection of optimal process conditions is often extremely time-consuming and non-systematic, resulting in poor control strategy. The present work aims to address these two problems, using an *in silico* approach.

After a brief introduction, where the state of the art and objectives of the work are discussed, the use of molecular dynamics to clarify the molecular mechanisms of protein stabilization and guide the choice of excipients will be investigated. Attention will be first focused on commonly used cryo- and lyoprotectants, with the aim to provide a better understanding of their stabilizing action, and identify molecular properties responsible for their effectiveness. The effect of buffers in common pharmaceutical formulations, and their role in modifying protein-excipients interaction, will also be addressed.

Ample coverage will then be given to the role of surfaces and surfactants, employing in this case also advanced molecular dynamics techniques, such as umbrella sampling and metadynamics. The mechanisms of interface-induced denaturation

of proteins will be studied, and a possible explanation for the role of surfactants in these phenomena will be proposed. In particular, an orientation-based mechanism of stabilization will be observed, where the protein is stabilized when surrounded by the hydrophilic heads of the surfactant. Attention will be particularly focused on the ice-water surface, because it plays a central role in the freezing step of the lyophilization process. It will be shown how the ice interface may perturb the secondary and tertiary structure of a protein by enhancing the mechanisms of cold denaturation.

Since the extent of the ice-water interface seems to be of utmost importance for the protein stability, some experimental work will also be presented, where the effect of the freezing protocol on protein stability is studied. Both conventional and controlled nucleation will be considered, for the case of human growth hormone and factor VIII as model proteins. It will be shown that the use of controlled nucleation is beneficial for cycle efficiency, product homogeneity and reconstitution time. However, the effect on protein stability strongly depends on the system being considered, with a potential role of the concentration of the formulation being frozen, as well.

Finally, a modelling approach based on energy and mass balance equations will be proposed to build the design space for the freezing process. The design space is a tool in the framework of the Quality by Design (QbD) concept that can be used to guide the selection of optimal operating conditions. The freezing step of freeze drying is considered in this Thesis because it determines product morphology, and therefore has an impact on cycle duration, thermal stress for the product being dried and protein activity. It will be shown that it is possible to predict the ice crystal size obtained as a result of freezing, as well as its effects on primary drying and protein stability. The proposed approach will be tested upon experimental data, and it will be demonstrated that it could help to improve both cycle efficiency and product quality. The critical parameters to be controlled for optimizing the freezing step of freeze-drying will be identified, and it will be shown that a trade-off should sometimes be achieved between quality and process performance.

Overall, the whole work has the aim to prove that modelling can help lyo-professionals in the selection of optimal freeze-drying conditions for protein-based therapeutics. It will be shown that simulations and experiments provide information at different levels. A model needs experimental data to be tuned and validated, but it can afterwards output information that cannot easily, or quickly, be accessed by current experimental techniques. Simulations and experiments are therefore not alternative, but complementary. The combination of these approaches may therefore be extremely beneficial, boosting up both process and formulation development.

Acknowledgements

I would like to start by thanking the people who helped me with the development of the research project presented in this thesis.

A special thank goes to my two advisors at Politecnico di Torino, Prof. Roberto Pisano and Prof. Antonello Barresi, who thoroughly supervised my work during these three years. Their professionalism, mentoring and helpfulness is sincerely acknowledged.

I would also like to thank the groups I collaborated with during my doctorate program. Some of the simulations presented in Chapters 3 and 4 were performed in collaboration with Prof. Joan-Emma Shea's group at the University of California, Santa Barbara (UCSB). Special thanks go therefore to Prof. Shea and her group, including Prof. James McCarty, Dr. Pritam Ganguly and Dr. Keila Cunha. Most of the knowledge I acquired about molecular dynamics is due to their precious teachings and availability.

Furthermore, some of the experimental work described in Chapter 5 was performed in Dr. Paul Matejtschuk's laboratories at the National Institute for Biological Standards and Control (Potters Bar, UK). I would therefore like to acknowledge Dr. Matejtschuk and his group, including Kiran Malik, Chinwe Duru and Ernest Ezeajughi, and all the people who helped with the experimental tests. For instance, Dr. Jackie Ferguson helped with the cell-based potency assay on human growth hormone, while Dr. Sanj Raut, Andrew Riches-Duit and Anwen Bullen contributed to the analyses of factor VIII stability in freeze dried plasma. Overall, they introduced me to the world of experiments with biologicals, and Dr. Matejtschuck and his group became my home away from home.

I would like to acknowledge support from the hpc@polito team (http://www.hpc.polito.it), the CINECA award under the ISCRA initiative (ProtExc-HP10C1MQDX and CysSurf-HP10CSOLZQ) and from the Center for Scientific Computing at the California Nanosystems Institute (CNSI), MRL: an NSF MRSEC (DMR-1720256) and NSF CNS-1725797, for high performance computing resources and support.

The master thesis students I supervised during these three years, and whose work is at least partially included in this manuscript, are also aknowledged: Alberto Romano, Andrea Paladini, Livio Marenco, Camilla Moino and Paolo Giorsello. I hope their thesis project was a pleasant experience, and I would like them to know

I was honored to contribute in their academic training.

All the friends who supported me during these years are gratefully acknowledged (in rigorous alphabetical order, in order to avoid choosing who should come first!): Alessio, Arianna, Benedetta, Beppe, Carlos, Daniele, Domenico, Elena, Fiora, Francesca, Francesco, Graziano, Lorenzo, Luca, Luigi, Luisa, Maite, Marco, Merve, Nunzio, Riccardo, Serena. They helped me with their advice, and brought happiness into this period of my life. Thanks for your friendship.

Last but not least, I would like to thank my whole family for their wholehearted support, especially in times of stress and anxiety. This journey would not have been possible without their continuous encouragement. Thanks for believing in me and being there whenever I needed you.

Contents

Chapter 1

Introduction

1.1 Freeze-Drying of Therapeutic Proteins: An Overview

The pharmaceutical industry is extremely active in research and innovation. In 2015, pharmaceutical companies around the world are estimated to have spent on R&D about US$ 150 thousand million [1], 5.5 times more than the aerospace and defense industries, 5 times more than the chemicals industry, and 1.8 times more than the software and computer services industry [2]. No other sector is so focused on R&D, as evident also from the fact that 5 of the 11 leading global R&D firms in 2014 were pharmaceutical companies [2, 3].

In addition to this, the number of research programs has increased steadily in the last few years. 56 new medicines were launched in 2015 [1], while in 2017 more than 7000 compounds were at different stages of development [3, 4]. These included 208 drugs for HIV/AIDS, 1919 for cancer, 401 for diabetes, 510 for mental disorders, 1308 for application in neurology and 563 for cardiovascular diseases [4]. However, a lot of research hurdles need to be overcome before a compound can be approved as a medicine.

According to some studies [5], the cost of developing a successful medicine today may even exceed US$ 2.6 thousand million, while it was about US$ 179 million in 1970s [3,6]. This sharp increase reflects the technical and regulatory challenges during the research and development phases before product approval. Pharmaceutical R&D is marked by high failure rates, and regulatory agencies require a stringent control of efficacy, quality, and safety, based on clinical trials.

These challenges have not reduced the industry's innovative energy, but have encouraged and promoted new strategies of development, focusing on collaboration with other companies and academia, and sharing of expertise. In recent years, a clear trend can be observed, where pharmaceutical companies are shifting their

production from small molecules and chemically synthesized drugs towards biopharmaceuticals. These include a variety of therapeutic products derived from living organisms, such as vaccines, recombinant proteins, and gene therapies.

The global biopharmaceutical market has grown considerably in the last few years, and is projected to exhibit a remarkable expansion in the near future. The global sales of biopharmaceutical medicines resulted in a total of US$ 228 thousand million in 2016 [7, 8], US$ 237 thousand million in 2018 and is estimated to be valued at US$ 389 thousand million in 2024 [9]. This remarkable growth can be attributed to the several benefits of biopharmaceuticals, including highly effective and potent action, few side effects, and the potential to cure previously untreatable diseases.

Monoclonal antibodies (mAbs) are expected to have the largest market share, but the segment of recombinant enzymes is believed to show rapid growth as well [9]. The use of biopharmaceuticals is still prevalent in developed countries, with North America currently being the dominant market. However, the Asiatic countries are expected to increase their market share in the future.

Moorkens *et al.* [8] analyzed the investment and development strategies of the top 25 pharmaceutical companies. They found that all these companies are actively engaged in the biopharmaceutical market, even though the industrial landscape is diverse. Pfizer, Novartis, Roche, Merck, Sanofi, Johnson & Johnson, AstraZeneca and GlaxoSmithKline invest in biotechnology both via their own development program, and via acquisition of biotechnological companies. Most of the biggest pharmaceutical companies are also involved in the development of next-generation biopharmaceuticals. For instance, Pfizer is working on a next-generation human growth hormone in collaboration with OPKO, Roche and Merck are developing Peginterferon alfa-2a and 2b respectively, Sanofi is doing research on modified structures of insulin etc.

The biggest pharmaceutical companies are also investing in emerging countries. As an example, Novartis is expanding its presence in Asia, Africa, and Latin America, Sanofi, Pfizer, and AstraZeneca are focused on China, Russia or both, and similar strategies are being carried on by all the biggest pharma groups. For the development of biopharmaceutical products, industries do not focus on a single strategy, but are generally involved in multiple investment programs. Most companies also look for collaborations with each other, and co-marketing. This is extremely important, because facilitates the sharing of expertise, know how and technologies required for the development of a biological medical product.

It is clear from these data that the market of biopharmaceuticals is steadily growing; however, not all the difficulties have been overcome, and some issues still need to be addressed. For instance, a problem that is intrinsically associated with biopharmaceuticals is their high instability. This is particularly true for therapeutic proteins, whose biological activity is strongly related to their 3D conformation. The native fold of a protein may be easily lost upon exposure to high/low temperature,

presence of interfaces, pH shifts etc., and any change in secondary or tertiary structure would result in diminished therapeutic potency. The formation of aggregates may even induce an undesired immune response in the patient. The development of technologies that may allow the preservation of biological products, and increase their stability, is therefore of utmost importance.

In this context, removal of water through drying is often used to improve stability, and for the long-term storage of labile molecules. It also provides additional benefits, including reduction in transportation and handling costs. Among drying techniques, freeze-drying or lyophilization is the most commonly used for the storage of protein therapeutics in the solid state, because it allows removal of water at low temperature. This represents a remarkable advantage when the active ingredient is sensitive to high temperature, as is the case of most of the proteins.

The freeze drying process consists of three steps, namely freezing, primary drying and secondary drying. Typically, the product is filled into vials, which are then loaded onto the shelves of the drying chamber. During freezing, the temperature of the shelves is then decreased linearly with time by means of a circulating fluid, until complete solidification of the formulation is achieved. Subsequently, the pressure inside the drier is lowered below the triple point of water, generally using a vacuum system comprising a condenser and a vacuum pump. The reduction in pressure marks the transition to the primary drying step, during which the frozen water is removed by sublimation. Finally, in the secondary drying phase the temperature is increased and the adsorbed water is removed by desorption, allowing further reduction in the residual moisture content.

Even though the freeze drying process is widely used for biopharmaceuticals, loss of the therapeutic activity may result from inappropiate choice of the formulation or of the process conditions. Starting from these premises, the present work aims to propose an approach, based on *in silico* modelling, capable of identifying those conditions that mostly preserve the biological activity of proteins to be freeze dried. In this first chapter, a state-of-the-art overview of the problems associated with the freeze-drying of proteins will be provided. First, the denaturation stresses that may arise during lyophilization will be described. The possible mechanisms of protein stabilization by excipients, and the current approach to the selection of formulation components, will be then introduced. Lastly, the effect of process variables on protein stability will also be briefly outlined. The objective is to define the context in which the present work is placed, and highlight the problems that this research project aims to address [1].

[1] Part of this chapter is based on a book chapter I have recently co-authored [10]

1.2 Denaturation Stresses During Lyophilization

Despite being a mild process, freeze drying can nevertheless lead to undesired disruption of the protein tertiary and secondary structure, or to the formation of aggregates [11–13]. This can result in a decreased therapeutic activity, and in some cases it may also lead to serious safety issues, such as generation of an immune response [12, 14].

The freeze-drying stresses are related to both the freezing and the drying steps. For instance, the interaction of nonpolar residues with water is less unfavorable at the low temperature used during freezing, and this may lead to cold denaturation of the protein [15–19]. The formation of ice may also be harmful to protein stability, for several reasons. First, proteins may adsorb to the interface, and undergo surface-induced denaturation [20]. Since a high cooling rate results in small ice crystals, with a large surface area, it has been often observed that protein unfolding is enhanced at higher cooling rates [21]. In addition to this, the crystallization of formulation components, such as glycine and mannitol, may also result in the formation of an additional interface, where surface-induced denaturation may occur [22, 23]. Another consequence of ice formation is the rapid increase in solutes concentration, which combined to selective crystallization of formulation components may lead to significant changes in ionic strength and relative composition of the amorphous phase, possibly destabilizing a protein [23]. As a result of the increased solute concentration, chemical reactions may actually accelerate in a partially frozen aqueous solution [24].

The effects for protein stability may be particularly significant when the component undergoing selective crystallization is a buffering species, as this could cause important pH changes. This may happen, for instance, with sodium phosphate and potassium phosphate [25–28]. The presence of polymers in the formulation may also result in problems of phase separation, due to polymers altered solubility at low temperature [29]. Phase separation leads to the formation of an additional interface where the protein may adsorb and denature, and also creates shear stresses that could be harmful to protein stability. Furthermore, the protein may preferentially partition into a phase with very low concentration of the stabilizer, reducing the protective effect.

Finally, the freeze drying process removes part of the hydration shell of the protein, which may disrupt the native state of the protein [30]. This is due to both a decreased charge density on the protein surface in a water-poor environment, which promotes aggregation, and the potential removal of water molecules that are integral part of active sites in the protein.

Because of the risks of protein denaturation associated to the freeze drying process, it is of utmost importance to design a formulation capable of minimizing the loss of therapeutic activity. In this context, we will describe the possible mechanisms of protein stabilization during both the freezing and drying steps of

lyophilization, and some examples of common pharmaceutical excipients will be provided.

1.3 Mechanisms of Protein Stabilization During Freeze-Drying

The most common excipients used for protein stabilization during freeze drying include:

- Sugars, such as the disaccharides sucrose, trehalose and lactose or the monosaccharide glucose;
- Polyols, such as sorbitol and glycerol;
- Polymers, including albumin, dextran, polyvinylpyrrolidone (PVP) or hydroxyethyl cellulose (HEC);
- Amino acids, such as glycine, proline, arginine etc.;
- Surfactants, especially the polysorbates Tween 20 and Tween 80.

When designing a formulation to be freeze dried, it is important to remember that freezing and drying expose proteins to different stresses; therefore the mechanisms of protein stabilization by excipients are not the same during the two stages of lyophilization. As a general guideline, those excipients that stabilize a protein in solution also have a protective action during freezing, as in both cases water is present. However, the mechanisms of protein stabilization are different in the dried state, and in this case the ability of the excipients to form a stiff, compact cake that inhibits the protein motions responsible for unfolding and aggregation becomes dominant [31]. In the following, the main mechanisms of cryo- and lyoprotection will be described. The role of surfactants will be discussed in a separate section, as their mechanism is significantly different.

1.3.1 Mechanisms of Cryoprotection

One of the most widely accepted mechanisms of protein stabilization during freezing is preferential exclusion. According to this theory, a protective osmolyte should be excluded from the protein surface, and preferential hydration should therefore ensue. The presence of these co-solutes creates a thermodynamically unfavorable situation, which is augmented by an increase in the surface area exposed by the protein [32]. As a consequence, the native fold is stabilized, because denaturation would result in a greater contact surface area between the protein and the solvent. Moreover, stabilization of the native structure is greater at high excipient

concentrations because preferential exclusion increases with the concentration of the osmolyte [33].

The preferential exclusion of an excipient from the protein surface can be quantified using the Kirkwood-Buff integrals G_{ij} [34, 35],

$$G_{ij} = \int_{r=0}^{\infty} [g_{ij}(r) - 1] 4\pi r^2 dr \tag{1.1}$$

where $g_{ij}(r)$ is the radial distribution function for component i with respect to component j at a given distance r between i and j. Using the Kirkwood-Buff integrals it is, in fact, possible to compute the preferential exclusion coefficient Γ_{23} between the protein (2) and the excipient (3), as

$$\Gamma_{23} = \rho_3 (G_{23} - G_{21}) \tag{1.2}$$

where G_{21} and G_{23} are the protein-water (1) and protein-excipient Kirkwood-Buff integrals, respectively, while ρ_3 is the number density of the excipient in the solution. A positive value of the preferential exclusion coefficient indicates preferential interaction between the protein and the excipient, while a negative Γ_{23} implies preferential exclusion.

Preferentially excluded excipients also hinder dissociation of proteins with a quaternary structure, again because the dissociation process would increase the surface area and the thermodynamic instability. However, those osmolytes that are excluded from contact with the protein may also enhance aggregation phenomena, because the self-association of proteins reduces the total surface area exposed. This undesired intensification of self-association is anyway prevented if protein unfolding is the key determinant in causing aggregation; in this case the excluded co-solute should be able to reduce aggregation by stabilizing the native structure [31].

In addition to this thermodynamic mechanism, many common cryoprotectants, such as polymers and sugars, also increase the viscosity of the solution, especially as a result of cryoconcentration. The high viscosity that is eventually reached dramatically restricts diffusion processes, and should consequently decrease the protein unfolding rate. Furthermore, some cryoprotectants may stabilize proteins by suppressing pH changes during freezing [26].

Some other hypotheses have been proposed more recently. For instance, it was found that the favorable interaction, rather than preferential exclusion, between lactate dehydrogenase (LDH) and high molecular weight polyethylene glycols (PEGs) (e.g., PEG 4000 and PEG 8000) cryoprotected LDH [36]. However, preferential exclusion remains by far the most widely used theory to explain protein stabilization both in solution and during freezing.

1.3.2 Mechanisms of Lyoprotection

During the drying steps of freeze drying, the preferential exclusion mechanism is no longer applicable because at this point there is essentially no water [37]. Therefore, different theories have been developed to explain protein stability in the dried state. For instance, according to the water replacement mechanism [38–40], the protective osmolytes should hydrogen bond with the protein at the end of the drying process to satisfy the hydrogen bonding requirement of the polar residues. In fact, a protein may form non-native intermolecular hydrogen bonds upon dehydration, resulting in unfolding. Formation of these intermolecular hydrogen bonds is inhibited if the excipients serve as water substitutes. However, since the highest number of hydrogen bonds can be formed if the stabilizer is in the amorphous phase, crystallization of the excipient may reduce protein stabilization [37].

One other major mechanism of protein stabilization by lyoprotectants is the formation of a viscous glassy matrix during lyophilization [41, 42], that increases protein stability by slowing down protein denaturation and unfolding [43]. Thus, in this case stabilization occurs via a pure kinetic mechanism, and a requirement is that the protein and stabilizer are in the same amorphous phase.

Finally, according to the water entrapment hypothesis, excipients may form a cage around the protein, which entraps and slows down water molecules [44–46]. This makes it possible both to maintain a high level of hydration, and to hinder the protein motions that may lead to denaturation. A scheme of the mechanisms of preferential exclusion and water entrapment is shown in Figure 1.1.

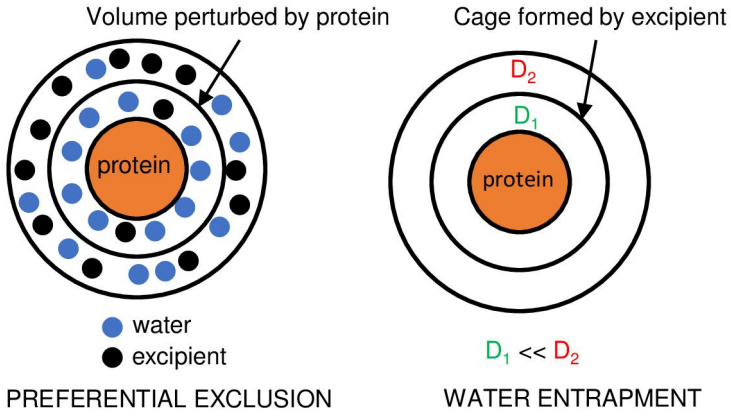

Figure 1.1: Graphic representation of preferential exclusion and water entrapment. D represents the diffusion coefficient of water molecules.

1.3.3 Role of Surfactants

As previously discussed, a major cause of biopharmaceuticals instability is related to phenomena of surface-induced denaturation and aggregation. Therefore, surfactants are often added to the formulation, in order to prevent surface-driven damage. The polysorbates Tween 20 and Tween 80, which are nonionic surfactants consisting of fatty acid esters of polyethylene sorbitan, are commonly used for protein stabilization.

Different mechanisms have been proposed to explain the protective action of surfactants [47, 48]. Because of their amphiphilic nature, surfactants preferentially locate at interfaces, thus reducing adsorption and/or aggregation phenomena. This effect is concentration-dependent, and correlates with the critical micelle concentration (CMC) of the surfactant. It was also suggested that surfactants may assist in protein refolding, acting like molecular chaperones, and should increase the free energy of protein unfolding [49]. Finally, it is also possible that surfactants associate with proteins in solution, preventing self-association. This last effect should not be related to the CMC of the surfactant, but to the molar binding stoichiometry between the surfactant and the protein [50–52].

However, numerous unresolved questions remain, together with some contradictory observations. For instance, Deechongkit *et al.* [52] observed that polysorbates affected the α-helix content of darbepoetin alpha, even below the CMC. In contrast, previous studies of recombinant human growth hormone (rhGH) [49] and anti-L-selectin antibody [53] by far-UV circular dichroism seemed to demonstrate a negligible effect of surfactant molecules on the proteins secondary structure.

1.4 Current Approach to Formulation Design

The challenge for a formulation scientist is to determine the optimal choice of excipients, and the best concentration of each excipient. Currently, the design and optimization of pharmaceutical formulations are mainly based on some practical guidelines, mostly derived from experience [54], along with the Design of Experiments (DoE) approach [55, 56].

Some practical guidelines can be useful for researchers who do not have extensive experience in formulation development, especially if they point out the main factors to consider when designing a freeze dried formulation, i.e., not only protein stability, but also glass transition and collapse temperature, tonicity, route of administration, container etc. However, when the formulation is critical, it is important to use a more effective methodology. In this context, DoE and statistical analysis are commonly applied to formulation development, because allow a quick and systematic evaluation of all the potential factors involved. Using DoE, it is possible to evaluate the effect of each formulation factor on each monitored output, as well as potential interactions. Afterwards, the critical factors could be identified

and adjusted, in order to optimize the formulation and guarantee its robustness.

However, a problem with current approaches for formulation design is that the molecular mechanisms at the basis of protein stabilization by excipients are generally not accessible by experimental techniques. This means that the formulation design is still mostly driven by experience and empirical observations, without real knowledge of the molecular-scale phenomena involved. Moreover, the number of factors to be considered is often huge, making the process extremely time-consuming. In this framework, this research project aimed to identify a possible solution to this problem, i.e., the combination of molecular-scale *in silico* modeling with experiments.

1.5 Effect of Process Variables on Protein Stability

Several process variables may affect protein stability during the freeze-drying process. Amon them, the cooling rate is a critical parameter that needs to be controlled during freezing. Generally, a higher cooling rate results in smaller ice crystals, and therefore larger ice-water interface [57,58]. The ice crystals size is also dramatically influenced by the nucleation temperature, with a higher nucleation temperature resulting in larger crystals [59, 60]. As previously mentioned, a major contribution to loss of therapeutic activity is related to adsorption at the ice-freeze concentrate interface [61]. Hence, it is clear that the ice crystals size is a critical parameter when protein stability is concerned.

However, while the cooling rate can easily be controlled in common freeze-drying equipment, the nucleation temperature is a stochastic variable, randomly distributed within the batch. This implies that large vial-to-vial variability is often observed, and that strict control of product morphology is not possible. To overcome this limitation, a number of techniques has been developed to allow the control of the nucleation temperature [58, 62–64]. The control of the nucleation temperature may also be applied to reduce the reconstitution time of highly concentrated lyophilized protein products [65], and it was reported that it reduced the degradation of lactate dehydrogenase (LDH) during freezing, even though did not promote significantly superior stability during the entire freeze-drying process [66]. Moreover, controlled nucleation applied to a highly concentrated monoclonal antibody solution helped to suppress glass fogging, i.e., the undesired migration of protein solutions up on the inner walls of glass vials during the freezing step of lyophilization [67]. However, the impact of controlled nucleation on protein stability during freeze-drying needs to be further investigated, and some results on this aspect will be discussed in this work.

The cooling rate may also influence other aspects related to protein behavior during lyophilization. For instance, it was already mentioned that buffer species may crystallize out selectively and cause pH shifting during freezing. If the protein

is pH-sensitive it is therefore preferable to keep the buffer species amorphous during freezing. A high cooling rate may prevent crystallization [68, 69]. Moreover, the cooling rate also affects the extent of crystallization, or the formation of different crystalline forms of an excipient, such as mannitol [70, 71] or glycine [72]. In turn, this may influence protein stability.

The presence of an additional annealing step during freezing, often introduced to promote complete crystallization of crystalline components, may also have an adverse effect on the biological activity of the therapeutic ingredient. During annealing, the frozen product is heated above the glass transition temperature of the formulation to promote recrystallization, and the removal of a component from the amorphous phase may loosen the hydrogen bonded matrix that preserves the protein structure, or form an additional interface onto which the protein could adsorb and unfold [73–75]. However, large ice crystals are formed during annealing [76], and this contributes to minimize the risk of protein denaturation at the ice-water interface during freezing, or at the air-water surface during drying.

During drying, the maximum temperature reached within the product is a critical parameter, as a too high temperature may result in product collapse. A collapsed cake is characterized by a lower rate of water removal by both sublimation and desorption, and generally contains a higher residual moisture compared to a well-formed product, which could speed up protein unfolding and aggregation. However, different opinions can be found in the literature in this regard, with some studies suggesting that collapse does not impact protein stability [77–79]. The maximum temperature reached by a product being dried is strongly influenced by its morphology, and, in particular, by the pore size of the dried layer. As the pore dimension corresponds to the ice crystals size of the frozen product, the cooling rate and nucleation temperature also influence the temperature behavior of the product during drying.

The residual moisture level in a lyophilized product is another key factor to be considered when protein stability is concerned. Usually, a lower moisture content results in improved stability, and, as a general rule, the moisture level in a freeze dried protein formulation should be below 2% [80]. However, a drier cake is not always more stable, as overdrying may occur. For instance, lyophilized BSA and bovine γ-globulin formulations were more stable at 10 % water content than at $<$ 1% [81].

Finally, it is important to remember that a freeze drying cycle should also be efficient and economically viable. In some cases, a trade-off should be achieved between efficiency and product quality, always keeping in mind the critical requirements of the final product. For instance, cooling rate and nucleation temperature during freezing also have an influence on the drying rate. A low cooling rate and high nucleation temperature speed up sublimation, but slow down the desorption process during secondary drying [59, 82–84]. Therefore, a balance should be achieved to maximize both process efficiency and product quality.

1.6 Current Approach to the Selection of Process Conditions

The design of a freeze drying process involves the selection of a number of variables that are strongly interrelated, and whose effect on process performance and product quality is often extremely complex. Because of this, regulatory agencies advice that greater attention should be given on process control. For example, in the past few years, the U.S. Food and Drug Administration (FDA) implemented the concepts of Quality by Design (QbD) into its processes, as outlined in its report "Pharmaceutical cGMPs for the 21st Century: A Risk-Based Approach" [85].

QbD is a systematic approach that emphasizes process understanding and control in order to guarantee the desired characteristics of the final product [86, 87]. According to the QbD concept, the focus is on control strategy and robustness, with the objective of improving manufacturing efficiency and product quality.

In spite of the importance of the subject, the current approach to the selection of freeze drying conditions remains, in many cases, empirical and nonsystematic, resulting in poor control strategy. In the following chapters of this thesis, we will show how the implementation of a QbD approach could allow an easier control of the freezing step of freeze drying, and a faster selection of the most appropriate freezing conditions.

1.7 A Model-Based Approach to Improve Formulation and Process Design

It is clear from the previous sections that several complex considerations are hidden behind the selection of an optimal formulation, or of the most suitable process conditions, for the freeze drying of biopharmaceuticals. In this framework, the objective of this research project was to identify and validate a model-based approach capable of helping the freeze drying community with this difficul task.

In this context, the use of molecular dynamics could be a valuable tool for formulation development. Molecular simulations [88, 89] are playing an increasing role in a wide number of fields, spanning physics, chemistry, biology, and life science. The key for their success is their ability of providing insight into problems which are often not directly accessible by experimental techniques, thus offering a better understanding of complex phenomena. In a common molecular dynamics (MD) simulation, structure files describing the protein (in many cases a Protein Data Bank file) or the excipient molecule, and topology files listing the forces acting on each atom, are needed to set up a simulation. A simulation box is then built, where the protein is surrounded by the desired number of excipient molecules and solvated in water, and counter ions are also added to guarantee the neutrality of the system.

A typical box is generally very small (side length on the order of ≈ 10 nm); due to this small size, edge effects (i.e., the notrivial percentage of molecules being on the outer faces of the box) become not negligible and periodic boundary conditions need to be used. This means that the simulated system is artificially surrounded by replicas of itself, and if an atom leaves the box from one side, it immediately enters the same box from the opposite one. After an energy minimization step, the system is equilibrated at the desired values of temperature and pressure. Finally, the production run can start, where Newton's law of motion is integrated over time for all the atoms in the simulation box, and the coordinates are written to an output file at regular intervals. A scheme of the process required for starting an MD simulation is shown in Figure 1.2.

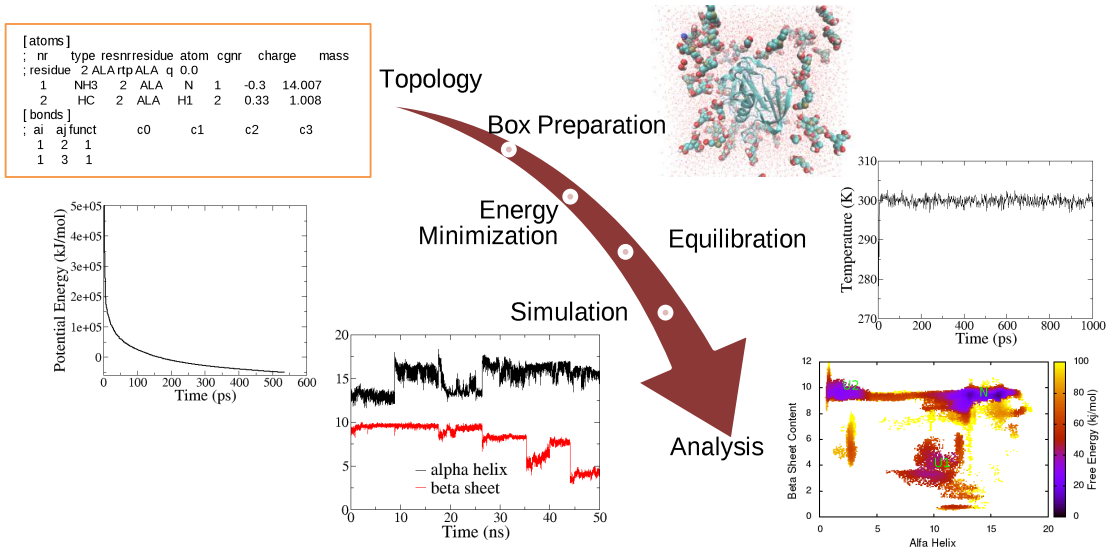

Figure 1.2: Scheme of the process required for starting a molecular dynamics simulation.

The forces acting on each atom are defined by the force field, which describes both the non-bonded and bonded interactions in the system under investigation. The bonded interactions can be easily and quickly computed because they always involve the same atoms. For instance, a covalent bond between two atoms i and j, with positions defined by $\mathbf{r}_i$ and $\mathbf{r}_j$, is generally described according to the harmonic potential V_{ij},

$$V_{ij}(r_{ij}) = \frac{1}{2} k_{ij}(r_{ij} - b)^2, \; r_{ij} = |\mathbf{r}_i - \mathbf{r}_j| \tag{1.3}$$

where k_{ij} is the force constant and b the equilibrium distance between the two atoms. The forces acting on angles θ_{ijk} and dihedrals ϕ_{ijkl} formed within each

molecule inside the simulation box are often defined according to the following potentials,

$$V_{ijk}(\theta_{ijk}) = \frac{1}{2}k_{ijk}(\theta_{ijk} - \theta_0)^2,\ \theta_{ijk} = \cos^{-1}\left(\frac{\mathbf{r}_{ij}\cdot\mathbf{r}_{jk}}{r_{ij}r_{jk}}\right) \tag{1.4}$$

$$V_{ijkl}(\phi_{ijkl}) = k_{ijkl}(1+\cos(\phi_{ijkl}-\phi_0)),\ \phi_{ijkl} = \cos^{-1}\left(\frac{(\mathbf{r}_{ij}\times\mathbf{r}_{jk})\cdot(\mathbf{r}_{jk}\times\mathbf{r}_{kl})}{|\mathbf{r}_{ij}\times\mathbf{r}_{jk}||\mathbf{r}_{jk}\times\mathbf{r}_{kl}|}\right) \tag{1.5}$$

where k_{ijk} and k_{ijkl} are the force constants, while θ_0 and ϕ_0 represent the equilibrium values. In some cases, the dihedral angles potential is expressed as a sum of powers of cosines, according to the Ryckaert-Bellemans function,

$$V_{ijkl}(\psi_{ijkl}) = \sum_{n=0}^{5} C_n(\cos\psi_{ijkl})^n,\ \psi_{ijkl} = \phi_{ijkl} - 180^\circ \tag{1.6}$$

The non-bonded interactions are more difficult to compute, as the atoms involved change during the simulation time. This happens because the molecules are not fixed in space, but are free to move within the simulation box. The non-bonded interactions include a repulsion term, a dispersion term, and a Coulomb term. The repulsion and dispersion terms are often combined in the Lennard-Jones (or 6-12) interaction,

$$V_{ij}(r_{ij}) = \frac{C_{12}}{r_{ij}^{12}} - \frac{C_6}{r_{ij}^6} = 4\varepsilon\left(\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^6\right) \tag{1.7}$$

where the parameters C_{12} and C_6 or σ and ε depend on pairs of atom types. The Buckingham (or exp-6) potential may also be employed,

$$V_{ij}(r_{ij}) = A_{ij}e^{-B_{ij}r_{ij}} - \frac{C_{ij}}{r_{ij}^6} \tag{1.8}$$

In addition, atoms having charges q_i and q_j also act through the Coulomb term,

$$V_{ij}(r_{ij}) = \frac{1}{4\pi\varepsilon_0}\frac{q_iq_j}{r_{ij}} \tag{1.9}$$

where ε_0 is the absolute permittivity of free space.

Some of the most common all-atom and united-atom force fields are AMBER [90], CHARMM [91], GROMOS [92] and OPLS [93]. All-atom force fields provide parameters for every single atom within the system, while united-atom ones provide parameters for all atoms except non-polar hydrogens. The resulting description of the system is extremely detailed, resulting in the generation of a huge amount of information about the system, but also in a high computational cost. In fact, the

main problem of classical atomistic MD is represented by the limited time and length scales that can be simulated.

We will show that molecular dynamics is a valuable tool for studying the interactions between a model protein and typical pharmaceutical excipients. For this type of simulations, the time scales which can be reached by classical MD simulations may be sufficient. A typical simulation box, in this case, would include a protein, solvated with water and surrounded by a given number of excipient molecules (see Figure 1.3).

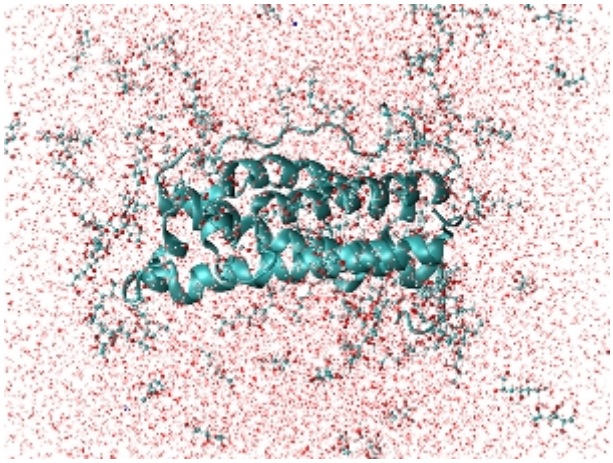

Figure 1.3: Example of an MD simulation box, where native human growth hormone is surrounded by sucrose molecules and solvated in water.

The MD trajectory could be then analyzed to identify, and quantify, the mechanisms of protein stabilization by different osmolytes. For instance, molecular level simulations allow evaluation of preferential interaction of proteins, in terms of either Kirkwood-Buff integrals or preferential exclusion coefficient Γ_{23}, with either water or cosolvents [94, 95]. The number and strength of the hydrogen bonds formed within the simulation box, which is related to the formation of hydrogen bonded clusters or to the degree of kinetic stabilization of the protein, can also be easily computed.

In some cases, however, the effect of an excipient on protein conformational changes may be the main interest of the *in silico* modeling, and these conformational changes may occur on timescales which cannot be easily reached by classical MD simulations. A possible approach to reduce the computational cost is coarse-graining. In this case, the number of degrees of freedom in the model is reduced, allowing simulations of larger systems for longer times [96]. To achieve this, typically whole groups of atoms are represented as single beads, and the coarse-grained force field describes their interaction. However, approximations must be made, and the resulting description of the system under investigation is often less accurate.

In addition to this, the approximations which result from coarse-graining may be adequate only for a limited selection of molecules. Therefore, this type of approach may be difficult to implement for the simulation of typical formulations, where the chemical nature of the excipients used is often very broad.

Another option to overcome the timescale limitation of computer simulations consists in the use of enhanced sampling methods [97]. Enhanced sampling techniques help speed up the transitions between different protein structures, and are extremely useful in the case of energy landscapes that feature many minima, and barriers between these minima that can be difficult to cross.

Some enhanced sampling approaches rely on the identification of few order parameters, or collective variables (CVs), which are critical for the system of interest and whose fluctuations are therefore enhanced by a bias during the simulation time. Among this class of techniques, metadynamics is frequently used [98]. Different versions of metadynamics exist, but in all cases an external history-dependent bias potential V is constructed in the space of the selected collective variables $s = (s_1, ..., s_d)$, which are function of the atomic coordinates $s_i(\mathbf{R})$, in order to push the system away from local minima and promote the sampling of new regions of the phase space. The potential is built as a sum of Gaussian kernels, having height ω and width σ_i, deposited along the trajectory,

$$V(s,t) = \int_0^t \omega \exp\left(-\sum_{i=1}^{d} \frac{(s_i(\mathbf{R}) - s_i(\mathbf{R}(t')))^2}{2\sigma_i^2} \right) dt' \tag{1.10}$$

A second class of techniques consists in simulating multiple replicas of the same system at different temperature, Hamiltonian or pressure, and then randomly exchanging the configurations with a given acceptance probability. Among the different possible approaches, the so-called parallel tempering replica exchange molecular dynamics (REMD) is often used [99,100]. REMD simulations enhance the sampling by running several independent replicas of the system of interest at different temperatures, and periodically exchanging the coordinates between the replicas with the following acceptance probability,

$$\alpha = \min\left(1, e^{\left(\frac{1}{k_B T_j} - \frac{1}{k_B T_i}\right)[U_j - U_i]}\right) \tag{1.11}$$

where we have imagined an exchange between two replicas at temperatures T_i and T_j, and having energies U_i and U_j. Typically, the set of replicas is constructed so that the replica at the lowest temperature represents the ensemble from which sampling is wanted, while the highest temperature is chosen so that the barriers in the energy landscape will be crossed over accessible simulation time scales.

In the next chapters we will show how MD simulations could be used to provide useful information about the mechanisms of protein stabilization by excipients, and therefore guide the choice of the formulation. More specifically, classical MD,

umbrella sampling (US) and parallel bias metadynamics (PBMetaD) simulations will be applied to different systems, including several class of typical pharmaceutical excipients and model proteins. Both US and PBMetaD belong to the class of enhanced sampling techniques relying on the identification of collective variables. In this thesis, the CVs will be related to the secondary and tertiary structure of proteins (for instance radius of gyration R_g, root mean square deviation from the backbone RMSD, dihedral angles, β-sheet or α-helix content) and to the effect of external cosolutes/surfaces on the native fold (distance from model surfaces, coordination number of cosolutes).

We will also demonstrate how the design of the freezing step of freeze drying could benefit from the QbD concept. In particular, a mechanistic model for the prediction of the ice crystals size distribution within a frozen product will be integrated into a design space approach. Design space is a tool that may be used in the QbD process to provide useful information about the effect of input variables on output critical parameters. The proposed approach aims to guide the selection of freezing conditions, that allow optimal preservation of proteins biological activity, while ensuring process economic efficiency at the same time.

In the concept herein proposed, *in silico* modelling is not alternative to experiments, but, rather, the two approaches should be seen as complementary. For instance, the knowledge of molecular-level phenomena provided by MD, which cannot be obtained experimentally, could be used to design a formulation to be lyophilized using a rational, knowledge-driven approach, in opposition to the present experience-based one. At the same time, the use of mathematical modelling to build a design space for the selection of optimal process conditions would allow a significant saving of time, dramatically decreasing the number of experiments to be performed.

However, due to the difficulty of reproducing in a simulation the complexity of reality, and the approximations inherent in any mathematical model of a real process, experimental tests remain an essential tool, which should be used to validate the simulation results. The computational investigation can therefore integrate the description offered by experimental data, making the whole process more time- and cost-effective.

1.8 Overview of the Work

In the following chapters, we will explain how computational modelling could be used to optimize formulation and process conditions for the freeze drying of biopharmaceuticals. More specifically, the following questions will be addressed,

1. What are the molecular properties related to cryo- and lyoprotection? Is the protection mechanism sequence-specific?

2. What is the role of buffers? Do they have any other property, other than pH control, which may affect protein stability?

3. How do surfactants prevent protein aggregation and surface adsorption? Do they also have an effect on the secondary structure of proteins? If so, what is the mechanism behind their action?

4. Which of the interfaces typically encountered during a freeze drying process is more dangerous for protein denaturation, and how do they induce conformational changes?

5. What is the role of the ice-water interface? How does it affect protein stability?

6. May a controlled nucleation approach, resulting in the formation of a small ice-water surface area, be beneficial for protein stability?

7. What freezing conditions result in the best preservation of protein residual activity? Are these conditions beneficial for process performance, as well?

A possible answer to these questions will be discussed in the present thesis, which will be divided into 6 Chapters.

In Chapter 2 we will show how MD simulations could be used to clarify the mechanisms of protein stabilization by typical pharmaceutical excipients. Both cryo- and lyoprotectants will be considered, using human growth hormone (hGH) and lactate dehydrogenase (LDH) as model proteins. Simulation results will be presented, suggesting that not all the excipients are equally effective during freezing and drying, and some molecular properties will be identified that correlate with the stabilizing action of the osmolytes. The interaction of buffer species with other components of a typical formulation for biopharmaceuticals will also be investigated, and the effects of buffers on protein preservation, other than pH control, will be discussed.

In Chapter 3 we will then shift our attention to the role of surfactants. Their effect on protein behavior will be studied both in bulk solution, and at the air-water, silica-water or ice-water interfaces. hGH and the GB1 peptide will be used as model proteins, while the polysorbates Tween 20 and Tween 80 will be the model surfactants. Simulation results will show evidence of an orientation-dependent mechanism, and will explain surfactants behavior as an effect of their amphiphilic nature.

In Chapter 4 attention will be focused more closely on the ice-water interface. The effect of the ice surface on protein stability will be studied using MD simulations, and an explanation based on a possible enhancement of cold denaturation phenomena by the ice interface will be proposed. Another possible interpretation of the preferential exclusion theory will also be introduced, suggesting that the

exclusion of excipients from specific patches on the protein surface is crucial for biopreservation.

In Chapter 5 the effect of the freezing protocol on protein stability, and in particular the benefits of the controlled nucleation approach known as vacuum induced surface freezing (VISF), will also be investigated. The human growth hormone and factor VIII (FVIII) will be selected as model proteins, and it will be shown that the benefits of controlled nucleation depend on the formulation being freeze dried.

Finally, in Chapter 6, a mechanistic model, based on energy and mass balance equations, will be presented, that allows prediction of the ice crystals size within a frozen product. Using this model, the design space for the freezing step of freeze drying will be built. It will be shown that the design space can guide the selection of freezing conditions, from the point of view of both process performance and protein stability.

Some final conclusions and discussion will eventually be presented in Chapter 7.

List of Symbols

A_{ij}	J	coefficient of Buckingham potential (Equation 1.8)
b	m	equilibrium bond length
B_{ij}	m^{-1}	coefficient of Buckingham potential (Equation 1.8)
C_n	-	coefficients of Ryckaert-Bellemans function (Equation 1.6)
C_6	J m^6	coefficient of Lennard-Jones potential (Equation 1.7)
C_{12}	J m^{12}	coefficient of Lennard-Jones potential (Equation 1.7)
C_{ij}	J m^6	coefficient of Buckingham potential (Equation 1.8)
d	-	number of collective variables being biased
D	m^2 s^{-1}	diffusion coefficient
g_{ij}	-	radial distribution function
G_{ij}	-	Kirkwood-Buff integrals
k	N m^{-1}	harmonic force constant
k_B	J K^{-1}	Boltzmann constant
q	C	atomic charge
$\mathbf{r}$	m	position vector
r	m	distance

$\mathbf{R}$	m	atomic coordinates
R_g	m	radius of gyration
s	-	collective variable
t	s	time
T	K	temperature
U	J	energy
V	J	potential

Greek Letters

Γ	-	preferential exclusion coefficient
ε	J	parameter of Lennard-Jones potential (Equation 1.7)
ε_0	F m^{-1}	absolute permittivity of vacuum
θ	-	angle
θ_0	-	equilibrium angle
ρ	m^{-3}	number density
σ	m	parameter of Lennard-Jones potential (Equation 1.7)
σ_i	-	Gaussian width
ϕ	-	dihedral angle
ϕ_0	-	equilibrium dihedral angle
ψ	-	dihedral angle (Equation 1.6)
ω	J	Gaussian height

Subscripts

1	water
2	protein
3	excipient

Chapter 2

Clarifying the Mechanisms of Cryo- and Lyo-Protection for the Biopreservation of Proteins

As discussed in the Introduction, biopharmaceuticals are playing a central role in the treatment of a wide number of human diseases. Having this in mind, it is clear that the stresses that may lead to loss of therapeutic activity during the production process should be minimized. This result is generally achieved through the addition of appropriate excipients.

A typical formulation for a protein-based therapeutics to be lyophilized includes both cryo- and lyoprotectants, together with a buffer to control the pH, a bulking agent to provide cake resistance, and, in some cases, a surfactant for the minimization of surface-induced denaturation. In this first chapter, we will focus our attention on the role of cryo- and lyoprotectants, with the aim to unveil the molecular mechanisms at the basis of protein stabilization.

Among the cryo- and lyoprotectants, sugars, polyols and amino acids are commonly used. As discussed in the Introduction, a number of theories has been proposed in the literature to explain their protective action, that encompass both thermodynamic and kinetic mechanisms of stabilization. Several experimental studies have also been carried out to investigate the effect of different formulations on protein activity. However, the molecular mechanisms at the basis of protein stabilization by excipients are generally hardly accessible by current experimental techniques. This means that, at present, the choice of the formulation is mainly empirical and based on experience, without deep knowledge of the molecular-scale phenomena involved.

In the present chapter, a possible approach for the design of the formulation for biopharmaceuticals is discussed. More specifically, we propose Molecular Dynamics (MD) as an efficient tool for the pre-screening of a wide number of possible

excipients. MD simulations can uncover the molecular mechanisms of protein stabilization by the surrounding osmolytes, thus integrating the picture provided by experimental data. Moreover, the preliminary *in silico* investigation would make it possible to decrease the number of formulations to be experimentally studied, making the whole process more time- and cost- effective.

In the concept herein outlined, MD simulations and experiments are envisioned as complementary approaches. The knowledge of molecular-level phenomena provided by MD, which cannot be obtained experimentally, could in fact be used to design a formulation to be lyophilized using a rational, knowledge-driven approach, in opposition to the present experience-based one. However, due to the limited time and length scales accessible by MD, and the difficulty of reproducing in a simulation box the complexity of reality, experimental tests remain an essential tool, which should be used to validate the simulation results.

The proposed approach will be here introduced using two model proteins, namely, human growth hormone (hGH) and lactate dehydrogenase (LDH). Human growth hormone is an aggregation-prone protein widely studied in literature [101–108]. It is a 191-amino acids protein, having 4 α-helices, that stimulates growth, cell reproduction and regeneration (see Figure 2.1a). LDH is a well-known tetrameric enzyme (see Figure 2.1b), sensitive to stresses encountered during the freeze-drying process [23, 26, 61, 66, 109–114].

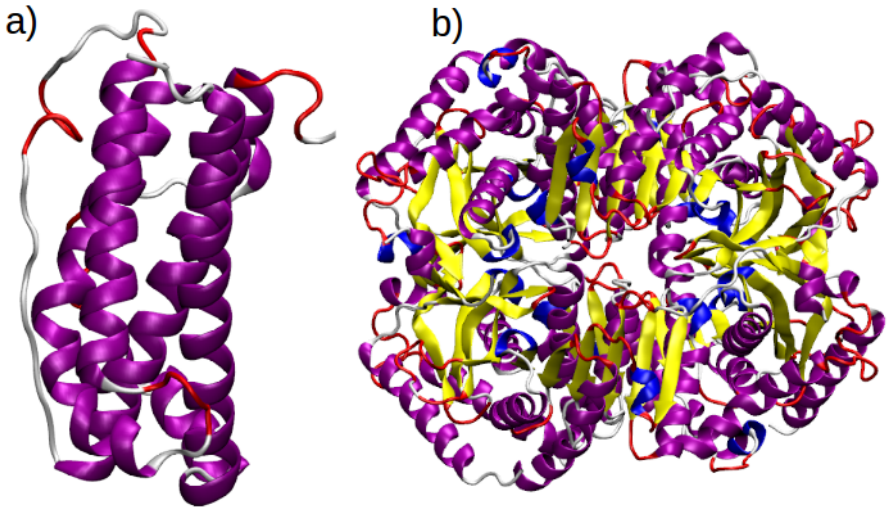

Figure 2.1: Cartoon representation of a) hGH and b) LDH. The various colors identify different types of secondary structure: α-helix in purple, 3_{10} helix in blue, extended beta in yellow, turn in red and coil in white

The mechanisms of hGH and LDH stabilization by excipients during freezing, or

in the dried state, will be first investigated using Molecular Dynamics, with reference to the previously discussed theories of preferential exclusion, vitrification, water replacement and water entrapment. Simulation results will show that not all the excipients are equally effective during freezing and drying. Moreover, a new theory will be proposed, where the water replacement, vitrification, and water entrapment hypotheses, so far considered as different, alternative mechanisms to explain lyoprotection, will be envisioned as complementary aspects of the same phenomenon. This phenomenon will be identified as the formation of a strong, compact hydrogen-bonded matrix. The results obtained for hGH and LDH will be shown to be in line with each other, indicating that the analysis performed should not be protein-specific, but should rather be generally applicable. Experimental data will be used to confirm the MD simulations, and a simple phenomenological model will be then presented to explain, using molecular properties, the behavior of different excipients with respect to protein stabilization. An example of the proposed approach for formulation design, where the *in silico* results are combined with an experimental investigation, will also be provided for LDH as model protein.

Finally, the effect of buffers on the previously discussed mechanisms of protein stabilization by excipients will be investigated. For this study, hGH will be used as model protein, and formulations including a mixture of sugars (sucrose or trehalose) and buffer species (phosphate and citrate buffers) will be studied. It will be shown that buffers may break the hydrogen bonding network formed by excipients, therefore affecting the mechanisms of protein stabilization in sucrose-based or trehalose-based formulations. This suggests that buffers have important properties, other than the control of pH, that could contribute to the overall stability of proteins. Some of these properties are related to their interaction with the other components of the formulation[1].

2.1 Mechanisms of Protein Stabilization by Excipients during Freezing and Drying

2.1.1 Simulation and Validation Approach

The interaction between different excipients, including sugars, amino acids, and polyols, and human growth hormone (hGH) or lactate dehydrogenase (LDH) was evaluated using the molecular dynamics software GROMACS [119] (vers. 5.0.7). The hGH and LDH topology files were obtained from the RCSB Protein Data Bank (PDB 3hHR [120] and 5KKC [121], respectively), while the excipients topology

[1]Part of the work discussed in the present Chapter is based on already published papers [115–118]

files were obtained from the ATB database [122]. Each simulation box contained 1 protein molecule in the native state, together with a given number of cosolute molecules, in water. All simulations were conducted using a cubic box with periodic boundary conditions. The simulation box was modeled with the Gromos 54A7 force field [123], using explicit SPC/E water [124]. Appendix A contains more information about the water models used in this work. The initial structures were energy minimized using the steepest descent algorithm, then equilibrated at the desired temperature using the velocity rescaling thermostat [125]. The long range electrostatics were calculated by the Particle Mesh Ewald technique [126] and the resulting configuration was equilibrated at 1 bar with the Parrinello-Rahman barostat [127]. The final simulations were integrated using a 1 fs time step and saving coordinates and velocities every 10 ps. A Lennard-Jones cut-off of 1.4 nm was used for all the simulations. This value was chosen because the Gromos force field was originally parameterized with a Lennard-Jones cut-off of 1.4 nm. Simulations details are listed in Table 2.1.

Table 2.1: Details of simulations

#	Protein	Exc. Type	# Exc. Mol.	T, K	Duration, ns	Box Dim., nm
1	hGH	sucrose trehalose cellobiose lactose glucose sorbitol glycine	130	258	45	8.5x8.5x8.5
2	hGH	sucrose trehalose cellobiose lactose glucose sorbitol glycine	390	272	45	8.5x8.5x8.5
3	hGH	sucrose trehalose cellobiose lactose glucose sorbitol glycine	850	233	45	8.5x8.5x8.5
4	hGH	sucrose	130	272	45	8.5x8.5x8.5
5	hGH	sucrose	260	272	45	8.5x8.5x8.5

6	hGH	sucrose trehalose cellobiose lactose sorbitol glycine	2600	272	50	16x16x16
7	LDH	sucrose trehalose cellobiose mannitol histidine β-cyclodextrin	700 700 700 700 700 210	258	40	11.5x11.5x11.5
8	LDH	sucrose trehalose cellobiose mannitol histidine β-cyclodextrin	700 700 700 700 700 210	293	40	8x8x8

In the case of simulations 1-5, the objective was to investigate the effect of the freezing process on hGH stability. Simulation 1 corresponds to the conditions encountered by the formulation during cooling, close to the nucleation temperature. Simulations 2, 4 and 5, on the contrary, reproduce the environment encountered during freezing, when the product temperature is equal to the freezing equilibrium value and cryoconcentration occurs. Finally, simulation 3 corresponds to the situation encountered by the product at the end of freezing, when the asymptotic temperature value is reached. A scheme of the relationship between simulations and the freezing curve is reproduced in Figure 2.2.
While simulations 1 to 3 were performed for seven different excipients, simulations 4 and 5 were carried out for sucrose only and were aimed to verify the relative importance of different mechanisms upon variation of excipient concentration. Simulations 6 were again performed using hGH as model protein, but with a larger simulation box, to provide insight into the water entrapment mechanism. Glucose was not considered in this case, because simulations 1-3 already identified this excipient as the least effective among the sugars being analysed (as will be shown in the following). Finally, simulations 7 and 8 were used to investigate the stability of LDH during freezing, or in the dried state, respectively. In order to reduce the computational cost, only the excipients that were found to be promising stabilizers for hGH were also considered for LDH. However, some other excipients

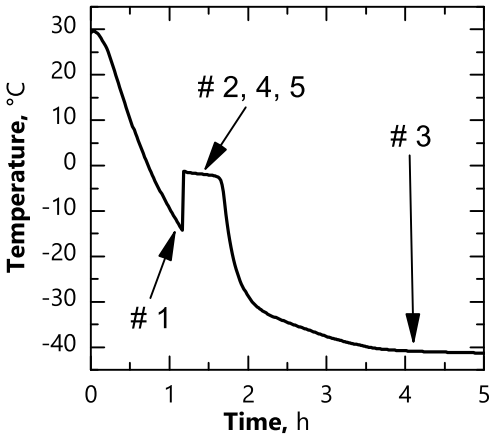

Figure 2.2: Temperature evolution of a solution in vials during freezing. Conditions for simulations of Table 2.1 are also shown.

(mannitol, histidine and β-cyclodextrin) were also included in simulations 7 and 8, mainly because of the availability of experimental data on these excipient-protein combinations.

Preferential Exclusion

In order to estimate preferential exclusion, the cumulative radial distribution function, and, thus, the average number of excipient, $n_{exc}(r)$, and water, $n_w(r)$, molecules within a distance r from the protein surface were computed. Hence, the degree of preferential exclusion was calculated plotting the function,

$$\beta = \frac{n_{exc}(r)/n_w(r)}{n_{exc}(\infty)/n_w(\infty)} \tag{2.1}$$

The excipient is preferentially excluded from the protein if β is smaller than 1 for small values of r, while preferentially interacts with the protein surface if $\beta > 1$.

To have a more immediate quantification of exclusion from the protein surface, the following parameter was also computed,

$$\Pi = \frac{\int_{r_0}^{r_{max}} \beta(r) dr}{\int_{r_0}^{r_{max}} 1 dr} \tag{2.2}$$

where r_0 corresponds to the minimum distance for which β becomes different from zero (0.224 nm in our simulations), while the upper limit $r_{max} = 1$ nm was arbitrarily chosen. A different choice of this upper limit for integration would anyway give similar results, provided that the difference $r_{max} - r_0$ is significantly large. A value of $\Pi < 1$ indicates preferential exclusion, and vice versa.

Vitrification

The vitrification theory was tested by computing the viscosity μ of the protein-free solution, relative to the viscosity μ_w of water at 300 K and 1 bar, using the periodic perturbation method [128]. In practice, a cosinusoidal acceleration is added to each particle each MD step, and the resulting velocity profile is measured. This profile is related to the solution viscosity according to the Navier-Stokes equation, and the viscosity can therefore be extracted. In the case of hGH, the viscosity was also measured indirectly from the diffusion coefficient D of the protein, using the Stokes-Einstein equation,

$$D = \frac{k_B T}{6\pi\mu R_g} \tag{2.3}$$

where k_B is the Boltzmann constant, T the absolute temperature and R_g the radius of gyration of the protein.

In the case of simulations 8, the mean square displacement $< u^2 >$ of LDH hydrogen atoms was also computed within each simulation box, on a nanosecond time scale. The value of $< u^2 >$ provides information on protein dynamics, since the hydrogen atoms reflect the motions of the side chains and backbone atoms to which they are bound.

Water Replacement

The water replacement mechanism was assessed by evaluating the parameter χ, which represents the relative contribution of hydrogen bonding between the protein and the excipient with respect to the total number of intermolecular hydrogen bonds formed by the protein.

$$\chi = \frac{\text{number of protein-excipient hydrogen bonds}}{\text{total number of protein-excipient and protein-water hydrogen bonds}} \tag{2.4}$$

To determine the presence of a hydrogen bond, a geometrical criterion was used, requiring that the distance between donor and acceptor was less than 0.35 nm, and that the angle formed between the hydrogen atom and the line joining the COMs of donor and acceptor was smaller than 30°.

Water Entrapment

To provide inside into the water entrapment theory, the radial distribution function $g(r)$ of the excipient molecules around the center of mass (COM) of hGH was computed in simulations 6.

The radial distribution function allows calculation of the local density $\rho(r)$ at a distance r from a given origin, once the average density ρ is known,

$$\rho(r) = \rho g(r) \tag{2.5}$$

In the interval $I = \{I_{min}, I_{max}\}$ where the radial distribution function of the excipient $g(r)$ is greater than 1, the excipient concentration is increased with respect to the average value. By contrast, a radial distribution function smaller than 1 indicates depletion of excipient molecules.

According to the water entrapment theory, an increase in excipient concentration around the protein surface should be related to formation of a cage. Moreover, water molecules confined within this cage should have a very low mobility. To verify this hypothesis, the mean square displacement of water molecules confined within the region of increased excipient concentration was compared to the value measured outside this region. The diffusion constant could also be calculated by least squares fitting a straight line through the mean square displacement as function of time.

Comparison with Experimental Data

For hGH, the simulation results were validated upon the experimental data reported in the works by Salnikova *et al.* [105] and Costantino *et al.* [103].

In the case of LDH, the experimental observations reported in the works by Nema and Avis [114], Izutsu *et al.* [110], and Al-Hussein and Gieseler [23] were used for validation. However, some experiments were also performed to demonstrate how the simulation approach could guide the selection of the formulation.

For these experiments, L-Lactate dehydrogenase from rabbit muscle (EC 1.1.1.27, a suspension in 3.2 M ammonium sulfate solution, pH approx. 7) was purchased from Sigma-Aldrich. LDH was dialyzed against 100 mM potassium phosphate buffer at pH 7.3, using the Sigma Aldrich Pur-A-Lyzer, 3.5 kDa MWCO kit. Dialysis was performed at 4 °C, and the buffer was changed 3 times (the first 2 times every 3 hours, while the third dialysis step was carried out overnight). The concentration of LDH after dialysis was determined using UV/VIS spectroscopy (6850 UV/VIS Spectrophotometer, Jenway, Stone, Staffordshire, UK). The peak at 280 nm was monitored, and an extinction coefficient of 1.44 ml/(mg cm) was used for calculations [66]. Three different formulations were prepared: A) 5 % w/w sucrose, B) 5 % w/w trehalose and C) 2.5 % w/w sucrose + 2.5 % w/w trehalose. All the formulations were prepared in 5 mM potassium phosphate buffer. The buffer species and excipients were also purchased from Sigma-Aldrich, as reagent grade or higher quality. The dialysed solution was then diluted in the formulations, so as to obtain a final LDH concentration of 10 μg/ml in each formulation. 12 vials/formulation (10R, 24 x 45, Nuova Ompi glass division, Stevanato Group, Piombino Dese, Italy) were filled with 2 ml of solution, and partially stoppered with silicon stoppers (West Pharmaceutical Services, Milano, Italy). Four more vials were filled with a 10 μg/ml LDH solution in 5 mM potassium phosphate buffer only (formulation

D). These vials were used as reference samples for the subsequent enzymatic activity assay. A freeze-dryer LyoBeta 25 (Telstar, Terrassa, Spain) was used for the freeze-drying cycles. In order to monitor the temperature of the product, two T-type copper/constantan miniature thermocouples were placed into selected vials within the batch. The batch contained 187 vials, and only the central ones were filled with the LDH formulations. Freezing was performed at -45 °C, using a 0.5 °C/min cooling rate, while primary and secondary drying were carried out at 10 Pa and -27 °C or 20 °C, respectively. A 4-hour ramp between primary and secondary drying was performed. The pressure within the chamber was monitored using both a Pirani and a capacitance (MKS Baratron) manometer. The activity of LDH was determined by monitoring the increase in absorbance at 450 nm, due to the reduction of NAD to NADH. 96-well-flat-bottom plates (Corning 96 well CellBIND Microplates, Merck) and a spectrophotometric multiwell plate reader (GDV Programmable MPT reader DV 990BV4, Roma, Italy) were used for the analyses. A standard curve built with 1.25 mM NADH standard was used for calculating the amount of NADH generated in each well. The reconstituted samples were diluted in 20 mM potassium phosphate buffer, pH 7.3. The dilution factor was chosen for each formulation so as to have the readings within the linear range of the standard curve.

2.1.2 Not All the Excipients are Equally Effective as Cryoprotectants

The efficiency of different excipients as cryoprotectants is related to their ability to be preferentially excluded from the protein surface [32,33,129,130]. This thermodynamic mechanism of stabilization should prevail during freezing, because there still is a not negligible amount of liquid water, and the excipient concentration is too low to allow the formation of a glassy matrix capable of kinetically hindering the protein movements. In MD, preferential exclusion of an excipient can be quantified using the β or Π parameters defined in Equations 2.1 and 2.2; $\beta < 1$ or $\Pi < 1$ indicate preferential exclusion, and vice versa. Simulations were used to evaluate different excipients for hGH or LDH, and the results are shown in Figures 2.3 and 2.4 for hGH, or 2.5 for LDH. While Figure 2.3 shows the β parameter for different excipients at the same concentration (simulations 1, 2 and 3 in Table 2.1), Figure 2.4 shows the preferential exclusion parameter for sucrose at different values of concentration (simulations 2, 4 and 5 in Table 2.1). Figure 2.3 can therefore be used to compare different excipients, while Figure 2.4 shows the effect of concentration on the mechanism of preferential exclusion. Finally, the values of Π for all simulations are listed in Table 2.2.
As regards hGH (Figure 2.3), the β parameter was found to be smaller than 1, at short distance from the protein surface, for almost all the simulations performed, indicating that the excipients were preferentially excluded from the protein. More

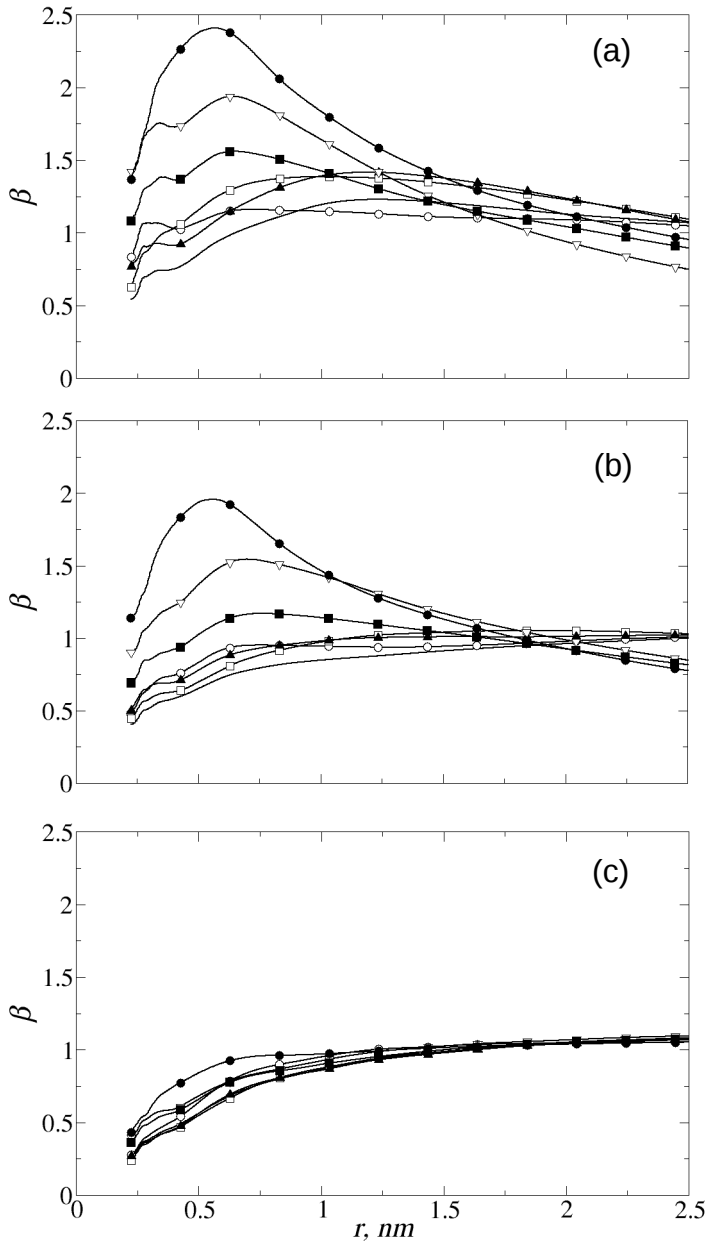

Figure 2.3: Plot of β parameter vs distance from hGH in the case of various excipient concentrations and temperatures: (a) 0.38 M at 258 K, simulation 1 in Table 2.1; (b) 1.15 M at 272 K, simulation 2 in Table 2.1; (c) 2.5 M at 233 K, simulation 3 in Table 2.1. The various curves refer to sucrose (—), trehalose (-○-), lactose (-□-), cellobiose (-▲-), glucose (-▽-), glycine (-●-), and sorbitol (-■-).

specifically, sucrose was slightly more excluded than trehalose, while cellobiose and

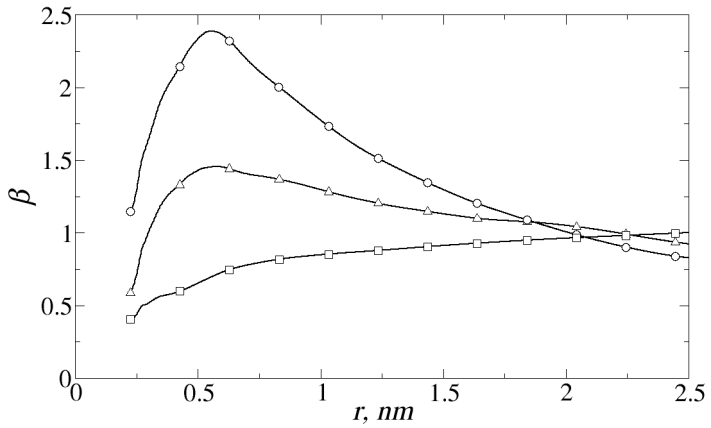

Figure 2.4: Plot of β parameter vs distance from hGH in simulations 4 (0.38 M at 272 K, -○-), 5 (0.80 M at 272 K, -△-) and 2 (1.15 M at 272 K,-□-), for sucrose as excipient.

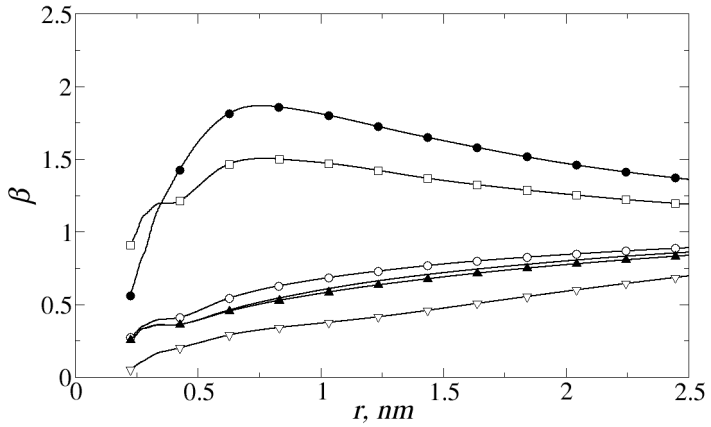

Figure 2.5: Plot of β parameter vs distance from LDH in simulations 7 (258 K) for sucrose (0.8 M, —), trehalose (0.8 M, -○-), mannitol (0.8 M, -□-), cellobiose (0.8 M, -▲-), β-cyclodextrin (0.25 M, -▽-), and histidine (0.8 M, -●-).

lactose had an intermediate behavior. Glucose, sorbitol and glycine were characterized by a significantly lower extent of preferential exclusion if compared to the disaccharides. This was also confirmed by the values of the Π parameter, as listed in Table 2.2. This means that, according to our analysis, glucose, sorbitol and glycine should not have a remarkable cryoprotective effect. Furthermore, when the concentration of water was low, as in simulation 3, the protein was preferentially hydrated, no matter what excipient was employed. Finally, the effects of changes in excipient

Table 2.2: Values of the Π parameter, as defined in Equation 2.2, for all simulations listed in Table 2.1

Simulation #	excipient	Π
1 (0.38 M at 258 K)	sucrose	0.93
	trehalose	1.10
	cellobiose	1.11
	lactose	1.20
	glucose	1.78
	sorbitol	1.44
	glycine	2.11
2 (1.15 M at 272 K)	sucrose	0.70
	trehalose	0.85
	cellobiose	0.83
	lactose	0.77
	glucose	1.37
	sorbitol	1.05
	glycine	1.70
3 (2.5 M at 233 K)	sucrose	0.63
	trehalose	0.71
	cellobiose	0.63
	lactose	0.62
	glucose	0.73
	sorbitol	0.71
	glycine	0.84
4 (0.38 M at 272 K)	sucrose	2.03
5 (0.80 M at 272 K)	sucrose	1.30
7 (0.80 M at 258 K)	sucrose	0.45
	trehalose	0.51
	cellobiose	0.44
	mannitol	1.36
	histidine	1.58
(0.25 M at 258 K)	β-cyclodextrin	0.26

concentration were studied. For this purpose 3 simulations were performed (number 4, 5 and 2 in Table 2.1) where the concentration of sucrose molecules was linearly increased. As shown in Figure 2.4 and Table 2.2, an increase in excipient concentration resulted in a corresponding enhancement of preferential exclusion from the protein.

As regards LDH (Figure 2.5 and Table 2.2), it is clear that mannitol and histidine were not excluded from the protein surface, but, rather, they preferentially

interacted with LDH, displacing the water molecules. In fact, their β parameter was greater than 1 for small values of the distance r from the protein surface, and they showed $\Pi > 1$. In contrast, the β-cyclodextrin molecules seemed to be the most excluded from the protein surface, showing values of β which were remarkably smaller than 1 close to the surface of LDH, and a correspondingly low value of Π (0.26). Also in the case of LDH, sucrose was slightly more excluded than trehalose from the protein surface, while cellobiose behaved similarly to sucrose.

The results for hGH and LDH compare fairly well with one another. This observation suggests a generally applicable behavior of the osmolytes toward different proteins, which is not sequence-specific. Moreover, the predictions of MD simulations are confirmed by experimental data. For instance, Nema and Avis [114] studied the residual activity of LDH after freeze-thawing cycles, in the presence of different excipients. They found that the addition of 0.9% (w/v) β-cyclodextrin protected the protein more than a 5 % (w/v) trehalose solution, with a remaining activity of 31 % and 26 %, respectively. As regards the disaccharides, the MD simulations suggest that sucrose is more excluded than trehalose from the protein surface. This, again, is in line with the experimental work by Nema and Avis, where a 5 % (w/w) sucrose formulation is reported to preserve the LDH activity much better (73 %) than a trehalose solution (26 %) at the same solid content. Nema and Avis also showed that the stabilizing effect of sucrose was concentration-dependent, with the highest LDH remaining activity ($\approx$ 100 %) measured at the highest solid content (34.2 % w/v, 1 M). This strong dependence on the solid content is also in accordance with the increase in preferential exclusion upon increase in the osmolyte concentration observed in Figure 2.4. Finally, the efficiency of mannitol and histidine as protectants, according to our simulations, should be quite poor. They are generally used in typical formulations not as protein stabilizers, but as buffer components, in the case of histidine, or bulking agents, in the case of mannitol. In line with our simulations, Nema and Avis observed that LDH showed only 6% recovery of enzymatic activity after freeze-thawing when mannitol was used as stabilizer, confirming its poor efficiency as a cryoprotectant. However, this diminished efficiency of mannitol as a protectant may also be related to its tendency to undergo crystallization.

Considering cryoprotection, i.e. preferential exclusion, the following "ranking" of excipients could therefore be written according to MD simulations:

- hGH as model protein
  sucrose, lactose > cellobiose, trehalose >> sorbitol, glucose >> glycine

- LDH as model protein
  β-cyclodextrin >> cellobiose, sucrose > trehalose >>histidine, mannitol

where only a slight difference exists between sucrose, trehalose, cellobiose and lactose (symbol >), while a huge separation can be observed between the disaccharides and polyols, monosaccharides or amino acids (glucose, sorbitol, glycine, histidine, mannitol) (symbol >>). β-cyclodextrin molecules were the most excluded from the surface of LDH, probably because of their large molecular volume that results in a steric hindrance effect.

2.1.3 Lyoprotection Becomes Dominant at High Excipient Concentration

Beside the thermodynamic mechanism of preferential exclusion, water replacement and vitrification could arise when the osmolyte concentration is sufficiently high.

Therefore, as regards lyoprotection, the solution viscosity μ and the ability of different excipients to hydrogen bond with the protein (χ value) were used as critical parameters, in line with the vitrification [42] and water replacement hypotheses [131, 132], respectively.

The solution viscosity for all simulations with hGH as model protein was computed using either the periodic perturbation method (PP) or the Stokes-Einstein equation (SE), and is reported in Table 2.3. The diffusion coefficient, as measured from protein mean square displacement, and the radius of gyration of the protein are listed, as well.
It is possible to notice that the viscosity increased with excipient concentration, as expected during the cryoconcentration process. The osmolyte concentration in simulation 3 is still lower than in a typical glass matrix, where the viscosity may be even 10 orders of magnitude higher that that of water at room temperature. It was decided to limit the excipient concentration because the simulation of a glass would not provide useful information, as the movements of the osmolyte molecules within the simulation box would be too slow, and a converged configuration would never be reached. Nevertheless, the simulations performed allow the identification of a clear trend, with the increase in viscosity being highly non-linear. The disaccharides showed higher values of viscosity than glucose and glycine, especially at the highest concentration tested. Furthermore, the viscosity of sorbitol was intermediate between those of disaccharides and glucose or glycine.

In literature, there are several correlations describing the concentration and temperature dependence of the viscosity of disaccharides solutions. Among these, the Genotelle equation may be used [133], which includes the effects of the sugar mole fraction x and of the reduced temperature $\phi = (30 - T)/(91 + T)$ (where T is in °C),

$$\log_{10} \mu = a_1 + a_2 x + \phi(b_1 + b_2 x^n) \tag{2.6}$$

This correlation was used, in its range of validity, to validate the results obtained

Table 2.3: Viscosity of the solutions as evaluated by simulations. Comparison between values of sucrose and trehalose viscosity as obtained by simulations or by Genotelle equation (GE) [133] is also shown. The coefficients of Genotelle equation were obtained from Longinotti and Corti [134]

Simulation Id.	excipient	μ/μ_w (PP)	D, m^2 s^{-1}	R_g, nm	μ/μ_w (SE)	μ/μ_w (GE)
1 (0.38 M at 258 K)	sucrose	6.57	$8.0\cdot 10^{-12}$	1.74	15.2	6.3
	trehalose	7.05	$9.5\cdot 10^{-12}$	1.75	13	7.6
	cellobiose	7.07	$1.3\cdot 10^{-11}$	1.74	9.37	
	lactose	7.28	$9.6\cdot 10^{-12}$	1.75	12.9	
	glucose	5.13	$3.0\cdot 10^{-11}$	1.75	4.68	
	sorbitol	5.27	$1.3\cdot 10^{-11}$	1.71	9.37	
	glycine	4.75	$3.5\cdot 10^{-11}$	1.75	3.51	
2 (1.15 M at 272 K)	sucrose	14.32	$1.6\cdot 10^{-12}$	1.73	85	12.8
	trehalose	14.58	$8.1\cdot 10^{-13}$	1.74	165	15.1
	cellobiose	18.88	$2.0\cdot 10^{-12}$	1.76	64	
	lactose	19.2	$9.0\cdot 10^{-13}$	1.74	149	
	glucose	4.85	$1.1\cdot 10^{-11}$	1.75	13	
	sorbitol	6.55	$1.6\cdot 10^{-11}$	1.73	8.4	
	glycine	3.35	$5.1\cdot 10^{-11}$	1.74	2.3	
3 (2.5 M at 233 K)	sucrose	1298	$1.6\cdot 10^{-13}$	1.62	763	
	trehalose	1410	$2.4\cdot 10^{-13}$	1.63	509	
	cellobiose	1837	$1.9\cdot 10^{-13}$	1.64	650	
	lactose	1870	$2.8\cdot 10^{-13}$	1.61	436	
	glucose	438	$8.8\cdot 10^{-13}$	1.64	137	
	sorbitol	638	$1.9\cdot 10^{-13}$	1.66	638	
	glycine	303	$8.0\cdot 10^{-13}$	1.64	152	
4 (0.38 M at 272 K)	sucrose	4.18	$1.1\cdot 10^{-11}$	1.75	12	3.1
5 (0.80 M at 272 K)	sucrose	6.78	$3.6\cdot 10^{-12}$	1.75	36.3	5.8

by simulations. The empirical parameters a_1, a_2, b_1, b_2 and n for sucrose and trehalose were obtained from the work by Longinotti and Corti [134], where the experimental viscosity of aqueous solutions of these disaccharides was analyzed over a wide range of concentrations and temperatures, covering the normal liquid and supercooled liquid regions. As can be seen comparing the values by Genotelle equation, listed in the last column of Table 2.3, and by simulations, the change in viscosity upon concentration and temperature was fairly well predicted by molecular dynamics. Moreover, the periodic perturbation method seemed to be more accurate than the Stokes-Einstein equation. This is not surprising, since the Stokes-Einstein equation is an approximate formula, that works well for spherical solutes at infinite

dilution only.

As previously mentioned, the vitrification hypothesis generally associates a higher viscosity to slower chemical and physical degradation rates of proteins. Viscosity is closely related to the α relaxation processes, i.e., the slowest and strongly temperature dependent motions of a glass [135]. However, some studies seem to suggest that no correlation exists between the α relaxation time τ_α and protein stability [136,137]. In contrast, it was reported [138–140] that the fast β relaxation processes have a strong relationship with the protein degradation rates. Glasses exhibit at least two high-frequency relaxations, i.e., the fast β and the slower Johari-Goldstein relaxation. The fastest motion occurs on a timescale of picoseconds, while the intermediate one occurs on microsecond to millisecond timescales. The correlation between β relaxation dynamics and protein stability in lyophilized products can be experimentally observed by measuring the mean square displacement $< u^2 >$ of hydrogen atoms using incoherent inelastic neutron scattering [141]. $< u^2 >$ is directly proportional to the amplitude of the fast β processes and was demonstrated to be strongly related with protein stability in sugar glasses [138, 139]. In those studies, $< u^2 >$ measurements were made on a nanosecond timescale, and a similar procedure was reproduced in simulations 8 using a computational approach. The results of this analysis are shown in Figure 2.6, where the value of $< u^2 >$ for the LDH hydrogen atoms is shown for all the formulations being considered. It is apparent that the disaccharides should provide the best kinetic stabilization of the protein structure, also if we consider the β relaxation time as critical parameter. According to our results, mannitol should be the worst stabilizer, followed by histidine and the β-cyclodextrin. The MD results obtained are therefore in accordance with the experimental work by Al-Hussein and Gieseler [23], where sucrose was proven to be a significantly better lyoprotectant than histidine or mannitol. Also Izutsu *et al.* [110] found that a 200 mM sucrose solution protected LDH (30 % remaining activity) from freeze-drying stresses more than a 400 mM mannitol formulation (14 % residual activity only), in line with the MD simulations.

The hydrogen bonding network within each simulation box was also analyzed. As evident from Figure 2.7, the relative contribution of hydrogen bonding between hGH or LDH and the osmolytes increased with increasing excipient concentration. In particular, it was small when the number of water molecules was sufficiently high, as in simulations 1, 2, 4, 5, and 7, while it became predominant when the number of water molecules was no more sufficient for a complete protein hydration, as in simulations 3 or 8. Hence, these data suggest that water replacement becomes dominant in the solid state, when most of the water has been removed, while preferential exclusion should be the prevailing mechanism in the liquid state.

Remarkably, the disaccharides, which formed the greatest number of hydrogen bonds with the protein, also increased the solution viscosity the most (Table 2.3), or resulted in the slowest β relaxation processes (Figure 2.6). Among the disaccharides, trehalose and cellobiose could form a larger number of hydrogen bonds

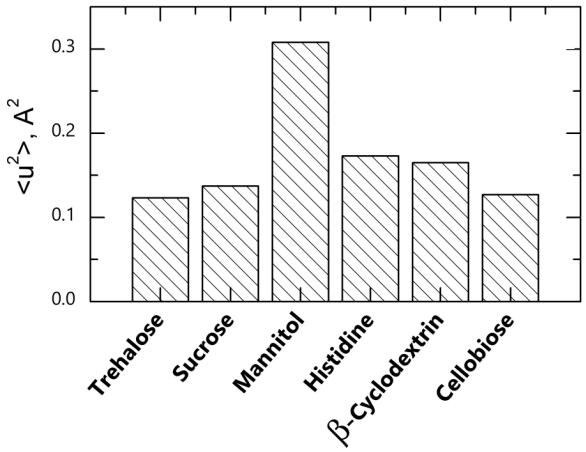

Figure 2.6: Amplitude of the mean square displacement of LDH hydrogen atoms on a nanosecond timescale, in simulations 8 (2.5 M for trehalose, sucrose, mannitol, histidine and cellobiose, 0.75 M for β-cyclodextrin, at 293 K).

with hGH than sucrose or lactose (see Figure 2.7a), and a similar conclusion can be drawn from Figure 2.7b for LDH if we compare cellobiose, trehalose and sucrose.

On the contrary, the excipients which were the least effective in hydrogen bonding to the protein were glucose, sorbitol and glycine in the case of hGH, or β-cyclodextrin, histidine and mannitol for LDH as model protein. In particular, β-cyclodextrin resulted in the lowest number of hydrogen bonds with LDH, maybe because of a steric hindrance effect, due to its relatively great dimensions. However, it should also be considered that the molar concentration of β-cyclodextrin in the simulation box was smaller if compared to the other excipients.

Moving now to the water entrapment mechanism, the radial distribution function $g(r)$ of excipient molecules around the center of mass of hGH was evaluated in simulations 6. The profiles obtained are shown in Figure 2.8.
We named $I = \{I_{min}, I_{max}\}$ the interval where $g(r) > 1$, i.e., the region where the excipient concentration was higher than the bulk value. More specifically, I_{min} was defined as the first point where the radial distribution function becomes equal to 1. By contrast, I_{max} was defined as the first point where $g(r)$ equals 1 again, after having passed through its global maximum.

As it is possible to see in Figure 2.8, different excipients showed a different behavior. First of all, the interval I of increased excipient density changed position according to the molecular volume of the protectant under investigation, as shown in Table 2.4. That is, smaller molecules, such as sorbitol and glycine, approached the protein the most and concentrated between approximately 1.5 and 3.5 nm from the protein center of mass. By contrast, bigger excipients, such as the disaccharides sucrose, trehalose, cellobiose and lactose, were found to crowd further from the

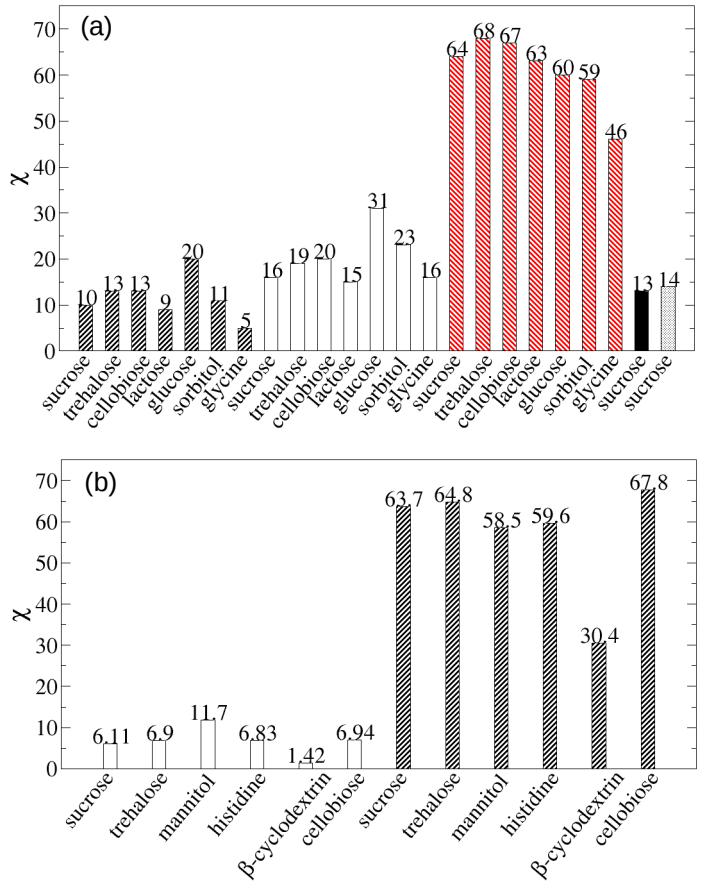

Figure 2.7: (a) χ parameter for hGH as model protein, during simulations 1 (0.38 M at 258 K, black striped bars), 2 (1.15 M at 272 K, plain white bars), 3 (2.5 M at 233 K, red striped bars), 4 (0.38 M at 272 K, plain black bar), and 5 (0.80 M at 272 K, dotted black bar). (b) χ parameter for LDH as model protein, during simulations 7 (0.8 M for trehalose, sucrose, mannitol, histidine and cellobiose, 0.25 M for β-cyclodextrin at 293 K, plain white bars), and 8 (2.5 M for trehalose, sucrose, mannitol, histidine and cellobiose, 0.75 M for β-cyclodextrin at 293 K, black striped bars).

protein, approximately between 2.5 and 4 nm from its center of mass.

Moreover, while some molecules, such as sorbitol, showed a marked increase in numerical density within the chosen volume, this effect was much smaller for other excipients, such as sucrose. The increase in excipient concentration which was observed could indicate, according to the water entrapment theory, that the excipients formed a cage around the protein surface. Moreover, water molecules confined within the cage should have a very low mobility. To verify this hypothesis,

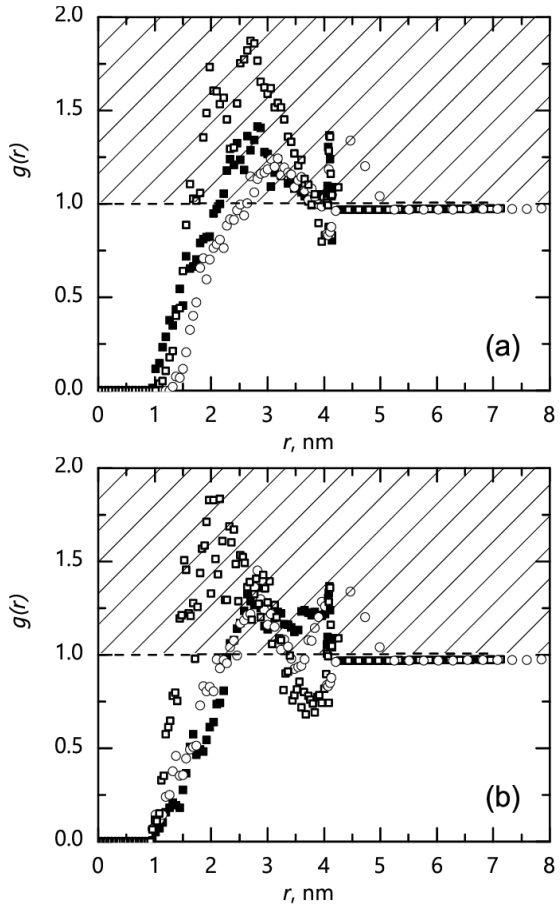

Figure 2.8: (a) Plot of the radial distribution functions of sorbitol (□), sucrose (○) and cellobiose (■) molecules around the center of mass of hGH. (b) Plot of the radial distribution functions of glycine (□), lactose (○) and trehalose (■) molecules around the center of mass hGH. The shaded area refers to zones of increased excipient concentration.

the mean square displacement and the diffusion coefficient of water molecules within a distance I_{max} from the center of mass of the protein was computed. In order to make a comparison, the mean square displacement and diffusion coefficient of water molecules outside the cage, i.e., being more distant than I_{max} from the protein center of mass, was also evaluated. Figure 2.9 shows the result of this analysis for the case of sorbitol as excipient, but similar profiles were obtained for all of the other excipients considered in this study.
The slope of the mean square displacement against time is directly proportional to the diffusion coefficient. Thus, as evident from Figure 2.9, the diffusion coefficient of water molecules confined inside the cage was significantly smaller than the same

Table 2.4: Values of I_{min} and I_{max} for the different excipients being considered in simulations 6 (1.15 M at 272 K) of Table 2.1

excipient	Molecular Volume, Å^3	I_{min}, nm	I_{max}, nm
sucrose	282.5	2.4	4.0
trehalose	281.6	2.3	4.1
cellobiose	283.1	2.2	4.0
lactose	279.2	2.2	3.4
sorbitol	160.9	1.4	3.8
glycine	68.1	1.4	3.3

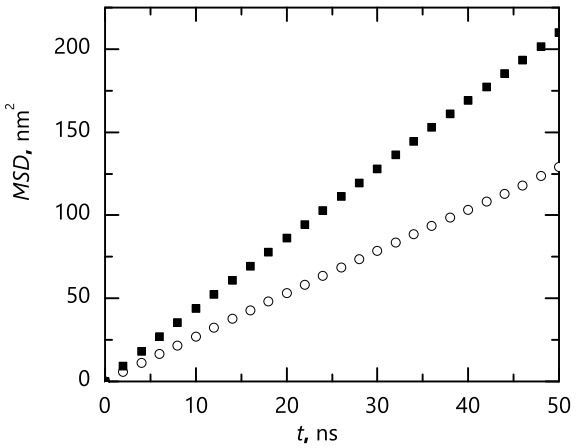

Figure 2.9: Mean square displacement of water molecules within (○) and outside (■) the region of increased excipient density, for the simulation involving sorbitol. The slope of each curve is related to the diffusion coefficient.

value measured outside the region of increased excipient density. Therefore, these results seem to support the water entrapment theory. In fact, a region of increased excipient concentration is formed around the protein surface, and water molecules within this region are substantially slowed down. Similar results were obtained for trehalose in a previous study [45].

However, not all the excipients under investigation crowded to the same extent around the protein surface. In order to quantify the different propensity of excipients in giving water entrapment, the following variables were evaluated: 1) cage ratio, defined as the average value of the radial distribution function of the excipient between I_{min} and I_{max}. The greater the cage ratio was, the more the excipient molecules crowded around the protein surface, 2)$c_{I_{min}}^{I_{max}}$, which is the volume by volume concentration of excipient molecules between I_{min} and I_{max}. The

larger it was, the more pronounced the cage effect was, 3) δ, calculated as the ratio between the diffusion coefficients of water molecules inside (numerator) and outside (denominator) the cage

$$\delta = \frac{D(< I_{max})}{D(> I_{max})} \tag{2.7}$$

The smaller δ was, the more intense water entrapment was, 4) $D(< I_{max})$, i.e., the diffusion coefficient of water molecules within the cage. The smaller it was, the more water molecules within the cage were immobilized.

These four variables are listed, for all the excipients investigated, in Table 2.5.

A remarkable result is that the crowding effect for the disaccharides was less important (smaller cage ratio) than for glycine and sorbitol. More specifically, sorbitol showed the largest cage ratio and lowest δ. In spite of this, neither sorbitol nor glycine were able to immobilize water molecules to the same extent as the sugars, and showed high values of $D(< I_{max})$. This observation can be explained considering that sorbitol and glycine molecules are smaller than the disaccharides. Hence, although these excipients crowded around the protein surface, the volume by volume density of the cage they formed was smaller than in the case of bigger molecules. This result is confirmed by the values of $c_{I_{min}}^{I_{max}}$ listed in Table 2.5. Hence, the lowest values of $D(< I_{max})$ were in any case observed for the sugars, that also showed the largest values of $c_{I_{min}}^{I_{max}}$. Among the disaccharides, cellobiose and trehalose immobilized water molecules the most, followed by lactose and sucrose.

Table 2.5: Parameters related to water entrapment for all the simulations (at conditions 6 in Table 2.1, 1.15 M at 272 K) performed.

Excipient	Cage Ratio	$c_{I_{min}}^{I_{max}}$, m^3/m^3	δ, %	$D(< I_{max})$, 10^{-5} cm^2/s
cellobiose	1.23	0.250	64	0.2066
trehalose	1.19	0.242	73	0.2008
sucrose	1.11	0.225	76	0.2221
lactose	1.18	0.235	66	0.2272
glycine	1.40	0.014	71	0.7865
sorbitol	1.47	0.087	61	0.4232

The values of the diffusion coefficient measured for confined, $D(< I_{max})$, and bulk, $D(> I_{max})$, water in this work compare fairly well with experimental data. For example, in quasi-elastic neutron-scattering studies of water at 273 K close to a C-phycocyanin protein [142], values of diffusion coefficient equal to about $0.8 \cdot 10^{-5}$ and $1.1 \cdot 10^{-5}$ cm^2/s were measured for confined and bulk water, respectively. These

values are not too far from those we obtained, for instance, in the case of glycine as excipient. The lower values observed for sorbitol and the disaccharides can be explained considering the remarkable vitrification and clustering effects provided by these excipients, which were not considered in the abovementioned experimental work.

It is interesting to observe that the excipients that formed the highest number of hydrogen bonds with the protein, generally also resulted in the largest values of μ (or lowest $< u^2 >$), and in the lowest values of $D(< I_{max})$. Considering lyoprotection, the following "ranking" of excipients could therefore be written, independently of the mechanism (vitrification, water replacement, water entrapment) considered:

- hGH as model protein
  cellobiose, trehalose > sucrose, lactose >> sorbitol, glucose >> glycine

- LDH as model protein
  cellobiose, trehalose > sucrose >>histidine, mannitol, β-cyclodextrin

where, again, only a slight difference exists between sucrose, trehalose, cellobiose and lactose (symbol >), while a huge separation can be observed between the disaccharides and the other types of excipients (symbol >>). This is confirmed by the well-known experimental observation that the disaccharides result in the best protein stabilization. The exipient-protein combinations simulated in this work were selected because of the existence of experimental data; a comparison with these data will be presented in section 2.1.5. However, a possible relation between the excipients molecular properties and their efficiency as protein stabilizers will first be investigated.

2.1.4 Molecular Properties can Explain the Excipients' Behavior as Protein Stabilizers

Taking into account the MD results previously presented, a simple phenomenological model could be proposed to explain the behaviour of different excipients as protectants. During our simulations, the excipient molecules moved randomly because of diffusion, and the smaller the excipient molecules were, the more they could approach the protein. For instance, in Table 2.7 the average minimum distance d between hGH surface and the excipient molecules is shown, for all the excipients under investigation in simulations 6. It is possible to notice that the smallest excipients, i.e., sorbitol and glycine, approached the protein the most.

In general excipient molecules were repelled by the protein surface, and this resulted in preferential exclusion. However, excipient molecules that hydrogen bonded to the protein remained in the volume perturbed by the protein, generating the water replacement scenario.

To support this last hypothesis, the effect of the surface charge properties of hGH on the arrangement of excipient molecules was studied. In order to do this, the minimum distance between the surface of each protein residue and the surrounding excipients was calculated. Moreover, the hydrophobic and hydrophilic surface area exposed by each amino acid was also evaluated. We then defined as hydrophobic residues those whose hydrophobic surface area was larger than the hydrophilic one, and vice versa. Finally, the average minimum distance between excipient molecules and the hydrophobic, d_{NP}, or hydrophilic, d_P, patches of the protein could be computed. The results obtained from this investigation are shown in Table 2.7.

Table 2.7: Average minimum distance between excipient molecules and the protein surface, d, or the hydrophobic, d_{NP}, and hydrophilic, d_P, patches of hGH, for all the excipients investigated in simulations 6 (1.15 M at 272 K) of Table 2.1.

Excipient	d, nm	d_{NP}, nm	d_P, nm
cellobiose	0.38	0.42	0.35
trehalose	0.41	0.43	0.39
sucrose	0.44	0.47	0.42
lactose	0.40	0.46	0.37
glycine	0.37	0.38	0.37
sorbitol	0.36	0.38	0.36

It is possible to notice that, for all the excipients, d_P was smaller than d_{NP}; therefore, the osmolyte molecules mainly approached the protein nearby its hydrophilic patches. This was true especially for the disaccharides, while the difference between d_P and d_{NP} was smaller for sorbitol and glycine. This effect of the surface charge properties of the protein may be related to the formation of hydrogen bonds between the excipient molecules and the hydrophilic patches exposed onto the protein surface, in accordance with the water replacement theory. In line with this hypothesis, the measured values of d_P were very close to 0.35 nm, that is the typical hydrogen bond length.

Even if some excipient molecules hydrogen-bonded to hGH, a void space was anyway formed between the osmolytes and the protein, because of steric hindrance and repulsive forces. Water confined in this void space was immobilized because of cage effect. This last observation is in accordance with the water entrapment theory.

Thus, according to the phenomenological model here outlined, the mechanisms that were proposed to explain protein stabilization in the dried state, i.e., vitrification, water replacement and water entrapment, may be different consequences of a same process. This process may consist in the formation of a matrix characterized by a high degree of structure ordering and a dense, compact hydrogen bonding network. This network is formed between excipient molecules, leading to

formation of a cage structure, and thus, to the water entrapment scenario. However, an equally strong network is established between excipient molecules and the hydrophilic patches of the protein, according to the water replacement theory. Finally, this highly structured matrix, where all the molecules are connected by means of hydrogen bonds, shows a very high viscosity, in line with the vitrification hypothesis.

According to these results, the mechanism at the basis of protein stabilization in the dried state may be the formation of a dense hydrogen bonding network, that kinetically prevents the protein motions responsible for unfolding and aggregation. From this viewpoint, the water replacement, vitrification, and water entrapment hypotheses, which have been so far considered as different, alternative mechanisms to explain the same phenomenon, may actually be complementary aspects of the same process.

As a further step, the presence of a correlation between the molecular properties of the excipients investigated, and their efficiency as protein stabilizers was also investigated. In the framework of the phenomenological model just described, both the molecular volume and the hydrogen bonding propensity of a molecule should be related to its behavior as a protectant. A large molecular volume V should correlate with a stronger preferential exclusion from the protein surface. This is the case, for instance, of β-cyclodextrin in Figure 2.5. Figure 2.10 shows the Π parameter, as defined in Equation 2.2, as function of the molecolar volume. An almost linear relation seems to exist between these two variables, as evident from the good R^2 values of the fitting (0.915 for Figure 2.10a and 0.996 for Figure 2.10b). A similar relation between the magnitude of preferential exclusion and the molecular volumes of excipients was also experimentally observed in a recent experimental work on monoclonal antibodies [143].

On the contrary, a high hydrogen bonding propensity should result in the formation of a compact, viscous matrix that lyoprotects the protein structure. As general guideline, a large number of hydrogen bond donors and acceptors is a crucial requirement for a lyoprotectant, but is not sufficient, as these hydrogen bonding sites must also be available on the excipient surface. It is clear that, for a given number of hydrogen bond donors/acceptors, the possibility to interact with the other molecules is enhanced if the excipient exposes a larger specific surface area. A parameter γ could therefore be defined, that takes into account all these considerations,

$$\gamma = \text{(number of hydrogen bond donors and acceptors)} \times \text{Area/Volume} \qquad (2.8)$$

The values of γ and V for the excipients considered in this work are shown in Figure 2.11. As evident from Figure 2.12, the γ parameter seems to correlate fairly well with the lyoprotective efficiency of the molecules under investigation. The disaccharides, in particular trehalose and cellobiose, show a significantly larger

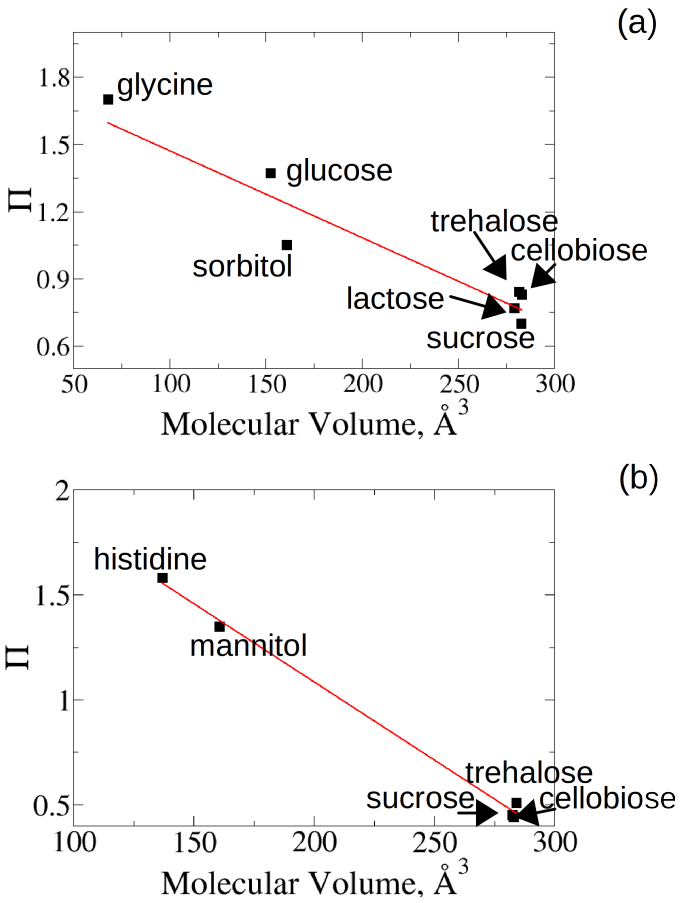

Figure 2.10: Π parameter, as defined in Equation 2.2, as function of the excipient molecular volume in the case of (a) sim. 2 (hGH, 1.15 M at 272 K) or (b) sim. 7 (LDH, 0.80 M at 258 K) in Table 2.1. The R^2 values of the fitting are 0.915 for (a) and 0.996 for (b).

hydrogen bonding propensity than polyols, monosaccharides and amino acids, and they correspondingly result in stronger and more compact glassy matrices. As it is not easy to know the real availability of hydrogen bonding sites in cyclic osmolytes, where the molecule surface can never be totally exposed to the external environment, the torus-like shaped β-cyclodextrin has not been considered.
From the molecular dynamics simulations here performed, the following conclusions could therefore be drawn:

- The mechanisms of protein stabilization change significantly during the freeze drying process, mostly as a result of the increase in excipient concentration.

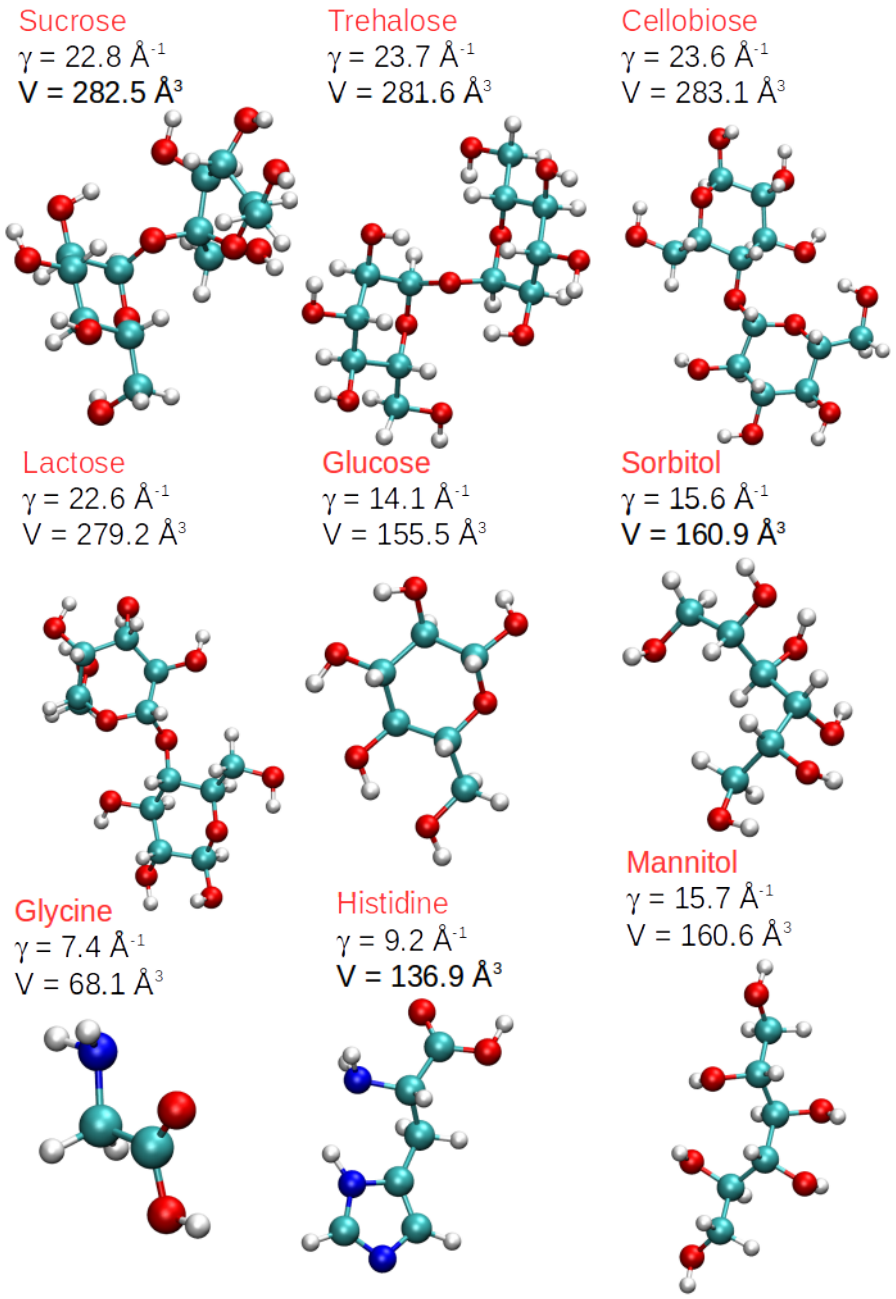

Figure 2.11: Values of molecular volume V and hydrogen-bonding propensity parameter γ for each excipient molecule. The ball-and-stick model of each molecule is also shown, where oxygen atoms are colored in red, carbon atoms in cyan, hydrogen atoms in white and nitrogen atoms in blue.

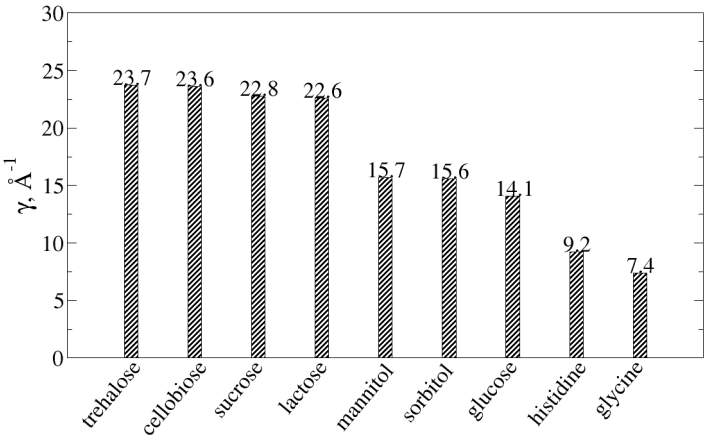

Figure 2.12: Values of the hydrogen-bonding propensity parameter γ for each excipient molecule, from biggest (left) to smallest (right).

Preferential exclusion prevails during freezing, while vitrification, water replacement and water entrapment become dominant in the dried state;

- Some evidence suggests that the mechanisms of lyoprotection should all be related to the formation of a dense, compact hydrogen bonding network;
- Not all the excipients are equally effective. The disaccharides should be better than polyols, monosaccharides and amino acids both during freezing and drying;
- Also among the disaccharides, small differences may exist, with sucrose and lactose being extremely good cryoprotectants, and trehalose and cellobiose being slightly better for lyoprotection;
- Some molecular properties seem to correlate with the protective effect of stabilizers. The higher the molecular volume is, the more the osmolyte is excluded from the protein surface. A high hydrogen bonding propensity is linked to the efficiency as a lyoprotectant.

In the following, we will show how the results obtained compare with experimental data, and illustrate how the MD simulations could guide the choice of the formulation for biopharmaceuticals.

2.1.5 Comparison with Experimental Data

The results obtained for hGH from MD simulations seem to be in line with the data reported in Salnikova *et al.* [105], who observed that the midpoint temperature of

thermally induced aggregation in the liquid state was delayed by 2°C in presence of trehalose, and by 6°C in presence of sucrose (Table 2.8). This is in line with our finding that sucrose should be a better stabilizer (more preferentially excluded) than trehalose in the liquid state. They also reported, for the hGH/sucrose formulation, a smaller amount of insoluble aggregate formation after freeze-drying with respect to the hGH/trehalose formulation.

Table 2.8: Comparison of sucrose and trehalose as excipients for hGH. Data from Salnikova *et al.* [105]

Formulation	Midpoint T °C	% insoluble aggregates after freeze-drying	Rate of insoluble aggregate formation in the solid state (% ins. aggregate/$\sqrt{\text{week}}$)
hGH/sucrose	80.4±1.1	0.35±0.04	2.38 ± 0.01
hGH/trehalose	76.2±0.1	5.0±0.2	1.97± 0.06

However, the trehalose formulation led to a smaller rate of aggregate formation in the solid state (Table 2.8).

According to the MD results, the smaller amount of high order aggregates in the sucrose-based formulation, immediately after freeze-drying, may again be related to the preferential exclusion of sucrose from the protein surface. However, as regards stability in the solid state, the preferential exclusion theory does not apply any more, and the formation of a compact, glassy matrix is at the basis of protein preservation. Because of its molecular conformation, that results in a high hydrogen bonding propensity, trehalose is therefore a very efficient lyoprotectant.

Salnikova *et al.* [105] did not analyze the effects of cellobiose and lactose, but according to our results cellobiose should behave similarly to trehalose, while lactose should be slightly less effective than cellobiose and trehalose as lyoprotectant. These hypotheses are consistent with the work by Costantino *et al.* [103], who experimentally determined the rate constants for hGH deterioration in the solid state and in the presence of different formulations (Table 2.9).
Costantino *et al.* analyzed a sorbitol-based formulation, as well, and they concluded that this polyol was not as good as protein stabilizer as the disaccharides, in good agreement with our MD results.

Let us now illustrate how the MD simulations could guide the choice of the formulation, using LDH as model protein. According to the MD analysis, the addition of mannitol or histidine to the LDH formulation should not be the best choice, as these two osmolytes do not provide remarkable thermodynamic or kinetic stabilization. From the simulation results, we would instead conclude that an effective formulation should include the disaccharides, that provide the best protein stabilization during both freezing and drying. However, a slight difference seems to

Table 2.9: Pseudo first-order rate constants for deterioration of hGH when stored in the presence of various excipients. Data from Costantino *et al.* [103]

Excipient	Psuedo first-order rate constant for deterioration (x10^{-3} $days^{-1}$)
trehalose	0.4±0.2
cellobiose	0.0±0.2
lactose	1.3±0.2
sorbitol	3.5±0.3

exist between the disaccharides as well, with sucrose being better for preferential exclusion and trehalose having a higher hydrogen-bonding propensity. From the simulation results, we would therefore conclude that a sucrose/trehalose mixture may provide the best stabilization during both freezing and drying, because of sucrose being preferentially excluded from the protein surface in the liquid state, and trehalose forming a very viscous matrix in the solid cake. Therefore, the stability of LDH in 3 different formulations was investigated: A) 5 % w/w sucrose, B) 5 % w/w trehalose and C) 2.5 % w/w sucrose + 2.5 % w/w trehalose. For reference, some vials were also filled with an LDH solution in 5 mM potassium phosphate buffer only, without any stabilizer (formulation D). After freeze drying, the remaining enzymatic activity of LDH in each formulation was monitored, as shown in Figure 2.13. The results are reported as remaining activity (%), relative to the buffer-only formulation (formulation D).

In the presence of sucrose, the preservation of LDH activity was about 30 % higher than in formulation D. We also observed that trehalose was a better excipient than sucrose, improving the stability of LDH by 43 % if compared to the buffer-only formulation. However, the best preservation of the enzymatic activity was achieved when both sucrose and trehalose were added to the formulation. In this case, the stability of LDH increased by about 50 % with respect to the buffer-only formulation. The synergistic effects resulting from the combination of multiple excipients seem therefore to be a subject that is worth investigating, and should be further addressed both computationally and experimentally.

Considering the results obtained in this Chapter, the following approach could be proposed for the design of a formulation. First, a pre-screening of the formulations could be carried out using MD simulations, in order to identify the mechanisms at the basis of protein stabilization by different additives. This information, which is at present not accessible by experimental techniques, can then guide the selection of an optimal formulation. Afterwards, the experimental analysis may focus on those formulations that, according to the simulation results, should be the most effective. This experimental campaign could therefore be more targeted, making the whole process extremely time and cost effective. MD simulations may

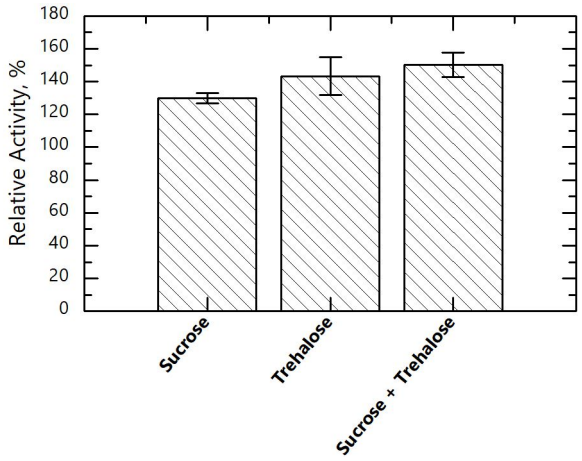

Figure 2.13: Residual activity of LDH after freeze drying in: A) 5 % w/w sucrose, B) 5 % w/w trehalose and C) 2.5 % w/w sucrose + 2.5 % w/w trehalose. The results are reported as remaining activity (%), relative to an LDH formulation containing only 5 mM potassium phosphate buffer, without any stabilizer.

therefore help lyo-professionals to design an appropriate formulation for biopharmaceuticals. Moreover, the proposed approach, which combines simulations and experiments, allows a knowledge-driven selection of the formulation, guided by real understanding of the stabilizing mechanisms at molecular level. In the case of the sucrose-trehalose formulation (formulation C), for instance, the improved stability may result from the combination of a good cryoprotectant, such as sucrose, with an effective lyoprotectant, such as trehalose.

2.2 How do Buffers and Sugars Interact?

In the previous section, the mechanisms of protein stabilization by different excipients have been investigated. However, the presence of buffers was not taken into account in that study.

Buffers are required in typical pharmaceutical formulations to stabilize the pH, and it is widely known that the choice of the buffer can sometimes be critical. For example, freezing a buffered protein solution may selectively crystallize one buffering species, causing pH changes. This happens, for instance, with sodium phosphate and potassium phosphate [25–28].

Although buffers are mainly used to control the pH of the formulation, they show other important features that may contribute to the stability of biopharmaceuticals [144]. These features include the ability to affect both conformational

and colloidal stability, and a factor to be considered in this framework is the potential interaction between buffers and the other components of the formulation. It is widely known that the presence of buffers can lead to a very different glass transition behavior of the surrounding matrix [145]. The glass transition temperature, T_g, which marks the transition between a rubbery (liquid-like) and glassy (solid-like) state, is extremely important in freeze drying, and biopreservation in general. A formulation with higher T_g generally shows lower mobility, and is thus preferred for the stabilization of unstable molecules. The glass transition behavior is determined by the formulation, and, in this regard, Weng and Elliott [145] found that the addition of HPO_4^{2-} to trehalose resulted in a higher T_g, while the opposite was observed for $H_2PO_4^-$. They explained this observation showing that HPO_4^{2-} strengthened the hydrogen-bonding network between trehalose molecules, while $H_2PO_4^-$ had the opposite effect. For sucrose as excipient the results were different, and in the case of either $H_2PO_4^-$ or HPO_4^{2-}, the hydrogen-bonding network was found to be strengthened.

However, while the aspects related to the control of pH and the glass transition behavior of the formulation have been clarified already, the effect of buffers on the mechanisms of protein stabilization by different stabilizers was not addressed. In this section, molecular dynamics will therefore be used to investigate the effect of buffers on the ability of sucrose and trehalose to protect a model protein, namely, the human growth hormone (hGH), from stresses due to the freeze drying process. The phosphate and citrate buffers were chosen for this investigation, as they are among the most widely used buffers in typical pharmaceutical formulations to be freeze-dried. It will be shown that buffers may break the ordered structure formed by excipients, thus promoting preferential exclusion and penalizing water replacement and water entrapment. Therefore, although buffering capacity is an important criterion, this is not the only aspect to be taken into account when selecting a buffer species for a pharmaceutical formulation.

2.2.1 Simulation Approach

The effect of phosphate and citrate buffers on the mechanisms of hGH stabilization by sucrose and trehalose was investigated. The human growth hormone topology was obtained from the RCSB Protein Data Bank (PDB 3hHR [120]), while the excipients topology files were obtained from the ATB database [122]. Simulation details are listed in Table 2.10.
All simulations were carried out using a cubic box with periodic boundary conditions. The simulation box was modeled with the Gromos 54A7 force field [123], using explicit SPC/E water (see Appendix A) [124], while the long-range electrostatics were calculated by the Particle Mesh Ewald technique (PME) [126]. During the simulations, the coordinates and velocities were saved every 10 ps, and a Lennard-Jones cut-off of 1.4 nm was used for all of the simulations.

Table 2.10: Details of simulations performed to evaluate the effect of buffers on protein stabilization mechanisms by sugars (numeration continues from Table 2.1).

#	Protein Mol. #	Exc. Type	Exc. Conc. wt %	Buffer	Buffer/Exc. Molar Ration	Box Dim. nm
9a 9b	1	sucrose trehalose	40	-	-	9x9x9
10a 10b	1	sucrose trehalose	40	$H_2PO_4^-$	0.6	9x9x9
11a 11b	1	sucrose trehalose	40	HPO_4^{2-}	0.6	9x9x9
12a 12b	1	sucrose trehalose	40	citrate	0.6	9x9x9
13a 13b	1	sucrose trehalose	40	citric acid	0.6	9x9x9
14a 14b	-	sucrose trehalose	40	citrate	0.6	4x4x4
15a 15b	-	sucrose trehalose	40	citric acid	0.6	4x4x4

In simulations 10 and 11, the two different components of the phosphate buffer, namely $H_2PO_4^-$ and HPO_4^{2-}, were analyzed separately. In fact, as evidenced by Weng and Elliott [145], $H_2PO_4^-$ and HPO_4^{2-} can interact differently with the other components of the formulation. The same approach was used for the citrate buffer, simulations 12 and 13. In any case, the negative charges of ion species were neutralized using Na^+ ions. Simulation 1 was used for comparison.

Simulations 14 and 15 were aimed to analyze the same formulations used in simulations 12 and 13, but this time no hGH molecules were introduced in the simulation box, which could therefore be smaller. These simulations were used as a control, to support the results obtained for the citrate buffer in the presence of a protein molecule. Analogous control simulations were not performed for the phosphate buffer, because they had already been carried out in a previous study [145].

For all simulations, the initial structures were energy minimized using the steepest descent algorithm, then equilibrated at 272 K for 1 ns using the velocity rescaling thermostat [125]. The resulting configuration was equilibrated at 1 bar for 10 ns with the Parrinello-Rahman barostat [127]. The final simulations were performed for 150 ns using a time step of 1 fs.

Simulations 9 to 15 were carried out for both sucrose and trehalose as excipients. The simulated condition corresponds to the environment encountered during

freezing, when the product temperature is equal to the freezing equilibrium temperature and cryoconcentration occurs. The buffer to protein ratio in the initial liquid formulations used for freeze drying is usually much lower than in the simulations here performed, but it anyway dramatically increases during cryoconcentration and the subsequent drying process, as a result of water removal. Here, it was decided to use a high concentration of buffers in order to produce a more visible and unequivocal effect on the calculated properties, clearly distinguishable from statistical fluctuations.

Hydrogen-Bond Lifetime

The hydrogen-bond lifetime τ was calculated from the average over all autocorrelation functions of the existence functions (either 0 or 1) of all hydrogen-bonds, defined as

$$C(\tilde{\tau}) = \langle s_i(t) s_i(t+\tilde{\tau}) \rangle \tag{2.9}$$

with $s_i(t) = \{0,1\}$ for hydrogen-bond i at time t. The integral of $C(\tilde{\tau})$ gives an estimate of the hydrogen-bond lifetime τ [146]

$$\tau = \int_0^\infty C(\tilde{\tau}) d\tilde{\tau} \tag{2.10}$$

The preferential exclusion, water replacement and water entrapment theories were also evaluated, as previouly described.

2.2.2 Phosphate Buffer

The Phosphate Buffer Affects the Hydrogen Bonding Network Formed by Excipients

As evidenced by Weng and Elliot [145], $H_2PO_4^-$ and HPO_4^{2-} behave differently, especially as regards their ability to hydrogen-bond with trehalose or sucrose. They observed that HPO_4^{2-} strengthened the hydrogen-bonding network between trehalose molecules, while $H_2PO_4^-$ had the opposite effect. Moreover, HPO_4^{2-} was found to aggregate into smaller clusters than $H_2PO_4^-$. Because of this, adding HPO_4^{2-} to trehalose resulted in a higher T_g with respect to pure trehalose. More specifically, Weng and Elliot [145] showed that the addition of Na_2HPO_4 to trehalose, until a 1:1 molar ratio, increased the glass transition temperature to 401.5 K, compared to 389.1 K for trehalose alone. By contrast, $H_2PO_4^-$ acted as a plasticizer, lowering the T_g of a 1:1 (molar ratio) mixture to 361.7 K, i.e., almost 30 K below the value for trehalose alone.

For sucrose as excipient the results were different, and in the case of either $H_2PO_4^-$ or HPO_4^{2-}, the hydrogen-bonding network was found to be strengthened

by their addition. Accordingly, Ohtake *et al.* [147] observed that the T_g of sucrose/sodium phosphate mixtures was always higher than for sucrose alone, at any ratio and pH value. As will be shown, these results were confirmed by our MD simulations.

In Table 2.11 the number N of excipient-excipient and excipient-buffer hydrogen bonds, and the hydrogen bond lifetime τ, which is an indicator of hydrogen-bond strength, are listed.

Table 2.11: Results of hydrogen-bond analysis for simulations performed using the phosphate buffer

#	Exc.	Buffer	N/τ, ps		n^b	n_c^b
			Exc-Exc	Ex.-Buffer		
9a	sucrose	none	1195/6.582	-	-	-
10a		$H_2PO_4^-$	910/6.823	890/6.204	0.195	9
11a		HPO_4^{2-}	810/6.263	980/8.354	0.017	7
9b	trehalose	none	1115/7.367	-	-	-
10b		$H_2PO_4^-$	850/6.751	855/5.246	0.275	9
11b		HPO_4^{2-}	745/5.428	990/8.980	0.004	7

Moreover, the number n_b of hydrogen bonds per molecule between buffer species and the average number n_c^b of buffer molecules forming a cluster in each simulation are also shown in Table 2.11. In this context, a molecule was considered part of a cluster when its distance to any element of the cluster was less than 0.35 nm. The value of 0.35 nm chosen as cut-off for the definition of cluster formation coincides with the typical hydrogen bond length. This means that the clusters identified by this analysis should mainly include molecules linked by hydrogen bonds.

From this analysis, it is possible to notice that the hydrogen bonds between trehalose molecules were stronger than those between sucrose molecules, as evident comparing the hydrogen bond lifetime τ for simulation 9 in case of sucrose (6.582 ps) and trehalose (7.367 ps). This difference between sucrose and trehalose is crucial because, as we previously discussed, determines their different behavior as protectants. While trehalose exerts its protective action on the protein structure mainly because of its strong hydrogen-bonding network, sucrose acts mainly by preferential exclusion.

Analyzing the results in Table 2.11, it is also possible to conclude that $H_2PO_4^-$ formed hydrogen-bonds with excipient molecules that were weaker than the excipient-excipient hydrogen-bonds. By contrast, HPO_4^{2-}-excipient hydrogen bonds were stronger than those between excipient molecules, as evident from analysis of the hydrogen-bond lifetime. This is true for both the sucrose-based and the trehalose-based formulation. $H_2PO_4^-$ also formed bigger clusters than HPO_4^{2-}, and characterized by a dense network of hydrogen bonds. HPO_4^{2-} ions were less likely to form

clusters and, thus, they tended to break the structure formed by excipients more than $H_2PO_4^-$. This emerged from the smaller number of excipient-excipient hydrogen bonds in simulation boxes containing HPO_4^{2-} rather than in those containing $H_2PO_4^-$. In spite of this, HPO_4^{2-} formed an even more structured matrix, with a large number of strong hydrogen-bonds between excipient molecules and HPO_4^{2-} ions. This is true for both sucrose and trehalose as excipients. These results are in line with the work by Weng and Elliot [145], where the same simulations were performed in absence of protein molecules. This suggests that the results obtained for the sugars-phosphate buffer interactions are not influenced by the presence of hGH.

Excipients Behavior in Presence of the Phosphate Buffer

From the previous analysis, it seems that different buffer species may have opposite behavior on sucrose or trehalose-based formulations, at least as regards the hydrogen-bonding network formed by the mixture. This is extremely important, because, according to the results previously discussed, the protection mechanisms that are generally proposed to explain protein stabilization in the dried state, i.e., vitrification, water replacement and water entrapment, may be directly related to the formation of a compact hydrogen bonding network. In the following, the effect of the phosphate buffer on the mechanisms of protein stabilization by excipients will be further investigated. In order to do this, preferential exclusion, water replacement, cage effect and water entrapment will be considered.

In Figures 2.14, 2.15, and 2.16 the preferential exclusion and water replacement parameters β and χ are shown.

As can be seen in Figure 2.14, sucrose was in general more excluded than trehalose from the protein surface. Moreover, the preferential exclusion of both sucrose and trehalose was generally increased upon addition of buffer species. This effect was particularly remarkable for trehalose. HPO_4^{2-} had a very positive effect on preferential exclusion, leading to the smallest values of β in the case of both sucrose and trehalose. These results were also confirmed by the Π parameter values listed in Table 2.12.

It is also interesting to compare these results with the previous findings about the effect of buffers on the hydrogen bonding network formed by excipients. In general, buffers that broke excipient structure the most, such as HPO_4^{2-}, had the most positive effect on preferential exclusion.

In Figure 2.15, the preferential exclusion of the buffer components from the protein surface was also evaluated. Preferential exclusion is rarely seen for buffers, mainly because buffers concentration in typical formulations is generally too low. However, at the high concentration simulated in this work, some of the buffer species under investigation showed preferential exclusion from the protein surface. This was particularly true for HPO_4^{2-}, as evident also from the Π values listed in

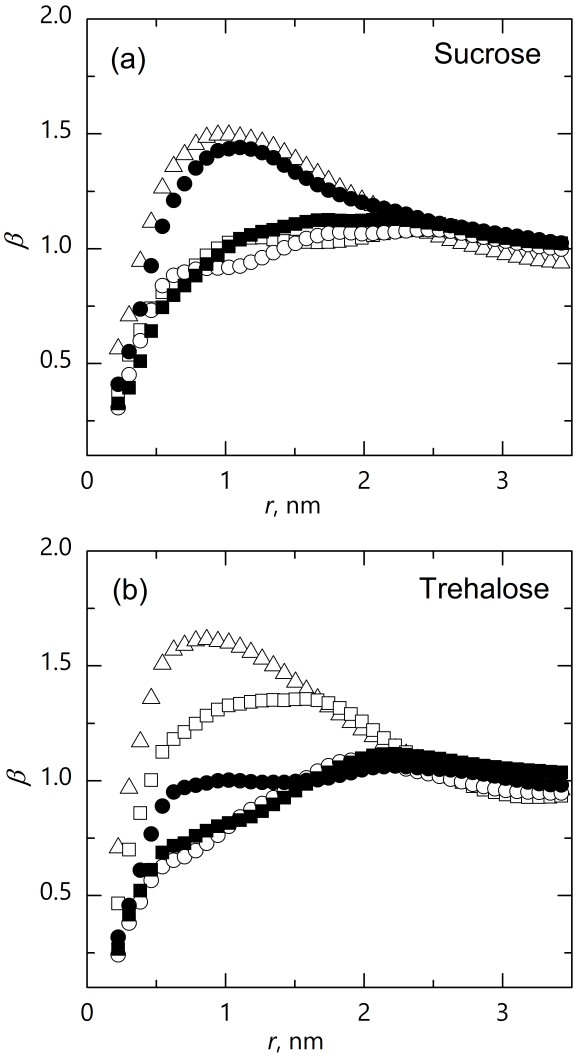

Figure 2.14: Plot of the preferential exclusion parameter β for sucrose (a), and trehalose (b). △: no buffer, □: $H_2PO_4^-$, ○: HPO_4^{2-}, ■: citrate,●: citric acid

Table 2.12.

Figure 2.15 also shows that the preferential exclusion of HPO_4^{2-} was less important in the presence of sucrose as excipient. This may be related to the competition between sucrose molecules and buffer components for being preferentially excluded from the protein surface.

Moving now to water replacement, Figure 2.16 shows that trehalose formed more hydrogen-bonds with hGH than sucrose. This result is in line with the observations discussed in the previous section. The addition of buffers had a negative effect on water replacement. More specifically, buffers that broke excipient structure the

Table 2.12: Π parameter values for simulations listed in Table 2.10

#	Exc.	Buffer	Π for the sugar	Π for the buffer
9a	sucrose	none	1.215	-
10a		$H_2PO_4^-$	0.800	0.856
11a		HPO_4^{2-}	0.769	0.758
9b	trehalose	none	1.407	-
10b		$H_2PO_4^-$	1.073	1.227
11b		HPO_4^{2-}	0.600	0.407

most, such as HPO_4^{2-}, had the worst effect on water replacement.

From the viewpoint of the water entrapment scenario, the same analysis perfomed in the previous section was repeated here in the presence of buffer species, and the results are listed in Table 2.13.

Table 2.13: Results of the water entrapment analysis, for simulations listed in Table 2.10

#	Exc.	Buffer	Cage Ratio	$c_{I_{min}}^{I_{max}}$ m^3/m^3	δ, %	$D(< I_{max})$ 10^{-5} cm^2/s	$[I_{min}, I_{max}]$ [nm, nm]
9a	sucrose	none	1.42	0.278	66	0.2534	[2.2, 4.1]
10a		$H_2PO_4^-$	1.02	0.206	64	0.1995	[2.7, 4.1]
11a		HPO_4^{2-}	1.00	0.206	89	0.1620	[2.6, 4.1]
9b	trehalose	none	1.34	0.252	70	0.2612	[2.3, 4.1]
10b		$H_2PO_4^-$	1.33	0.271	67	0.1687	[2.2, 4.1]
11b		HPO_4^{2-}	1.06	0.219	83	0.1529	[2.6, 4.1]

From the results obtained, it seems that the phosphate buffer reduced the cage effect. Both the cage ratio and $c_{I_{min}}^{I_{max}}$ were generally smaller in the presence of the buffer species. This suggests that the accumulation of excipient molecules around the protein surface decreased upon addition of buffers. According to the theory previously discussed that the water entrapment scenario is directly related to the existence of a highly structured matrix, the reduction in cage effect caused by buffers could again be related to the disruption of the hydrogen-bonding network formed by the excipient molecules. However, buffer components also reduced $D(< I_{max})$, meaning that water molecules were further slowed down upon addition of the phosphate buffer. This can be explained considering that buffer species are bigger and less mobile than water, thus hindering the motion of nearby molecules. The reduction in $D(< I_{max})$ was particularly pronounced for those buffer components, such as HPO_4^{2-}, that formed a great number of hydrogen bonds with the excipient molecules.

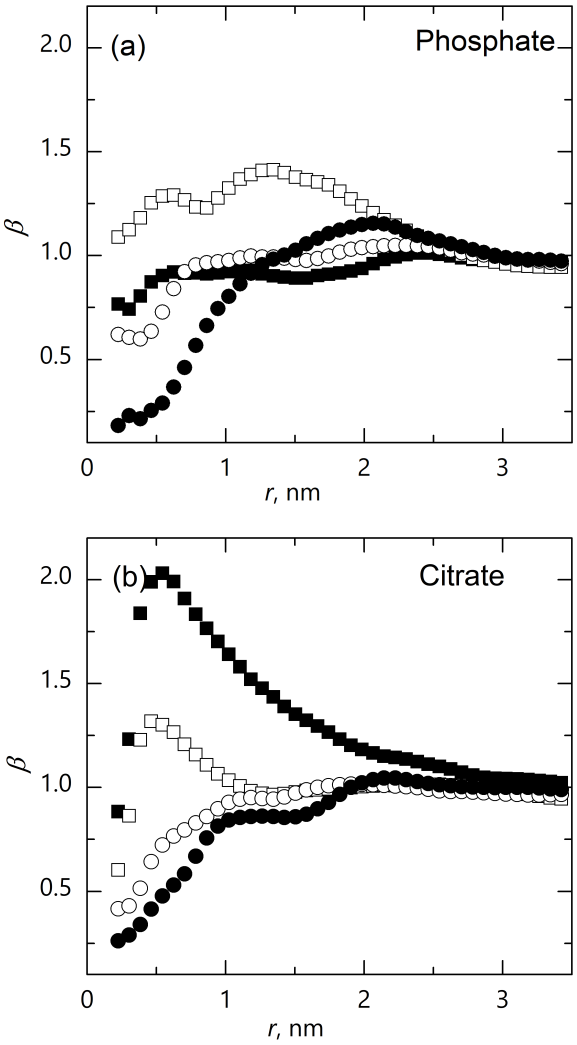

Figure 2.15: Plot of the preferential exclusion parameter β for phosphate (a), and citrate (b) buffer species. ●: HPO_4^{2-}/citrate in presence of trehalose, ○: HPO_4^{2-}/citrate in presence of sucrose, □: $H_2PO_4^-$/citric acid in presence of trehalose, ■: $H_2PO_4^-$/citric acid in presence of sucrose.

Finally, the effect of the surface charge properties of hGH on the arrangement of both sugar and buffer molecules was also investigated, as it was done already in the previous section.

In order to do this, the hydrophobic and hydrophilic surface areas exposed by each amino acid were evaluated. We then defined as hydrophobic residues those whose hydrophobic surface area was larger than the hydrophilic one, and vice versa. Finally, the average minimum distance between excipient/buffer molecules and the

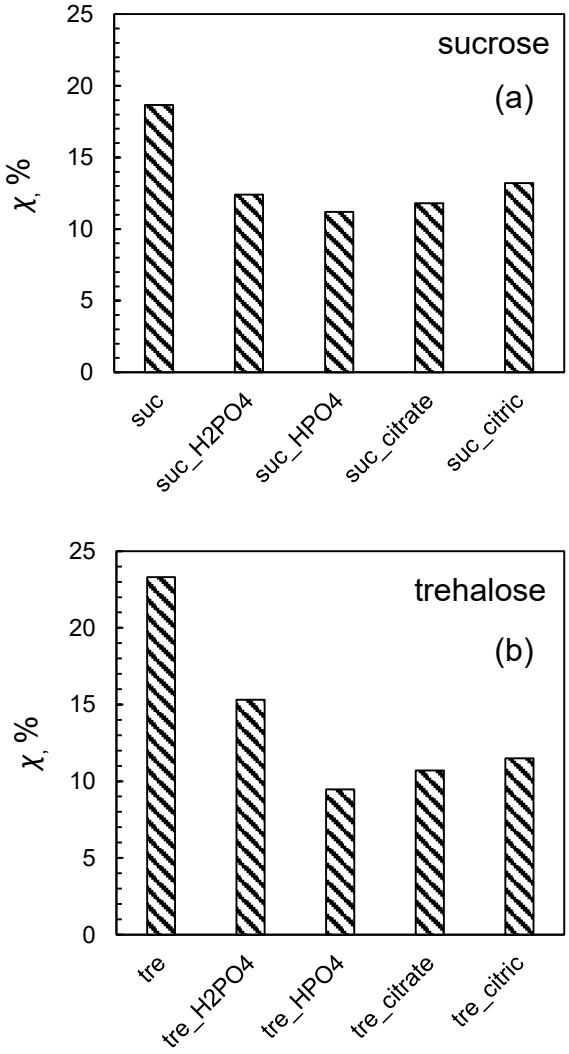

Figure 2.16: Plot of the water replacement parameter χ for sucrose (a), and trehalose (b).

hydrophobic, d_{NP}, or hydrophilic, d_P, patches on the protein surface was computed. The results obtained are listed in Table 2.14.
Both sugars and buffer species approached the hydrophilic patches of the protein the most ($d_P < d_{NP}$). Moreover, sugars were closer to the protein surface than the phosphate buffer species. It can finally be noted that the addition of buffers significantly distanced trehalose from the protein surface, while this effect was less evident in the case of sucrose. This observation is in line with the significantly

Table 2.14: Average distance between sugars or phosphate buffer molecules and the hydrophobic, d_{NP}, or hydrophilic, d_P, patches of the protein

#	Exc.	Buffer	Distance from hGH surface			
			d_{NP}, nm		d_P, nm	
			sugar	buffer	sugar	buffer
9a	sucrose	none	0.40	-	0.26	-
10a		$H_2PO_4^-$	0.41	0.63	0.28	0.47
11a		HPO_4^{2-}	0.40	0.68	0.26	0.52
9b	trehalose	none	0.34	-	0.21	-
10b		$H_2PO_4^-$	0.37	0.62	0.26	0.49
11b		HPO_4^{2-}	0.41	0.77	0.30	0.64

reduced efficiency of trehalose in providing the water replacement mechanism (Figure 2.16), and its increased preferential exclusion (Figure 2.14), upon addition of buffers.

2.2.3 Citrate Buffer

The Citrate Buffer Affects the Hydrogen Bonding Network Formed by Excipients

The same analysis performed for the phosphate buffer was repeated for the citrate buffer. The results of the hydrogen-bonds analyses are shown in Table 2.15. The analysis performed makes it possible to conclude that citric acid tended to aggregate significantly more than citrate, with clusters formed by about 35-45 or 14 molecules in the case of citric acid and citrate, respectively. However, in the case of citric acid, the clusters were substantially less structured, as evident from the much lower value of n_b in the case of citric acid, ≈0.2-0.3, than in the case of citrate ions, ≈0.6. This result suggests that citrate ions tend to form more compact and structured matrix if compared to citric acid, in line with the experimental observation that the T_g for citrate (342 K for sodium citrate and 388 K for disodium citrate) is remarkably higher than for citric acid (284 K) [148].
Moreover, it is possible to notice that citrate ions broke the ordered structure formed by the excipients much more than citric acid. The number of excipient-excipient hydrogen bonds in simulation boxes containing citrate was much smaller than in simulation boxes containing citric acid. However, in the case of sucrose as excipient, citrate ions formed an even more structured matrix, with a large number of strong sucrose-citrate hydrogen bonds. This was true both in presence ($N \approx 1290$, τ =7.136 ps, as in sim. 12a) and in absence of the protein ($N \approx 158$, τ =6.485 ps, as in sim. 14a). In the case of trehalose, the presence of citrate disrupted the

Table 2.15: Results of hydrogen-bond analysis for simulations performed using the citrate buffer

#	Exc.	Buffer	N/τ, ps		n^b	n_c^b
			Exc-Exc	Ex.-Buffer		
9a	sucrose	none	1195/6.582	-	-	-
12a		citrate	620/5.739	1290/7.136	0.619	14
13a		citric acid	1025/6.430	515/7.833	0.233	37
14a		citrate	57/4.979	158/6.485	0.583	14
15a		citric acid	105/5.947	49/7.588	0.292	37
9b	trehalose	none	1115/7.367	-	-	-
12b		citrate	630/6.428	1165/6.100	0.602	14
13b		citric acid	920/6.352	520/7.305	0.212	36
14b		citrate	84/5.128	123/4.928	0.625	14
15b		citric acid	90/4.357	51/4.623	0.292	43

hydrogen-bonding network formed by the excipient but, even though the number of excipient-citrate hydrogen bonds was very high, these bonds were weaker (τ = 6.100 ps, sim. 12b) than those formed by trehalose molecules in absence of buffer (τ = 7.367 ps, sim. 9b).

As regards citric acid, it formed very strong hydrogen bonds with both sucrose and trehalose, as confirmed by the very high values of τ, but, due to its high tendency to form big clusters, the total number of citric acid-excipient hydrogen bonds was much lower than in the case of citrate. Again, these results were true both in presence (sim. 13a,b) and absence (sim. 15a,b) of the protein. According to this analysis, the addition of citrate to sucrose-based formulations should strengthen the hydrogen-bonding network of the mixture. This is confirmed by the experimental results by Kets *et al.* [149], who showed that the addition of sodium citrate to a sucrose solution resulted in a glass transition temperature that was higher than for pure sucrose. Moreover, in line with our analysis, they experimentally observed a remarkable average strength of hydrogen bonding in the sucrose/citrate glass. By contrast, according to the MD simulations, citric acid should have the opposite effect.

In the case of trehalose, both citrate ions and citric acid molecules should weaken the hydrogen-bonding network formed by the excipient. As for the phosphate buffer, the results obtained for the interaction between sugars and the citrate buffer should not be influenced by the presence of the protein. This is confirmed by the similar results obtained in simulations 12-13 and 14-15, respectively.

Excipients Behavior in Presence of the Citrate Buffer

The same analyes of protein-excipient interaction, already performed for the phosphate buffer, was here repeated for the citrate buffer.

As can be seen in Figure 2.14, the preferential exclusion of both sucrose and trehalose was generally increased upon addition of the citrate buffer, with citrate ions having a very positive effect on preferential exclusion. The only exception was the mixture sucrose-citric acid, whose preferential exclusion was similar to that of sucrose alone. Again, it can be observed that buffers that broke excipient structure the most, such as citrate, had the most positive effect on preferential exclusion. These results were also confirmed by the Π parameter values listed in Table 2.16.

Table 2.16: Π parameter values for simulations listed in Table 2.10

#	Exc.	Buffer	Π for the sugar	Π for the buffer
9a	sucrose	none	1.215	-
12a		citrate	0.736	0.698
13a		citric acid	1.088	1.923
9b	trehalose	none	1.407	-
12b		citrate	0.656	0.524
13b		citric acid	0.829	1.130

Figure 2.15 shows that citrate also showed significant preferential exclusion from the protein surface. However, this preferential exclusion became less important in the presence of sucrose as excipient. Again, this may be related to the competition between sucrose molecules and buffer components for being preferentially excluded from the protein surface. This result is in line with the experimental observation that citrate could promote protein conformational stability as a result of preferential exclusion, especially at high concentration [150–152].

As regards water replacement, Figure 2.16 shows that the addition of citrate buffer, and especially of citrate ions, had a negative effect on water replacement. Furthermore, the results for the water entrapment scenario are listed in Table 2.17. Similarly to the phosphate species, also the citrate buffer reduced the cage effect. However, citrate ions also reduced $D(< I_{max})$, meaning that water molecules were further slowed down upon addition of the citrate buffer. In line with previous observations, the reduction in $D(< I_{max})$ was particularly pronounced for those buffer components, such as citrate, that formed a great number of hydrogen bonds with the excipient molecules.

Finally, the effect of the surface charge properties of hGH on the arrangement of sugar and citrate buffer molecules was also investigated, as it was done already in the previous section. The results obtained are listed in Table 2.18.

Table 2.17: Results of the water entrapment analysis, for simulations listed in Table 2.10

#	Exc.	Buffer	Cage Ratio	$c_{I_{min}}^{I_{max}}$ m^3/m^3	δ, %	$D(< I_{max})$ 10^{-5} cm^2/s	$[I_{min}, I_{max}]$ [nm, nm]
9a	sucrose	none	1.42	0.278	66	0.2534	[2.2, 4.1]
12a		citrate	1.02	0.213	80	0.0717	[2.7, 4.2]
13a		citric acid	1.27	0.249	69	0.1482	[2.3, 4.1]
9b	trehalose	none	1.34	0.252	70	0.2612	[2.3, 4.1]
12b		citrate	1.06	0.223	83	0.0689	[2.7, 4.1]
13b		citric acid	1.10	0.220	83	0.1610	[2.5, 4.1]

Table 2.18: Average distance between sugars or citrate buffer molecules and the hydrophobic, d_{NP}, or hydrophilic, d_P, patches of the protein

#	Exc.	Buffer	Distance from hGH surface			
			d_{NP}, nm		d_P, nm	
			sugar	buffer	sugar	buffer
9a	sucrose	none	0.40	-	0.26	-
12a		citrate	0.41	0.60	0.28	0.41
13a		citric acid	0.45	0.40	0.28	0.25
9b	trehalose	none	0.34	-	0.21	-
12b		citrate	0.40	0.64	0.26	0.48
13b		citric acid	0.43	0.46	0.29	0.33

As previously observed for phosphate, sugars were generally closer to the protein surface than citrate buffer species, with the only exception of citric acid. Citric acid clustered very close to the protein surface, especially in the case of the sucrose-based formulation. This result is in line with the preferential interaction observed between citric acid and hGH in Figure 2.15.

2.2.4 Conclusions

Overall, these results suggest that buffers have important properties, other than the control of pH, that can affect protein preservation. Examples of these properties were also observed experimentally. For instance, both the citrate and the phosphate buffer increased the stability of green fluorescent protein against denaturation at high temperature [153], and they had a positive effect on ovotransferrin [154], as well.

However, at the same time, these buffers sometimes caused protein destabilization. For example, it has been reported that phosphate destabilized lysozyme [155], and that the rate of aggregation of monoclonal antibodies was significantly increased in the presence of citrate [156]. It is also known that buffers can interact with other typical components of the formulation, such as sugars, modifying their behaviour. This was observed both during freezing and in the dried state. During freezing the $T_g^{'}$, i.e., the glass transition temperature of the maximally freeze-concentrated state, changed upon addition of phosphate species to sugars solutions [157, 158]. In the dried state, the effect of different phosphate species on the glass transition behavior of trehalose and sucrose [145, 147], and the increase in T_g observed upon addition of citrate to sucrose [149], have been mentioned already. Moreover, the citrate buffer was shown to interact with arginine in another study, producing an amorphous glass which increased protein stability [159, 160]. Also, the addition of phosphate species to sugars was found to improve the preservation of lyophilized liposomes [161].

These experimental data, combined with the MD results, suggest that the interaction of buffers with other formulation components could affect the overall protein stability. In the case of sugars, MD simulations suggest that this interaction may be related to the creation of a sugar-buffer hydrogen bonding network, which disrupts, and substitutes, the original matrix formed between sugar molecules. This affects the efficiency of sugars, such as the commonly used sucrose and trehalose, in providing protein stabilization. It should also be taken into account that the different components of a buffer, according to the MD investigation, may have a different behavior, and therefore the effect on the sugar matrix should be strongly pH-dependent.

List of Symbols

a_1, a_2	-	coefficients of Genotelle equation (Equation 2.6)
b_1, b_2	-	coefficients of Genotelle equation (Equation 2.6)
$c_{I_{min}}^{I_{max}}$	-	volume by volume concentration
C	-	autocorrelation function of the existence functions of hydrogen bonds
d	m	distance from the protein
d_{NP}	m	distance from the hydrophobic patches
d_P	m	distance from the hydrophilic patches
D	$m^2\ s^{-1}$	diffusion coefficient
g	-	radial distribution function
I	-	interval of increased excipient concentration
I_{max}	m	upper limit of interval I
I_{min}	m	lower limit of interval I

k_B	J K^{-1}	Boltzmann constant
n	-	exponent of Genotelle equation (Equation 2.6)
n_b	-	number of hydrogen bonds per molecule between buffer species
n_c^b	-	average number of buffer molecules forming a cluster
n_{exc}	-	number of excipient molecules
n_w	-	number of water molecules
N	-	number of hydrogen bonds
r	m	radial distance
r_0	m	lower integral limit for parameter Π
r_{max}	m	upper integral limit for parameter Π
R_g	m	radius of gyration
s	-	existence function of hydrogen bonds
t	s	time
T	K	temperature
T_g	K	glass transition temperature of amorphous solids
T_g'	K	glass transition temperature for maximally freeze-concentrated solutions
$\langle u^2 \rangle$	m^2	mean square displacement
V	m^3	volume
x	-	mole fraction

Greek Letters

β	-	preferential exclusion parameter
γ	-	hydrogen bonding propensity parameter
δ	-	ratio of diffusion coefficients
μ	Pa s	dynamic viscosity of solution
μ_w	Pa s	dynamic viscosity of water
Π	-	integral preferential exclusion parameter
ρ	m^{-3}	numerical density
τ	s	hydrogen bond lifetime
τ_α	s	α-relaxation time
ϕ	-	reduced temperature
χ	-	water replacement parameter

Chapter 3

Role of Surfactants: Evidence of an Orientation-Dependent Mechanism

The aim of this chapter is to provide insight into the molecular mechanisms underlying protein stabilization by surfactants, using Molecular Dynamics.

The human growth hormone (hGH, already introduced in Chapter 2) and the GB1 hairpin (a 16-residue peptide) are used as model proteins. hGH has been selected because it was the subject of a wide experimental investigation [101–108], also in the field of protein-surfactant interactions [49]. DeFelippis *et al.* [108] hypothesized that hGH forms a molten globule intermediate during its equilibrium denaturation. Molten globule intermediates are compact, globular molecules with substantial secondary structure but no specific tertiary structure, and with more non-polar residues exposed to the solvent than the native state [162]. The molten globule intermediate of hGH was found to be prone to self-associate [108], and Bam *et al.* [49] observed that this self-association is reduced upon formation of a Tween-rhGH complex.

The GB1 peptide (shown in Figure 3.1a) corresponds to residues 41-56 of the GB1 protein, and is also well-studied, both experimentally and computationally [163–175].

The polysorbates Tween 20 and Tween 80 were selected as model surfactants. Both have four hydrophilic poly(ethylene oxide) head groups (see Figure 3.1b,c), named X, Y, Z and W heads, which contain x, y, z and w number of ethylene oxide units, respectively. The hydrophobic tail consists of lauric acid in the case of Tween 20, and oleic acid for Tween 80. The sum of x, y, z and w averages 20, and a wide range of isomers is possible. For this work, the isomer having x=y=z=w=5 was chosen as reference molecule. Tween 80 is more surface-active, and has a lower critical micelle concentration (CMC) than Tween 20, while Tween 20 generally shows stronger interactions with proteins in the bulk [52, 176, 177]. For this reason,

Tween 20 will be used for simulations of hGH in the bulk, while Tween 80 will be selected as model surfactants when studying the interaction of GB1 with different surfaces [1].

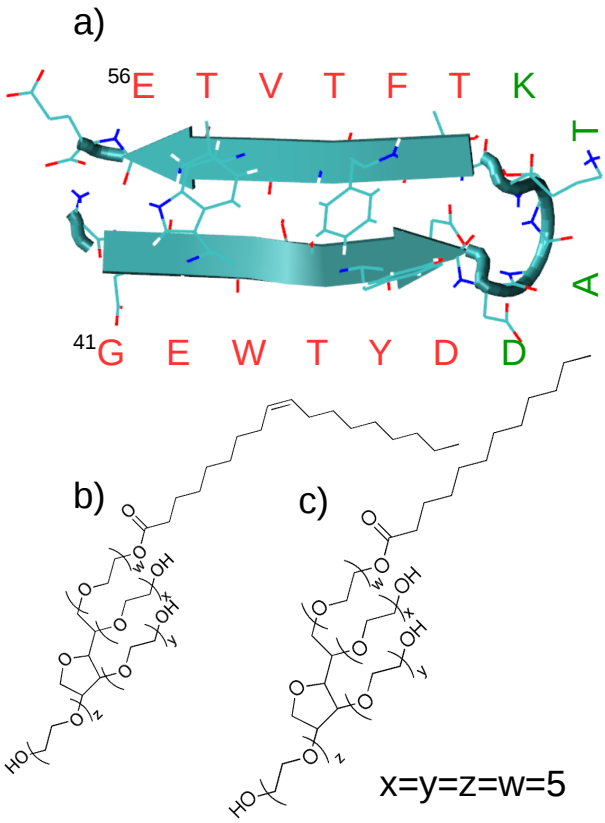

Figure 3.1: Illustration of a) the GB1 hairpin and b) the Tween 80 or c) Tween 20 molecule. In panel a), the GB1 sequence is also shown, with the β strand residues in red, and those in green being part of the turn.

3.1 Mechanisms and Unresolved Issues in Protein Stabilization by Surfactants: A Literature Overview

A major cause of biopharmaceuticals instability is related to the amphiphilic, surface-active nature of protein molecules. Because of this, phenomena such as

[1]Part of this chapter is based on already published papers [178, 179]

interface-solution partitioning and surface-induced unfolding are very common, and could lead to formation of aggregates. This is particularly true in *in vitro* environments, where surface-induced protein adsorption and denaturation become a major concern. Surfaces are present in most *in vitro* experiments: the glass container of a test tube, inorganic or organic surfaces used in biosensors, and even the air-water interface. Surface-induced denaturation can have serious consequences in the field of pharmacology, where loss of protein native structure leads to loss of drug activity.

Drawing from the cell's defense mechanism against protein denaturation, one can mitigate protein denaturation *in vitro* through addition of protective osmolytes, or surfactant molecules. Surfactant molecules are a particularly attractive means, as these molecules are readily synthesized and have been shown to counter surface-induced denaturation in a number of studies. Surfactants are amphipathic molecules, with a hydrophobic tail and an anionic, cationic, or non-ionic hydrophilic head. Non-ionic surfactants, and, among these, the polysorbates Tween 20 or Tween 80, are commonly used in pharmaceutical studies to stabilize protein-based drugs in experiments where surface-induced denaturation is an issue.

The polysorbates consist of fatty acid esters of polyethylene sorbitan, and their amphiphilic nature drives both their surface activity, and their tendency to interact with proteins. The hydrophobic nature of the polysorbates is provided by a fatty acid side-chain, while the hydrophilic nature is given by the ethylene oxide subunits. At the critical micelle concentration (CMC), the polysorbates form micelles, while at lower concentration they are predominantly monomeric.

While surfactants are used routinely, the mechanism by which they prevent proteins denaturation is poorly understood, and several possible mechanisms have been proposed [47,48]. For instance, it has been hypothesized that surfactants may reduce the driving force of protein adsorption and/or aggregation at interfaces. Surfactants should preferentially locate at interfaces, thus precluding protein adsorption. This stabilizing effect is concentration-dependent, and is strongly related to the CMC of the surfactant.

Other stabilization mechanisms were also proposed, such as assistance in protein refolding, which may inhibit protein-protein interactions. Bam *et al.* [49] suggested that surfactants, such as the polysorbates, may act like molecular chaperones. Molecular chaperones are a class of molecules that catalyze correct folding and prevent undesired aggregation of proteins, while not altering the native conformation. Moreover, it was also suggested that interaction of surfactants with the native state of the protein could minimize aggregation by increasing the free energy of protein unfolding [49,176]. More recently, Deechongkit *et al.* [52] observed that polysorbates had an effect on the α-helix content of darbepoetin alpha, even below the CMC. This was in contrast with previous results, where studies of recombinant human growth hormone (rhGH) [49] and anti-L-selectin antibody [53] by far-UV circular dichroism seemed to demonstrate a negligible effect of surfactant molecules on the proteins secondary structure.

It has also been hypothesized that surfactants may associate with proteins in solution, in this way stabilizing them against close approach and inhibiting aggregation. As regards this last mechanism, it has been suggested that the polysorbates should bind hydrophobic patches on the surface of proteins, thus preventing interactions [48,180]. In this case, the concentration of surfactant required to protect the protein from aggregation should not correlate with the CMC of the surfactant, but, on the contrary, the degree of protection should be maximized at the molar binding stoichiometry between the surfactant and the protein [50, 51]. Deechongkit *et al.* [52] confirmed this result, suggesting that the protein should bind to or interact with individual polysorbate molecules, rather than with micelles. They observed that the polysorbates prevented surface loss and aggregation at polysorbate concentrations below the CMC. In any case, the protein-polysorbate interactions that have been experimentally observed are generally considered weak, hydrophobic interactions, as deduced, for instance, by differential scanning calorimetry (DSC) and isothermal titration calorimetry (ITC) [52, 177]. Chou *et al.* [176] studied the saturable binding of both Tween 20 and Tween 80 with Albutropin, finding a molar binding stoichiometry (surfactant:protein) of 10:1 and 9:1, respectively. Bam *et al.* [50] studied the interactions of Tween with recombinant human growth hormone (rhGH) and recombinant human interferon-γ (rhIG) using electron paramagnetic resonance (EPR), and found that the binding stoichiometries were dependent on the hydrophobicity of the protein surface and the surfactant.

This brief overview shows that several unresolved questions remain about protein-surfactant interactions, together with some contradictory observations. Hence, the actual mechanisms at the basis of protein protection by surfactants are not yet completely understood.

3.2 Overview of the Study

In this chapter, the effect of Tween 20 on hGH aggregation will be compared to the stabilization mechanisms of typical protectants, namely, the disaccharides sucrose and trehalose. These excipients were chosen because, according to the results discussed in Chapter 2, they should correspond to a very performant cryoprotectant and lyoprotectant, respectively. The stabilizing action of these stabilizers in the bulk solution will be investigated using molecular dynamics, and an explanation for this behavior will be proposed. It will be shown that the interaction of stabilizers with the peptide chain, and in particular with the aggregation-prone regions (APRs) of the protein, may be at the basis of the observed stabilization. MD results will be presented, that seem to suggest an orientation-dependent mechanism, where the surfactant ability to act as a molecular chaperone and catalyze correct folding should be related to its orientation toward the protein surface.

The folding of the GB1 hairpin in the bulk, at different surfaces, and in the

presence and absence of Tween 80 will also be investigated. The air-, ice- and silica-water surfaces were chosen because they are relevant during the freeze-drying process, and in biotechnology settings in general. The air-water interface is the most common hydrophobic interface encountered by proteins in a laboratory setting, particularly during stirring or shaking of the solution [51, 181]. The ice-water interface is relevant for freezing and freeze-drying processes, with the surface of ice an agent for loss of protein activity [61, 182]. Silica interfaces are relevant as proteins *in vitro* are exposed to the glass surfaces of beakers, syringes (for drug delivery), and other containers. An understanding of the mechanism by which surfactants counteract the denaturation of proteins at these surfaces would therefore have practical applications in the realm of the pharmaceutical and the biotechnology industry, and in the field of freeze drying, as well.

Using MD simulations, it will be shown that the extent to which surfactants prevent denaturation is dependent on the nature of the surface. A surfactant orientation-dependent stabilization mechanism will again be proposed, in which the surfactant stabilizes the native fold by surrounding the protein with its hydrophilic heads pointing toward the protein.

3.3 Surfactants Inhibit Protein Aggregation in Bulk

3.3.1 Simulation Setup

The topology of native human growth hormone was downloaded from the RCSB Protein Data Bank (PDB 3hHR [120]). A topology for unfolded hGH was also obtained from simulations where the protein was exposed to the stresses typically encountered during freeze-drying, i.e., contact with the ice-water interface and removal of water. Both the native and the unfolded structure were used in order to study the effect of stabilizers on both states of hGH. The native and unfolded conformations (shown in Figure 3.2) differ in the amount of non-polar surface area and in secondary structure, with the unfolded conformation exposing more non-polar area ($S_{np} \approx 50$ nm^2) and having smaller α-helix content (about 45 %) than the native one ($S_{np} \approx 42$ nm^2, α-helix ≈ 68 %). The Tween 20, sucrose and trehalose topology files were obtained from the ATB database [122].
Details of the simulations performed are listed in Table 3.1. The number of cosolute molecules used in each simulation box was chosen so as to reproduce typical values of cosolute to protein ratio used in experimental work [104, 105].
All the simulations were performed at low temperature, i.e., 273 K. This temperature also allows a precautionary estimate of the interaction potential, as it has been shown that an increasing attractivity is observed upon a decrease in temperature [183]. Periodic boundary conditions were used, and the simulation box was

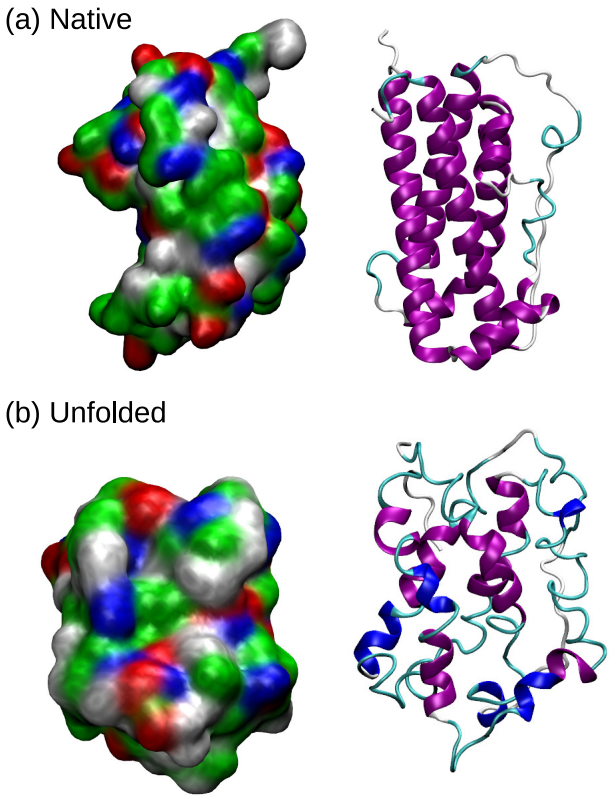

Figure 3.2: Illustration of native (top) and unfolded (bottom) hGH conformations used for this study, where both a surface (left) and a cartoon (right) representation have been used. In the surface representation, polar residues are colored in green, nonpolar residues in white, acidic residues in red and basic residues in blue. In the cartoon representation, α-helices are in purple, 3_{10}-helices in blue, turns in cyan and coils in white.

modeled with the Gromos 54A7 force field [123] as implemented in Gromacs 5.0, using explicit SPC/E water [124] (see Appendix A). Na^+ ions were added to guarantee neutrality of the system. The initial structures were energy minimized using the steepest descent algorithm, then equilibrated for 1 ns at the desired temperature using the velocity rescaling thermostat [125]. The long-range electrostatics were calculated by the Particle Mesh Ewald technique (PME) [126]. Finally, the resulting configuration was equilibrated for 20 ns at 1 bar with the Parrinello-Rahman barostat [127]. A Lennard-Jones cut-off of 1.4 nm was used for all the simulations.

Table 3.1: Details of simulations performed to evaluate the interaction potential (PMF) between hGH molecules (sim. 1-5), or to investigate the effect of different cosolutes on hGH (sim. 6-9)

#	Protein	Exc.	# Exc. Mol.	# Water Mol.	Box dim. nm	Sim. time ns
1	2 native	none	0	24601	6x6x22	375
2	2 unfolded	none	0	24431	6x6x22	25 windows 375
3	2 unfolded	Tween 20	20	23978	6x6x22	25 windows 600
4	2 unfolded	Sucrose	260	20653	6x6x22	40 windows 600
5	2 unfolded	Trehalose	260	20514	6x6x22	40 windows 600
6	1 native	Tween 20	10	19235	8x8x8	40 windows 100
7	1 unfolded	Tween 20	10	11269	8x8x8	100
8	1 unfolded	Sucrose	130	9596	8x8x8	100
9	1 unfolded	Trehalose	130	9599	8x8x8	100

3.3.2 Simulation Approach

The aggregation propensity of hGH was investigated in simulations 1-5, using umbrella sampling [184]. Simulations 1 and 2 were used to compare the aggregation behavior of native and unfolded hGH in bulk water. Simulations 3-5 were focused on unfolded hGH, and the presence of other cosolutes, i.e. Tween 20, sucrose or trehalose, was also considered.

In an umbrella sampling (US) simulation, the free energy barriers in the system under investigation can be more easily surmounted, allowing a statistically relevant exploration of the free energy landscape. This result is achieved by adding a bias $V(s)$ on selected degrees of freedom s, generally called collective variables (CVs). If the shape of the free-energy landscape were known since the beginning, it would be possible to build a bias potential that makes the biased free-energy landscape completely flat and barrier-less. However, it is difficult to have an a priori guess of the free-energy landscape. Because of this, it is advantageous to combine results from several independent trajectories, each with a bias potential in the form of a harmonic restraint,

$$V(s) = \frac{k}{2}(s - s_0)^2 \tag{3.1}$$

so that only values of s close to s_0 will be explored. In each trajectory, the exploration of the CV region is therefore limited to a small window, that overlaps with nearby windows explored within the set of independent trajectories performed (see Figure 3.3). The statistics of these independent simulations could then be combined using the weighted histogram analysis method (WHAM) [185]. In this work, the interaction potential between protein molecules (potential of mean force, PMF), was computed using this technique.

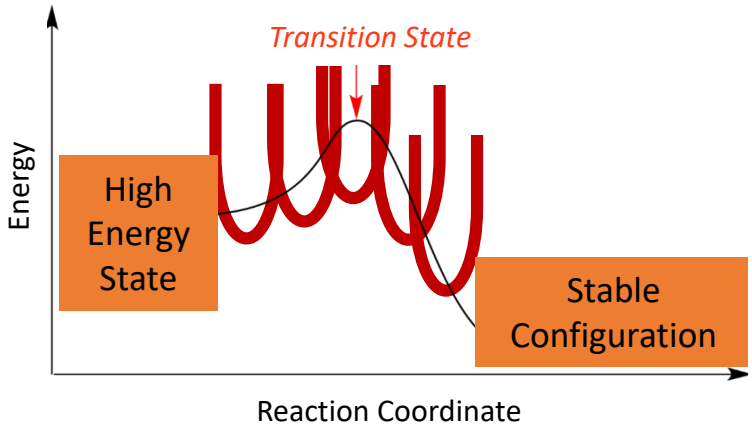

Figure 3.3: Scheme of umbrella sampling. Several independent trajectories are run along the free-energy landscape, where the exploration of the CV is limited to a small window by the application of a harmonic restraint.

For simulations 1 to 5, at the beginning of the US procedure, the two protein molecules were placed at a short distance in the z direction. The center of mass (COM) distance between the two protein molecules was defined as a reaction coordinate. The pull code in GROMACS [119] was used to generate, from an initial pulling trajectory, snapshots for the umbrella sampling simulations. Position restraints were imposed on one protein molecule, which was used as an immobile reference for pulling simulations. The other protein molecule was pulled away from the reference one along the z-axis with a harmonic force constant of 1000 kJ mol^{-1} nm^{-2} and a pull rate of 0.01 nm ps^{-1}. A final COM distance of approximately 11 nm between the two protein molecules was achieved, while the initial distance was about 6 nm for simulations 1,2 and 4,5, and 3.5 nm for simulation 3. In the case of simulation 3 a shorter minimum distance was selected to allow a better understanding of protein-protein interactions in presence of surfactants. Along the reaction coordinate, 25 windows for simulations 1 and 2, or 40 windows for simulations 3 to 5, were selected in the range of distances explored. The number and position of each window was chosen so as to have sufficient overlapping between nearby windows. The data within each trajectory were collected every 10 ps. The overall

simulation time for a single PMF profile was 375 ns for simulations 1 and 2, and 600 ns for simulations 3 to 5 (15 ns for each window). After removing the first 5 ns for equilibration, the PMFs were constructed with the WHAM method. Statistical errors on the computed PMFs were estimated by bootstrap analysis [186].

For a better investigation of hGH behavior in presence of the selected cosolutes, four other simulations were performed (sim. 6-9), see Table 3.1. Each simulation box contained 1 protein molecule, together with a given number of cosolute molecules, in water. The system was simulated in an NPT ensemble for 100 ns. In this case, the aim was to study the unbiased evolution of the system, and no umbrella sampling was performed.

The conformations assumed by hGH during simulations 6-9 were grouped by performing a cluster analysis based on the Daura algorithm [187], after discarding the first 40 ns of the trajectory. More specifically, the conformations were grouped together if the root mean square deviations (RMSDs) of the protein backbone were less than 0.1 nm compared to each other.

3.3.3 The Role of Stabilizers in the Aggregation Propensity of hGH Molecules

In simulations 1 and 2, the interaction potential between hGH molecules was evaluated, in the case of native and unfolded conformations, respectively. In Figure 3.4, V represents the interaction potential, while r is the distance between the proteins' COMs.

In the case of native protein molecules, Figure 3.4a, it is possible to notice that the potential of mean force showed a typical DLVO behavior [188–190]. More specifically, a primary minimum, responsible of irreversible aggregation, was observed at $r \approx 6.5$ nm. Considering an aggregation process, i.e., moving from right (large r) to left (small r) in the graph, this minimum was preceded by a maximum, which should act as an energy barrier. However, this maximum was not sufficiently pronounced to guarantee efficient inhibition of aggregation. At a distance of about 9 nm, a shallow secondary minimum, responsible of reversible flocculation, was evidenced by the US simulation. This PMF profile suggests that native hGH should be characterized by low stability towards aggregation.

The situation was even worse in the case of unfolded hGH (Figure 3.4b), where the secondary minimum was more pronounced, and lay at a larger value of r (≈ 9.5 nm), than for the native conformation. The shallow primary minimum, localized at $r \approx 7$ nm, was not preceded by a significant energy barrier. These results suggest that the unfolded molecule should be more unstable than the native conformation, being highly prone to irreversible aggregation. The addition of stabilizers is therefore of utmost importance when dealing with hGH, as this molecule tends to form aggregates, and the process is promoted by unfolding.

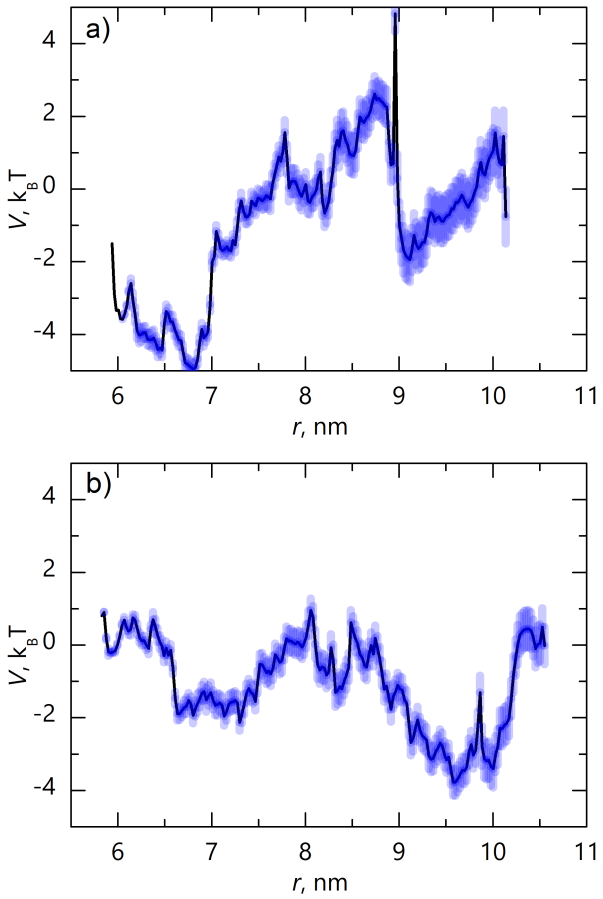

Figure 3.4: Potential of mean force for a) native and b) unfolded protein molecules, in absence of stabilizers. Statistical errors, shown as shaded areas in blue in the figures, were estimated by bootstrap analysis [186].

In Figures 3.5a and 3.5b, the potential of interaction between unfolded hGH in the presence of either sucrose or trehalose is shown, respectively.
It can be observed that, in the presence of both sucrose and trehalose, the primary minimum ($r \approx$ 6-7 nm) was preceded by a high energy barrier (at $r \approx$ 10 nm). If compared to the PMF between native hGH in absence of stabilizers, Figure 3.4a, the energy barrier was about 10 times higher in the presence of trehalose ($\approx$ 20 k_BT vs $\approx$ 2 k_BT), and almost 15 times higher ($\approx$ 30 k_BT vs $\approx$ 2 k_BT) in the presence of sucrose. This significant energy barrier hinders protein aggregation. Thus, in the presence of disaccharides, the hGH molecule should be stabilized against aggregation. As discussed in Chapter 2, sucrose should be a very good cryoprotectant, while trehalose should be slightly better as lyoprotectant. Despite this, their effect on the interaction potential between hGH molecules was very

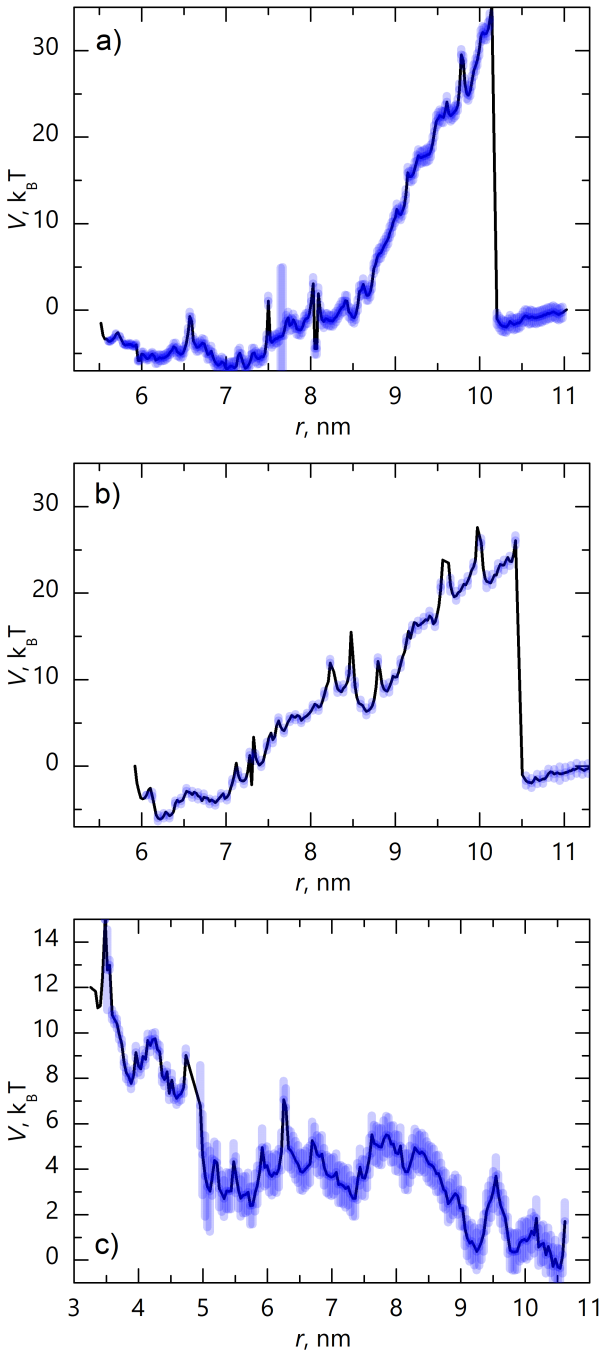

Figure 3.5: Potential of mean force between unfolded hGH molecules in presence of a) sucrose, b) trehalose and c) Tween 20. Statistical errors, shown as shaded areas in blue in the figures, were estimated by bootstrap analysis [186].

similar. The only difference was the extent of stabilization provided, with a higher energy barrier observed in the presence of sucrose.

On the other hand, the effect of Tween 20 on the PMF profile was substantially different, as shown in Figure 3.5c. In this case, a monotonic profile was observed. The potential of mean force in the presence of Tween 20 was characterized by the absence of an attractive branch, and, thus, only repulsion between protein molecules was observed for any value of r. This result seems to suggests that the aggregation of hGH is thermodynamically unfavorable in the presence of surfactant molecules. Therefore, Tween 20 should be a better stabilizer than the disaccharides against protein aggregation. In the following, the behavior observed for both the disaccharides and Tween 20 will be further investigated, with the aim to find an explanation to their effect on the PMF.

3.3.4 The Interactions Between hGH and Tween 20

Simulations 6 to 9 in Table 3.1 were used to investigate in more detail the interaction between hGH and Tween 20, and a comparison was also made to the case of sucrose and trehalose as excipients. In order to do this, the secondary structure of hGH during these simulations was analyzed using the DSSP tool [191,192]. This analysis was performed on the protein conformation that, according to the cluster analysis based on the Daura algorithm [187], was the most sampled during the equilibrated trajectory (last 60 ns). The secondary structure assumed by the protein in each simulation was then compared to the native fold. It was therefore possible to calculate the percentage of secondary structure that remained native-like. The α-helix content in the dominant protein structure for each simulation was also evaluated. This is an important parameter to be considered, because Costantino *et al.* [103] observed that the α-helix content in hGH should correlate with the degree of structural conservation. The results of this investigation are shown in Table 3.2, and a snapshot of each conformation is diplayed in Figure 3.6.
From the simulation results it seems that Tween 20 promoted a partial refolding of unfolded hGH. While at the beginning of simulation 7 the secondary structure of the unfolded protein (Figure 3.2b) had 58 % similarity with the native fold (Figure 3.2a), at the end of simulation 7 this value increased to 76 % (Figure 3.6b). The α-helix content increased as well, from 45 % to 63 %. If aggregation is mediated by unfolding, this result could at least partially explain the effect of Tween 20 on the prevention of protein aggregation. However, in the presence of Tween 20, native hGH (simulation 6) experienced a conformational change, and modified its secondary structure. These modifications led to an increased α-helix content, from the 68 % of the folded crystal structure to a final value of 70 % (Figure 3.6a).

On the contrary, sucrose did not lead to significant recovery of secondary structure and α-helix content. Only a 3 % increase in α-helix content was observed in

Table 3.2: DSSP analyses of simulations 6 to 9. The analyses were performed on the protein conformation that, according to the Daura algorithm [187], was the most sampled during the equilibrated trajectory (last 60 ns). In the last column, the percentage of the total population represented by the dominant cluster is shown.

Protein Structure	Sec. Str. %	α-helix %	% of population
Native	100	68	- (Fig. 3.2a)
Unfolded (Beginning sim. 7, 8, 9)	58	45	- (Fig. 3.2b)
Tween 20 - folded (End sim. 6)	87	70	77 (Fig. 3.6a)
Tween 20 - unfolded (End sim. 7)	76	63	80 (Fig. 3.6b)
Sucrose - unfolded (End sim. 8)	62	48	82 (Fig. 3.6c)
Trehalose - unfolded (End sim. 9)	72	54	80 (Fig. 3.6d)

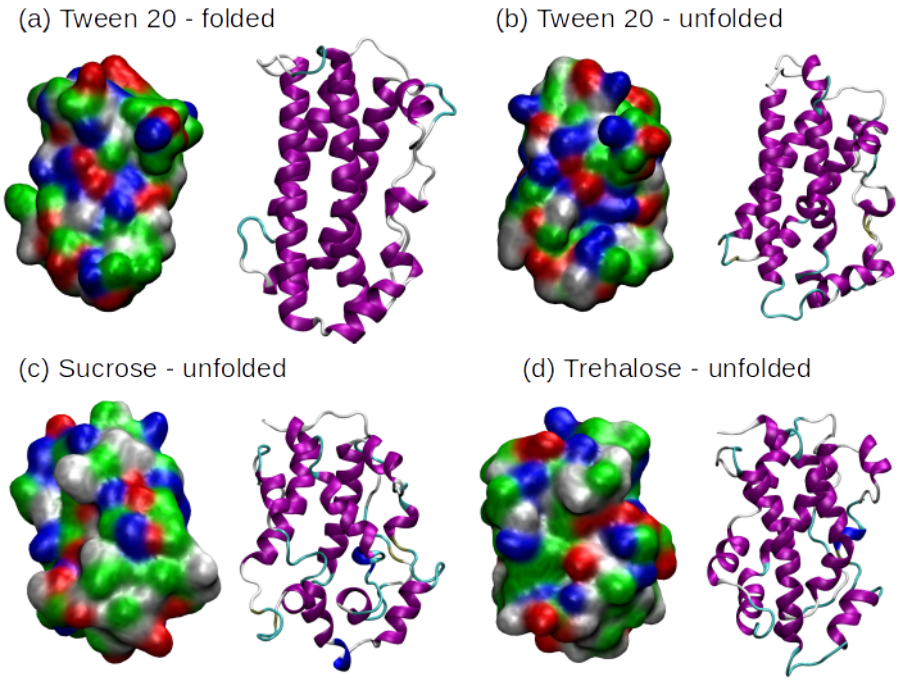

Figure 3.6: Illustration of the dominant structures obtained at the end of simulations 6 (a), 7 (b), 8 (c) and 9 (d). Colors used in the surface and cartoon representations are the same as in Figure 3.2.

simulation 8, from 45 % of the initial structure to 48 % at the end of the simulation (Figure 3.6c). Trehalose, was found to be better than sucrose, but not as effective as Tween 20 in promoting recovery of secondary structure. In simulation 9 the α-helix content of the unfolded protein molecule increased by 9 % during the

simulation, and there was a 14 % recovery in secondary structure (from 58 % to 72 %, see Figure 3.6d).

The stability of hGH during simulations 6 to 9 was also investigated using as critical parameter the non-polar surface area of the protein. Kasimova *et al.* [106] demonstrated that a large hydrophobic surface becomes exposed to the solvent in the unfolded conformations of hGH. This uncompensated exposure of non-polar surface leads to a highly unstable situation and a large tendency to aggregate. The non-polar area ratio R_{np} between two protein conformations was thus computed for each protein residue,

$$R_{np} = \frac{S_{np}(1)/S(1)}{S_{np}(2)/S(2)} \tag{3.2}$$

where S_{np} and S are the non-polar and total surface area of the residue under investigation, respectively, while 1 and 2 refer to the two protein conformations being compared. More specifically, structure 2 acts as a reference for structure 1. A value of R_{np} larger than 1 corresponds to an increase in the non-polar surface area exposed by the residue, and vice versa. Using R_{np}, two other variables were defined,

$$H_{all} = \frac{\sum_{all}(R_{np} - 1)}{n_{all}} \tag{3.3}$$

$$H_{APRs} = \frac{\sum_{APRs}(R_{np} - 1)}{n_{APRs}} \tag{3.4}$$

In H_{all}, the summation runs over all of the protein residues, while in H_{APRs}, only those residues that belong to the aggregation-prone regions (APRs) of the protein are included. The APRs are specific amino acid sequences which tend to drive aggregation [193–196], because of their charge, aromaticity, hydrophobicity, average packing density, surface area and secondary structure characteristics. A number of computational approaches have been developed to predict potential APRs in proteins. According to the AMYLPRED [197] server (http://aias.biol.uoa.gr/AMYLPRED/), the APRs for hGH correspond to residues 14-18 (MLRAH), 74-87 (ELLRISLLLIQSWL), 91-96 (QFLRSV), 161-168 (YGLLYCFRK) and 173-181 (VETFLRIVQ). While these APRs are buried within the core of the native protein, they are surface exposed in the unfolded conformation of hGH that was used in simulations 2-5 and 7-9, with the only exception of residues 91-96. A common feature of these APRs is that they are highly hydrophobic. This also explains why a large hydrophobic surface becomes exposed to the solvent upon unfolding of hGH. In Equations 2 and 3, n_{all} and n_{APRs} are the number of residues involved in the summation. For interpretation of the results, it should be considered that a positive value of H_{all} or H_{APRs} corresponds to an increase in the non-polar area exposed by the whole protein or by the APRs, respectively. The results of the analyses performed are shown

in Table 3.3, where conformations 1 and 2 of Equation 3.2 are listed in the first two columns. Again, the analysis was performed on the dominant protein structure during the equilibrated trajectory, according to the Daura algorithm [187].

Table 3.3: Values of H_{all} and H_{APRs} for simulations 6-9. In the last column, the percentage of the total population within the equilibrated trajectory represented by conformation 1 (according to the cluster analysis [187]) is shown.

Conformation 1	Conformation 2	H_{all}	H_{APRs}	% of population
Unfolded (Beginning sim. 7, 8, 9)	Folded	0.195	0.167	-
Tween 20 - folded (End sim. 6)	Folded	-0.05	-0.07	77
Tween 20 - unfolded (End sim. 7)	Folded	0.05	-0.14	80
Tween 20 - unfolded (End sim. 7)	Unfolded (Beginning sim. 7, 8, 9)	-0.141	-0.333	80
Sucrose - unfolded (End sim. 8)	Unfolded (Beginning sim. 7, 8, 9)	-0.195	-0.262	82
Trehalose - unfolded (End sim. 9)	Unfolded (Beginning sim. 7, 8, 9)	-0.146	-0.214	80

As can be observed in Table 3.3, Tween 20 promoted a reduction in the non-polar area of both folded and unfolded hGH. Moreover, this effect was particularly remarkable for the aggregation-prone regions, as evident from the more negative values of H_{APRs} with respect to H_{all}. This was particularly true in simulation 7, during which the originally unfolded conformation of hGH transformed into a much more compact structure, where the APRs were even more buried within the protein core than in the native state. This is evident from the negative value of H_{APRs} obtained comparing the structure at the end of simulation 7 (molecule 1), with folded hGH (molecule 2).

The non-polar surface area exposed by unfolded hGH decreased also in the presence of sucrose and trehalose. For both disaccharides this was particularly true for the APRs, but in the case of sucrose a greater reduction in non-polar surface area was observed. However, from the results shown in Table 3.3 it seems that neither sucrose nor trehalose was as effective as Tween.

The effect of Tween 20 on the protein structure was further investigated. In order to do this, the evolution of hGH non-polar surface area S_{np} during simulation 7 was evaluated, see Figure 3.7a. Moreover, the hGH total surface area S and radius of gyration R_g during simulation 6 are also shown in Figures 3.7b and 3.7c. The radius of gyration is defined as,

$$R_g = \sqrt{\left(\frac{\sum_i m_i |\mathbf{R}_i - \mathbf{R}_{COM}|^2}{\sum_i m_i}\right)} \tag{3.5}$$

where $\mathbf{R}_{COM}$ is the position of the center of mass, $\mathbf{R}_i$ the position of atom i having mass m_i, and the summation runs over all atoms in the protein.

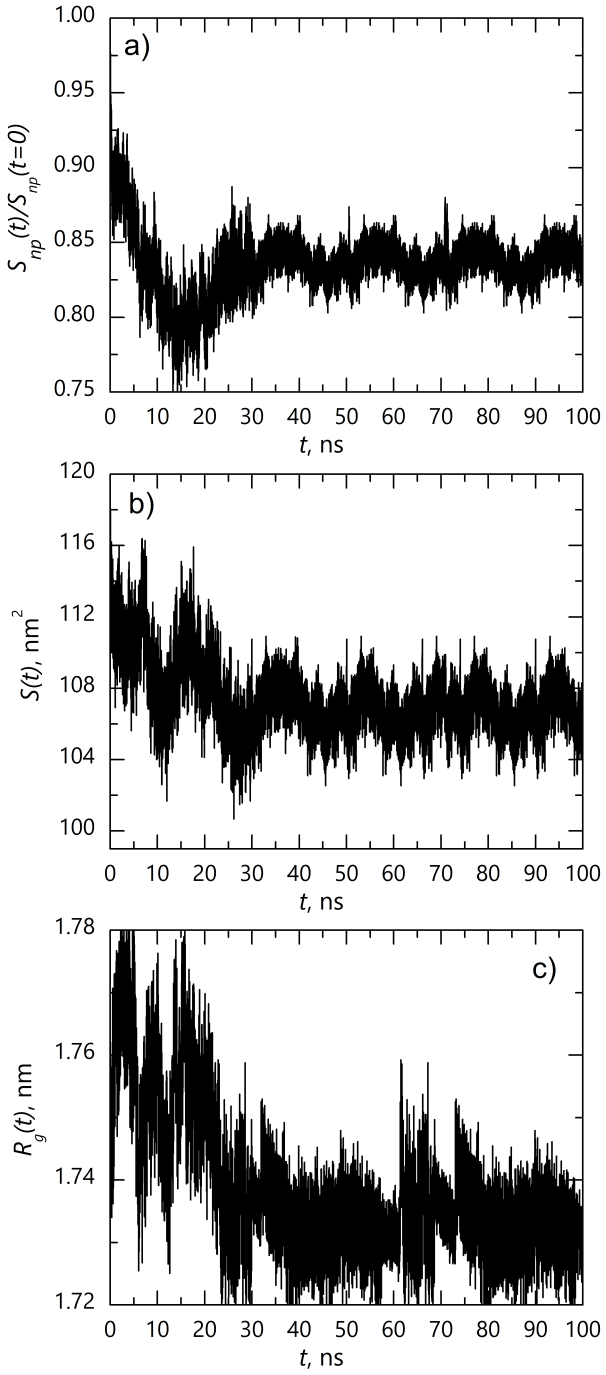

Figure 3.7: a) Evolution of unfolded hGH non-polar surface area $S_{np}(t)$ during simulation 7. b) Evolution of native hGH total surface area $S(t)$ during simulation 6. c) Evolution of native hGH radius of gyration $R_g(t)$ during simulation 6.

The results obtained suggest that Tween 20 promoted a significant reduction in the non-polar surface area exposed by unfolded hGH (Figure 3.7a), fostering the formation of a more compact structure. This more compact structure is characterized by an increased similarity with the native state of hGH (see Table 3.2). However, Tween 20 also promoted a shift towards a more compact structure of native hGH, as both the total area and radius of gyration of the folded protein decreased during simulation 6 (Figures 3.7b and 3.7c).

The arrangement of Tween 20 molecules around the protein surface was then investigated, as this could be related to the observed conformational changes. The tendency of surfactant molecules to crowd around the protein is evident from Figure 3.8.

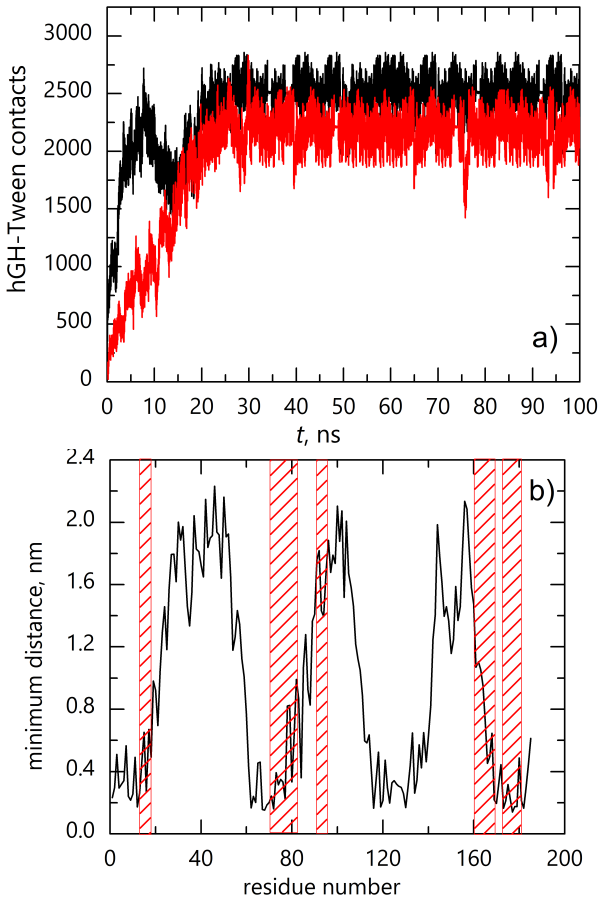

Figure 3.8: a) Number of contacts between Tween 20 molecules and folded (red line, sim. 6) or unfolded (black line, sim. 7) hGH. b) Minimum distance between Tween 20 molecules and the different residues of unfolded hGH, in simulation 7. Areas corresponding to the aggregation-prone regions of hGH are shaded in red.

In Figure 3.8a, the number of contacts between Tween 20 and unfolded (black curve) or folded (red curve) hGH during simulations 7 or 6 is shown. In this context, the number of contacts is defined as the number of atoms that are closer than 0.6 nm to one another. It can be observed that the number of protein-surfactant contacts increased by 5 times during the simulation time, and, remarkably, this increase was substantially faster in the case of unfolded hGH. However, the arrangement of Tween 20 molecules around the surface of unfolded hGH was not uniform, as evident from Figure 3.8b, where the minimum distance between the molecules of Tween 20 and the different residues of unfolded hGH during simulation 7 is shown. Tween 20 was generally closer to the APRs of the protein, shaded in red in Figure 3.8b, rather than to the other protein residues. This was true for all of the APRs, with the only exception of residues 91-96. This is not surprising because, as previously discussed, residues 91-96 were not surface exposed in the unfolded conformation of hGH.

The distribution of Tween 20 molecules around the protein seems to suggest a preferential interaction between the surfactant and the APRs. This could explain the negative values of H_{APRs} listed in Table 3.3. In the case of folded hGH, no specific pattern in the distribution of Tween molecules around hGH was observed (data not shown).

Moreover, the number of contacts between Tween 20 molecules increased during the simulation time, as shown in Figure 3.9. The increase in hGH-Tween and Tween-Tween contacts suggests the formation of a protein-surfactant complex, which should be less surface-active, and thus, more stable, than the protein alone.

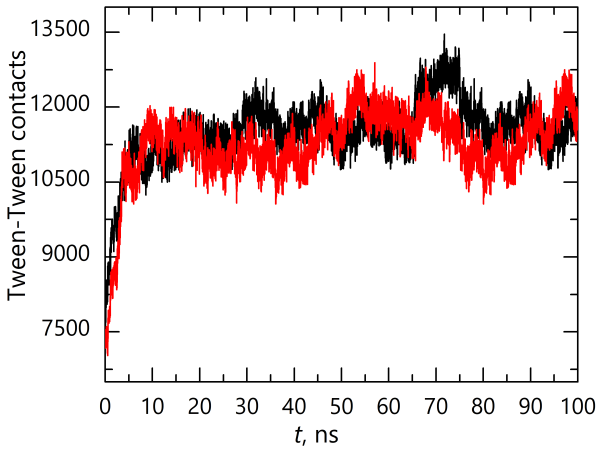

Figure 3.9: Number of contacts between Tween 20 molecules during simulations 6 (red line, native hGH) and 7 (black line, unfolded hGH).

The results obtained support the conclusion that Tween 20 should be able to effectively prevent protein aggregation. According to the MD simulations, the surfactant molecules tend to approach the protein surface, as shown in Figure 3.8. A protein-surfactant complex, stabilized against surface-induced denaturation, is thus formed.

In the case of unfolded hGH, Tween 20 molecules mainly crowded nearby the surface-exposed aggregation prone regions, as suggested by Figure 3.8b. Moreover, it seems that Tween 20 also promoted protein refolding, increasing the α-helix content of unfolded hGH and the secondary structure similarity to the native conformation, as shown in Table 3.2. This is in accordance with the previously proposed hypotheses that surfactants promote protein refolding, acting as molecular chaperones [49], and affect the α-helix content of proteins [53]. However, Tween 20 also affected native hGH, leading to a more compact conformation (Figure 3.7), but also modifying its secondary structure content (Table 3.2).

A possible explanation for this behavior was further investigated. For this purpose, the orientation of Tween 20 molecules toward the protein was studied in the case of both native and unfolded hGH. In Figure 3.10a, the radial distribution functions of the hydrophilic (red curve) and lipophilic (black curve) groups of Tween 20 around the center of mass of unfolded hGH are shown. These radial distribution functions were calculated as an average over the last 60 ns of the trajectory. It is evident that the hydrophilic groups of Tween 20 were closer to the protein than the lipophilic ones. This suggests that Tween 20 was oriented with its hydrophilic head in the direction of the surface-exposed APRs of unfolded hGH. As the APRs are hydrophobic, this creates a highly unstable situation, fostering the evolution of unfolded hGH towards a more native-like state, where the APRs are buried within the protein core.

This result is counter-intuitive, as the consequence of this orientation is the exposure of the Tween 20 hydrophobic tails to the solvent. A possible explanation can however be found if the non-polar surface area exposed by the entire protein-surfactant complex is considered. Figure 3.10c shows that there was a decrease in the non-polar surface area exposed by the whole complex. This may indicate that the reduction in the non-polar surface area of unfolded hGH induced by Tween 20 (Figure 3.7a and Table 3.3), and the contemporary clustering of surfactant molecules around the protein surface (Figures 3.8a and 3.9) offset the unfavorable exposure of the hydrophobic tails to the solvent. Therefore, the counter-intuitive arrangement of Tween 20 molecules around the protein surface promoted a so marked reduction in the non-polar surface area exposed by hGH that the whole process was energetically favourable.

By contrast, the situation was reversed for native hGH (Figure 3.10b). In this case, the hydrophobic (black curve) groups of the surfactant molecules were closer to the center of mass of the protein than the hydrophilic heads (red curve). This indicates that Tween 20 preferentially oriented with its lipophilic groups in

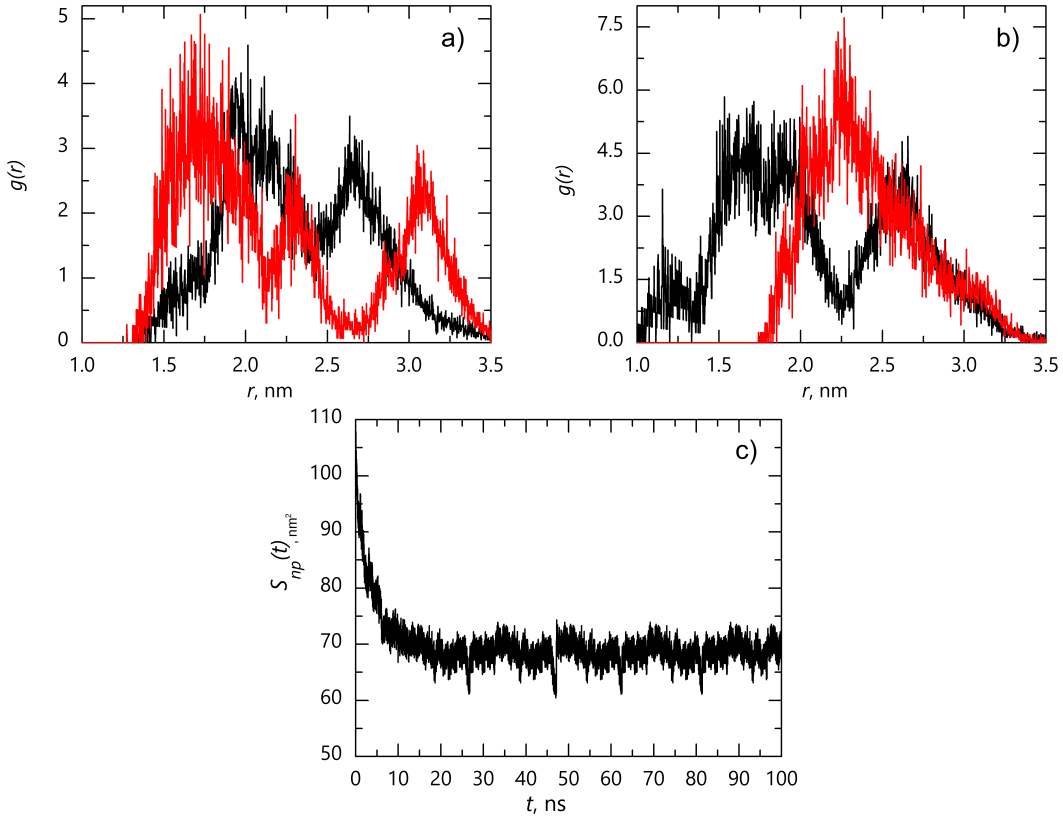

Figure 3.10: a) Radial distribution functions of the hydrophilic (red line) and lipophilic (black line) groups of Tween 20 around the center of mass of unfolded hGH. b) Radial distribution functions of the hydrophilic (red line) and lipophilic (black line) groups of Tween 20 around the center of mass of folded hGH. c) Evolution of the non-polar surface area $S_{np}(t)$ exposed by the entire protein-surfactant complex during simulation 7.

the direction of the native protein surface. This may explain the conformational changes discussed in Table 3.2, and the structural modifications observed for native hGH in presence of Tween 20.

In the following, the role of surfactants will be studied for the case of another model protein, namely, the GB1 hairpin. While hGH was simulated in the bulk only, the folding behavior of the GB1 peptide will be studied using metadynamics at the air-water, ice-water and silica-water surfaces, both in presence and absence of Tween 80. GB1 was selected for this study because it is a small peptide, more suitable than larger proteins, such as hGH, for a metadynamics investigation. Tween 80 was chosen because, as previously mentioned, it is more surface-active, and has a lower critical micelle concentration (CMC), than Tween 20.

3.4 The Effect of Surfactants on Surface-Induced Denaturation of Proteins

3.4.1 Simulation Setup

The GB1 hairpin was simulated in aqueous bulk solvent, and at the vacuum-water, ice-water, and silica-water interfaces, both in the presence and absence of the surfactant Tween 80. All simulations were performed using Gromacs 5.1.4 [198] patched with the Plumed 2.4.1 plug-in [199] using the GROMOS53A6 force field [200] for the protein and the GROMOS53A6$_{OXY+D}$ force field [201] for the surfactant. An initial structure for GB1 was taken from the PDB (PDB ID: 1GB1 [202]), while the topology file for Tween 80 was obtained from Tang *et al.* [203]

Ten Tween 80 molecules were inserted into the simulation box. The resulting concentration, about 32.4 mM, is higher than the critical micelle concentration (CMC), which has been reported to lie in the range 0.010-0.015 mM at 25 °C for Tween 80 [204]. This is important because, if the surfactant prevents surface-induced unfolding by coating the interface and thus inhibiting protein adsorption, surfactant concentrations near or above the CMC are generally necessary to provide stabilization.

The GenIce algorithm [205] was used to obtain an initial configuration of hexagonal (Ih) ice with proton disorder and zero dipole moment, and the generated ice layer was oriented with the basal {0001} plane in the direction of the liquid phase. The silica-water interface considered in this study was hydrophilic (OH terminal group), and the same force field described in Das *et al.* [206] was used. The Ih ice water molecules and silica heavy atoms were kept frozen in place during the simulations. For simulating the air-water interface, a 5 nm vacuum space was included along the z axis above the liquid phase.

The simulations were performed using periodic boundary conditions. Long-range electrostatic interactions were treated using the Particle Mesh Ewald (PME) algorithm [126]. A cut-off radius of 1.2 nm was used for both the real-space Coulombic and Lennard-Jones interactions. In all simulations, 1 native protein molecule was introduced into the simulation box, solvated with SPC/E water molecules (the number of water molecules in each simulation box is listed in Table 3.4), and neutralized using Na^+ ions. The liquid phase was equilibrated for 5 ns at 1 bar and 260 K in the NPT ensemble, using Berendsen pressure and temperature coupling [207].

A complete list of the simulations performed, in which the dimensions of both the simulation box and the liquid phase are specified, is shown in Table 3.4. In simulations 1 through 8, the effect of an interface on protein stability was investigated, while simulations 9 and 10 were performed in the bulk and used as reference. In simulations 1-3 and 9, no Tween 80 molecules were introduced into the box, while, in simulations 4-8 and 10, 10 surfactant molecules were present in the liquid phase. Simulations 5, 7 and 8 were all performed at the ice-water interface, but a different

set of collective variables (CVs) was biased (column 4 in Table 3.4).

Table 3.4: List of the PBMetaD simulations performed*.

Sim. #	Surface	# Tween 80	CVs	Box Dim., nm	Liq. Phase Dim., nm (# water mol.)
1	Air-Water	-	d, $\alpha\beta$, R_g, β_s	8.0 x 8.0 x 13	8.0 x 8.0 x 8.0 (20614)
2	Ice-Water	-	d, $\alpha\beta$, R_g, β_s	7.8 x 8.1 x 11	7.8 x 8.1 x 8.0 (17202)
3	Silica-Water	-	d, $\alpha\beta$, R_g, β_s	8.1 x 8.8 x 13	8.1 x 8.8 x 8.0 (23168)
4	Air-Water	10	d, $\alpha\beta$, R_g, β_s	8.0 x 8.0 x 13	8.0 x 8.0 x 8.0 (16360)
5	Ice-Water	10	d, $\alpha\beta$, R_g, β_s	7.8 x 8.1 x 11	7.8 x 8.1 x 8.0 (16605)
6	Silica-Water	10	d, $\alpha\beta$, R_g, β_s	8.1 x 8.8 x 13	8.1 x 8.8 x 8.0 (22696)
7	Ice-Water	10	d, $\alpha\beta$, R_g, β_s, CN_t	7.8 x 8.1 x 11	7.8 x 8.1 x 8.0 (16605)
8	Ice-Water	10	d, $\alpha\beta$, R_g, β_s, CN_h	7.8 x 8.1 x 11	7.8 x 8.1 x 8.0 (16605)
9	-	-	$\alpha\beta$, R_g, β_s	8.0 x 8.0 x 8.0	8.0 x 8.0 x 8.0 (16946)
10	-	10	$\alpha\beta$, R_g, β_s	8.0 x 8.0 x 8.0	8.0 x 8.0 x 8.0 (16360)

*300 ns simulation time

3.4.2 Simulation Approach

To improve the sampling of the conformational state space, the metadynamics enhanced sampling method was used [98, 208]. Metadynamics significantly reduces the amount of computational time needed to obtain a meaningful statistical sampling of configuration space by introducing a history-dependent bias potential that acts on a set of relevant order parameters or collective variables (CVs, see Figure 3.11). In the context of protein adsorption, the efficacy of metadynamics largely depends on a judicious choice of CVs, which can include:

- the distance between the surface and the protein center-of-mass;
- various metrics of protein secondary structure;
- the coordination of surfactant molecules with either the protein or surface.

To account for as many CVs as needed we use a variant of metadynamics called parallel bias metadynamics [209] (PBMetaD) introduced by Pfaendtner and Bonomi. PBMetaD alleviates the difficulty of depositing a high-dimensional bias by

instead constructing multiple one-dimensional biases each acting on a single CV in parallel. This approach, and metadynamics in general, have been successfully applied whenever the exploration of the free-energy landscape is hindered by the presence of energy barriers. More specifically, PBMetaD and other enhanced sampling techniques have been widely used to study protein-surface interactions [210–215]. To further accelerate the sampling, multiple-walkers (MWs) [216] PBMetaD was performed, using 3 multiple walkers. PBMetaD simulations were conducted at 260 K in the NVT ensemble, with the temperature controlled by stochastic velocity rescaling [125]. An integration timestep of 2 fs was used for the MD equations of motion, and each walker was simulated in a production run of 100 ns.

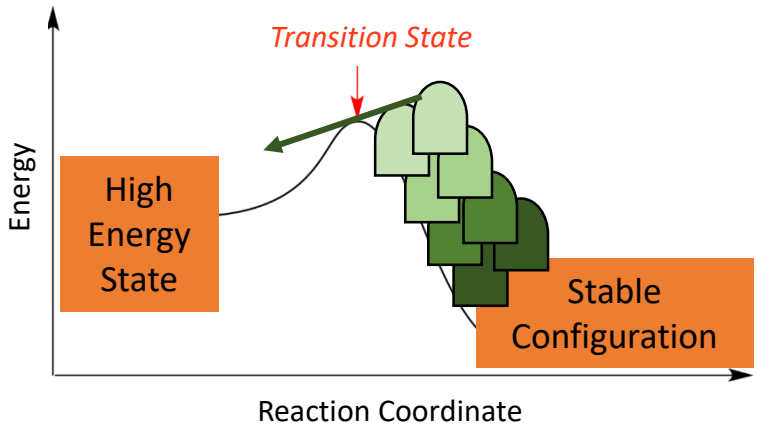

Figure 3.11: Scheme of metadynamics. A Gaussian history-dependent bias potential acting on a set of collective variables is used to flood the free energy landscape.

The fundamentals of PBMetaD are presented in the following section.

Parallel Bias Metadynamics

In a PBMetaD simulation several one-dimensional bias potentials $\{V(s_i)\}$ are deposited, each acting on a different collective variable s_i, which is a function of the atomic coordinates $s_i(\mathbf{R})$. The individual bias potentials have the familiar form of a metadynamics bias potential

$$V(s_i,t) = \int_0^{t'} dt\ \omega_i(t) \exp\left(-\frac{(s_i(\mathbf{R}) - s_i(\mathbf{R}(t)))^2}{2\sigma_i^2}\right) \tag{3.6}$$

with σ_i a Gaussian width specific to CV s_i and $\omega_i(t)$ a deposition hill height which decreases as the bias accumulates according to

$$\omega_i(t) = \omega_{0,i}\ \exp\left(-\frac{V(s_i,t)}{k_B \Delta T_i}\right) \times \frac{\exp\left(-\frac{V(s_i,t)}{k_B T}\right)}{\sum_{j=1}^{n} \exp\left(-\frac{V(s_j,t)}{k_B T}\right)} \tag{3.7}$$

which is the usual well-tempered [217] prescription, modified by an additional conditional probability giving a higher weight to CVs that have a lower bias potential acting on them. k_B is the Boltzmann constant; $\omega_{0,i}$ is an initial hill hight for CV s_i, and ΔT_i is an input parameter in units of temperature related to the so-called bias factor $\gamma_i = \frac{T+\Delta T_i}{T}$. The PBMetaD bias potential acting on all CVs has the form

$$V_{PB}(s_1, s_2, ..., s_n, t) = -k_B T \ln \sum_{i=1}^{n} \exp\left(-V(s_i, t)/k_B T\right). \tag{3.8}$$

The advantage of PBMetaD is that it makes it possible to include more CVs than is feasible with conventional metadynamics while still keeping to the spirit of the well-tempered approach of having a bias that becomes quasi-stationary in the long-time limit. Up to five CVs were included in the PBMetaD simulations

- d: the distance between the protein center of mass (COM) and the surface. This CV helps to sample states during the process of adsorption and has been used in a number of studies on peptide adsorption [211, 212, 218].

- $\alpha\beta$: the so-called AlphaBeta similarity CV defined as

$$\alpha\beta = \sum_i \frac{1}{2}\left[1 + \cos\left(\theta_i - \theta_i^{ref}\right)\right] \tag{3.9}$$

  where θ_i are the backbone ϕ and ψ dihedral angles and θ_i^{ref} is a reference value determined from the PDB structure. This CV allows for the conformational exploration of the protein backbone dihedral angles [219].

- R_g: the protein radius of gyration. This CV helps to sample extended or compact protein configurations [213, 219].

- β_s: the antiparallel β-sheet content. Biasing this CV enhances the sampling of configurations with different degrees of antiparallel β-sheet content. The antiparallel β-sheet CV is defined as the number of 3+3 residues in an antiparallel β-sheet configuration [220, 221] computed as

$$\beta_s = \sum_\mu g[r_{dist}(\{R_i\}_{i\in\Omega_\mu}, \{R^0\})]$$

  where the summation runs over all possible segments involved in the antiparallel β-sheet structure, $\{R_i\}_{i\in\Omega_\mu}$ are the atomic coordinates of a set Ω_μ of six residues of the protein, and $g(r_{dist})$ is a switching function

$$g(r_{dist}) = \frac{1 - \left(\frac{r_{dist}}{r_0}\right)^8}{1 - \left(\frac{r_{dist}}{r_0}\right)^{12}} \tag{3.10}$$

  A cutoff distance of $r_0 = 0.08$ nm was used, and r_{dist} is the distance RMSD with respect to a reference antiparallel β-sheet configuration $\{R^0\}$.

- CN_h/CN_t: finally, the coordination number of the Tween 80 hydrophobic tails (CN_t) or heads (CN_h) around the protein surface was considered as a CV

$$CN = \sum_{i=1}^{N_A} \left[\sum_{j=1}^{N_B} \frac{1 - \left(\frac{r_{ij}}{r_0}\right)^6}{1 - \left(\frac{r_{i,j}}{r_0}\right)^{12}} \right] \tag{3.11}$$

  where $r_{i,j}$ is the distance between species i in group A and j in group B. Group A was the protein center of mass and group B consisted of the center of masses of either the hydrophobic tail or hydrophilic head group on each surfactant molecule. This CV was used to count the number of contacts between the protein and hydrophobic/hydrophilic surfactant atoms. The cutoff distance r_0 was set to 1 nm.

In all PBMetaD simulations the initial Gaussian height $\omega_{0,i}$ was set to 2 kJ/mol, the bias factor γ to 10, and the Gaussian deposition rate to 1 hill/ps. The previously defined collective variables were biased, and the σ (Gaussian width) values used were 0.3 nm, 0.5, 0.03 nm, 0.2 and 0.7 for d, $\alpha\beta$, R_g, β_s and the coordination number, respectively. In the long-time limit the biases converge to $V(s_i, t \to \infty) = -\frac{\Delta T_i}{T+\Delta T_i} F(s_i) + C$ where $F(s_i)$ is the free energy as function of CV s_i and C is an irrelevant constant. Convergence was assessed by monitoring the fluctuation of the one-dimensional free energy profiles $F(s_i)$ for the different CVs during the last 10% of the simulation time (see Appendix B, Figures B.1-B.10).

In order to obtain meaningful information from biased simulations it is necessary to extract unbiased probability distributions which can be inverted to obtain free energies. In the context of PBMetaD, unbiased multi-dimensional free energy surfaces can be reconstructed by applying the time-dependent reweighting method of Tiwary and Parrinello [222]. This reweighting method was used to obtain all of the two-dimensional free energy surfaces presented in the remainder of this Chapter.

3.4.3 Folding of GB1 in the Absence of Surfactants

Simulations were first performed in the bulk as a reference point for protein conformational stability. In agreement with previous simulations (using a number of force fields) [163, 169–172], GB1 adopts a hairpin structure. The free energy surface (FES) as a function of the radius of gyration R_g and the antiparallel β-sheet content β_s is shown in Figure 3.12a.

The most sampled structure (A9) is a compact hairpin, with radius of gyration 0.78 nm and a large β-sheet content ($\beta_s = 3$). Here, as well as in the following, the protein conformations have been labeled with a letter followed by a number. The number identifies the simulation (first column of Table 3.4), while the letter is used to distinguish different structures within the same simulation. Letter A always corresponds to the most folded structure within each simulation.

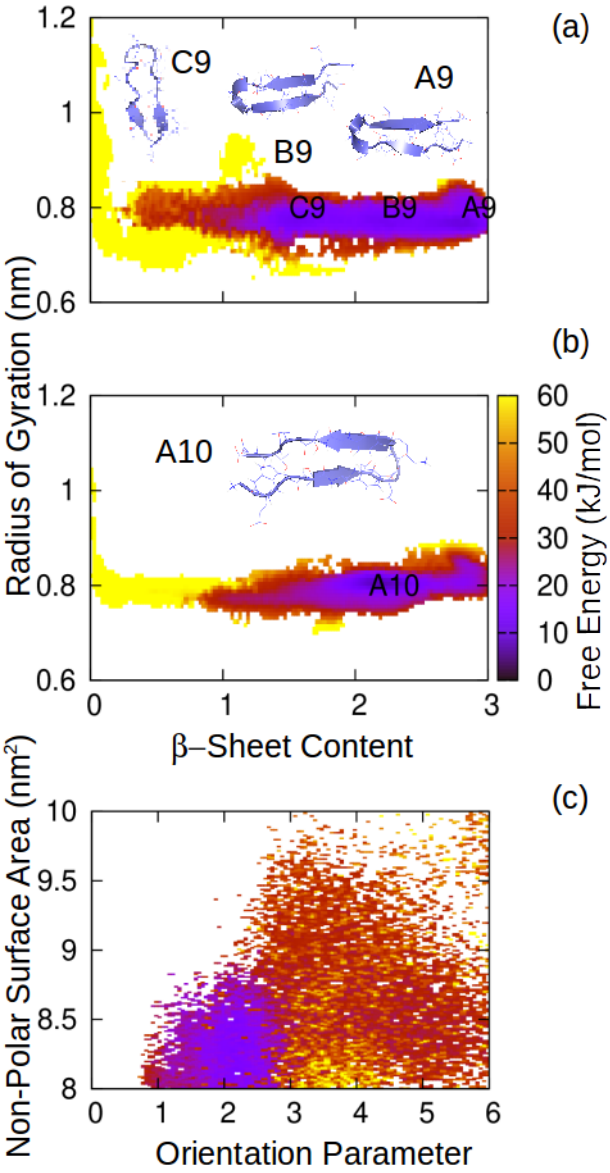

Figure 3.12: Free energy surface (FES) as function of the GB1 radius of gyration and of its β-sheet content during (a) simulation 9 in Table 3.4 (bulk solution, no surfactants) and (b) simulation 10 in Table 3.4 (bulk solution, 10 Tween molecules). In panel (a) and (b), the letters on the FES identify the most sampled protein conformations, and a cartoon of each structure is also shown. (c) FES as function of the non-polar surface area exposed by the GB1 hairpin and of the orientation parameter during simulation 10 in Table 3.4.

The free energy surfaces of the peptide, in the presence of the three surfaces and

in the absence of surfactant, is shown in Figure 3.13 as a function of antiparallel β-sheet content (β_s) and the distance d between the protein COM and the surface (left panel) and as a function of R_g and β_s (middle panel). Representative structures are shown in the right hand panel.

The protein-surface interactions were also analyzed, and the average distance from the interface for the 6 closest and the 3 furthest protein residues is shown in Figure 3.14.

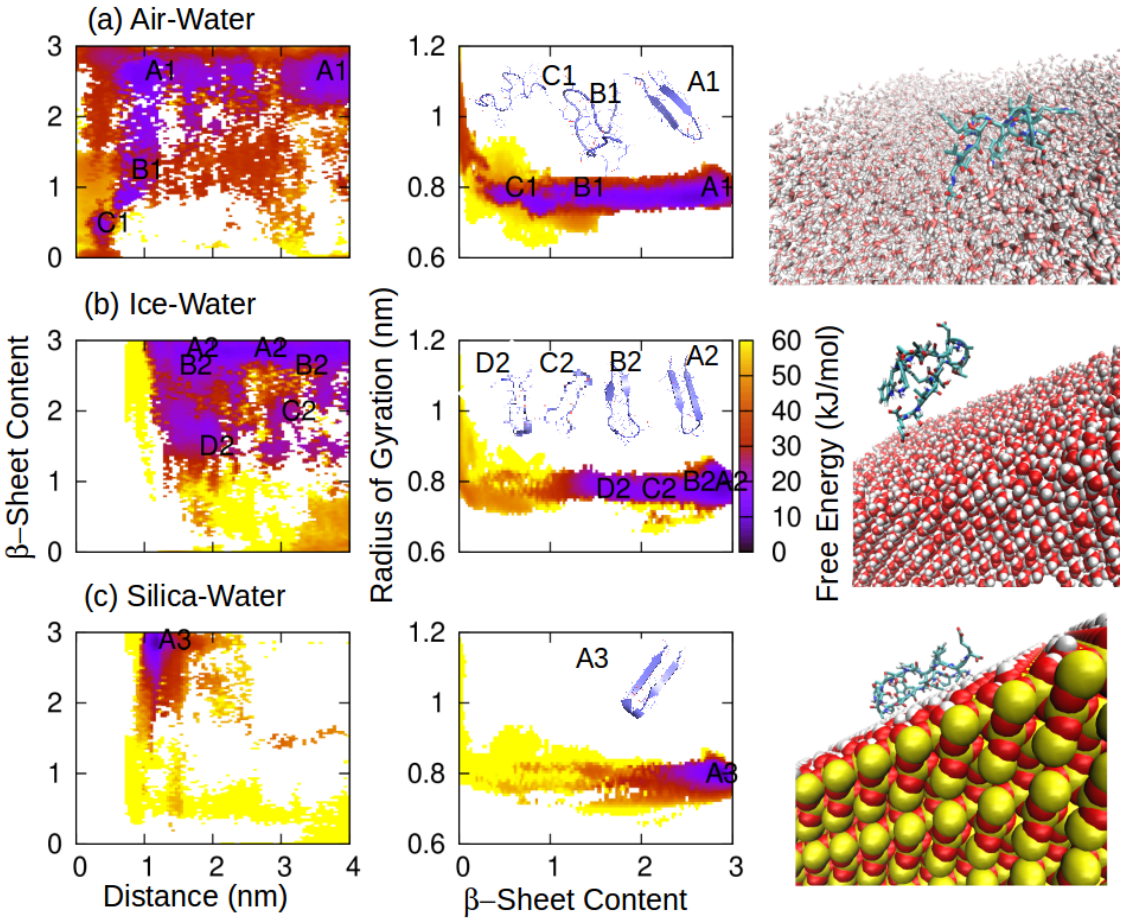

Figure 3.13: Free energy surface (FES) as function of distance and β-sheet content (left) or β-sheet content and radius of gyration (center) for the (a) air-water, (b) ice-water and (c) silica-water interface. On each line, a snapshot of the system being investigated is also displayed (right). The letters on the FES identify the most sampled protein conformations, and a cartoon of each structure is also shown in the central panel. The results refer to simulations 1, 2 and 3 in Table 3.4, where no Tween 80 molecules were present in the simulation box.

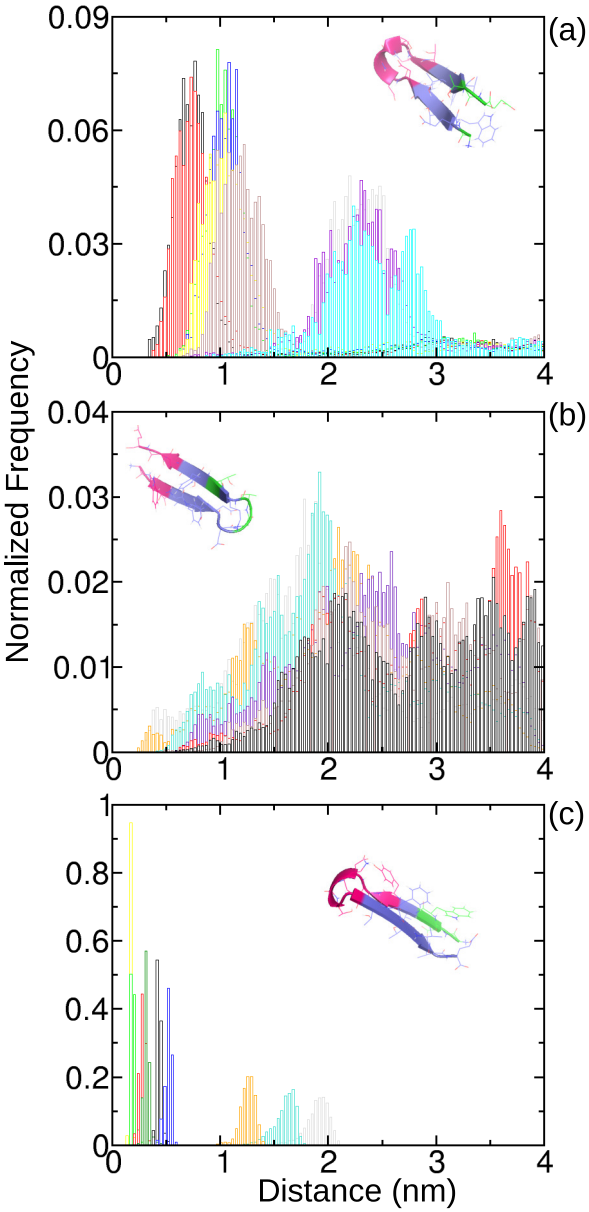

Figure 3.14: Minimum distance between different protein residues (grey: 41G, turquoise: 42E, orange: 43W, dark green: 45Y, blue: 46D, yellow: 47D, red: 48A, black: 49T, light green: 50K, brown: 51T, indigo: 54V, violet: 55T, cyan: 56E) and the (a) air-water, (b) ice-water or (c) silica-water interface. In the cartoons of the protein structures shown within each graph, the residues closest to the surface are evidenced in red, while the furthest ones are colored in green.

Air-Water Interface

The FES for the case of the air-water interface is shown in Figure 3.13a. The protein is destabilized compared to the bulk, populating conformations with low

β-sheet content ($\beta_s = 1.25$ or 0.5, respectively for conformations B1 and C1). The most unfolded conformation (C1, also shown in the right hand panel in Figure 3.13a) is surface bound (less than 0.5 nm from the surface), while the native-like conformation A1 lies at distances greater than 1 nm from the surface. Analysis of the distance of the amino acids from the surface (Figure 3.14a) shows that the region including the hairpin turn, i.e. residues 46-51 (DDATKT), approached the air-water interface the most, while residues 41G, 55T and 56E were, on average, the furthest away. The amino acids that interacted with the surface, in particular residues 48-50 (ATK), showed a large increase in non-polar surface area. A significant increase in the exposed hydrophobic area was also observed for the amino acid pair 43W-54V, which is involved in the hairpin hydrophobic core. The observed loss of structure is therefore related to the adsorption of the protein at the interface, and the subsequent exposure of its hydrophobic groups to the vacuum. Indeed, structure C1 shows a larger radius of gyration ($R_g \approx 0.8$ nm) and non-polar surface area (9 nm^2) than conformation A1 (0.75 nm and 7.5 nm^2, respectively).

Ice-Water Interface

Interestingly, in contrast to the air-water interface, only partial unfolding of the protein is observed at the ice-water interface (Figure 3.13b). In fact, conformation A2 to D2 showed similar radius of gyration $R_g \approx 0.78$ nm. The least native-like conformation at the ice-water interface (conformation D2, $\beta_s = 1.5$), showed significantly more secondary structure than the conformation C1 sampled at the air-water surface, retaining the turn structure. The orientation of the protein to the ice surface is different than in the case of the air-water interface, with the terminal residues 41-43 (GEW) and 54-56 (VTE) now closest, and the turn residues 48A, 49T and 51T being the furthest (Figure 3.14b). It can be observed that at the ice-water surface, the hydrophobic pair formed by amino acids 43W-54V is affected by the surface, leading to partial unfolding. However, unlike the air-water interface, the ice surface does not strongly interact with the turn residues. As a result, the structures observed in the presence of ice had smaller variations in the non-polar surface area and in the radius of gyration as compared to the case of the air-water surface. Moreover, as can be seen from the FES as a function of β_s and distance from the surface (left handed panels), the protein does not approach the surface as closely as in the air-water case. The denaturing effect of the surface is long-ranged, with partially unfolded conformations observed 3 nm from the surface of ice (Figure 3.13b). This is further confirmed by the average distance between the protein residues and the surface, as shown in Figure 3.14b. The range of distance values explored by the different residues during the simulation was extremely broad when compared to the case of the air-water interface (Figure 3.14a).

Silica-Water Interface

In sharp contrast to the ice-water and air-water interface, only the native structure is significantly sampled at the hydrophilic silica-water interface (Figure 3.13c). Comparison to the bulk FES (as a function of R_g and β-sheet content, Figure 3.12a) shows that the protein is in fact stabilized compared to the bulk in the presence of the silica surface. The FES as a function of β-sheet content and distance from the surface (Figure 3.13c) indicates that the hairpin remains near the surface, as evidenced from the small range of d (around 1 nm) sampled. The protein is oriented with the turn region, residues 45-50 (YDDATK), towards the surface and residues 41-43 (GEW) towards the bulk (Figure 3.14c). Adsorption of the protein does not translate into a loss of secondary structure. It is also interesting to note that the protein approached the silica-water interface more closely than the ice-water surface, with a minimum value of the protein COM-surface distance of $d \approx 1$ nm and $d \approx 1.5$ nm for the silica and ice surfaces, respectively. In the case of the air-water interface (Figure 3.13a), the protein could stretch out in the vacuum, and the distance from the interface can hence be smaller than in the case of solid surfaces.

GB1 Behaves Differently at Different Surfaces

Overall, these results suggest that the unfolding process of the GB1 hairpin at the air-water interface was mainly driven by hydrophobic interactions. The denaturation process involved disruption of the hairpin turn, and exposure of residues 43W and 54V, which are part of the hydrophobic core in the native protein. A similar exposure of the hydrophobic residues 43W-54V promoted partial unfolding at the ice-water surface. However, the turn structure was retained in this case. In contrast, the GB1 hairpin was significantly stabilized by the silica-water surface, as a result of weak interactions between the protein and silica, with the surface restricting the conformations that the protein can adopt and thereby stabilizing the protein through a confinement effect.

3.4.4 Role of Surfactants against Surface-Induced Denaturation

The effect of the surfactant on protein stability will now be investigated. The effect of Tween 80 on the FES at the air-water, ice-water and silica-water interfaces is shown in Figure 3.15.

Air-Water Interface

In the case of the air-water interface (Figure 3.15a), the surfactant molecules were extremely effective in stabilizing the native protein conformation, with the peptide now only significantly populating structures with high β-sheet content ($\beta_s = 3$ or 2.3

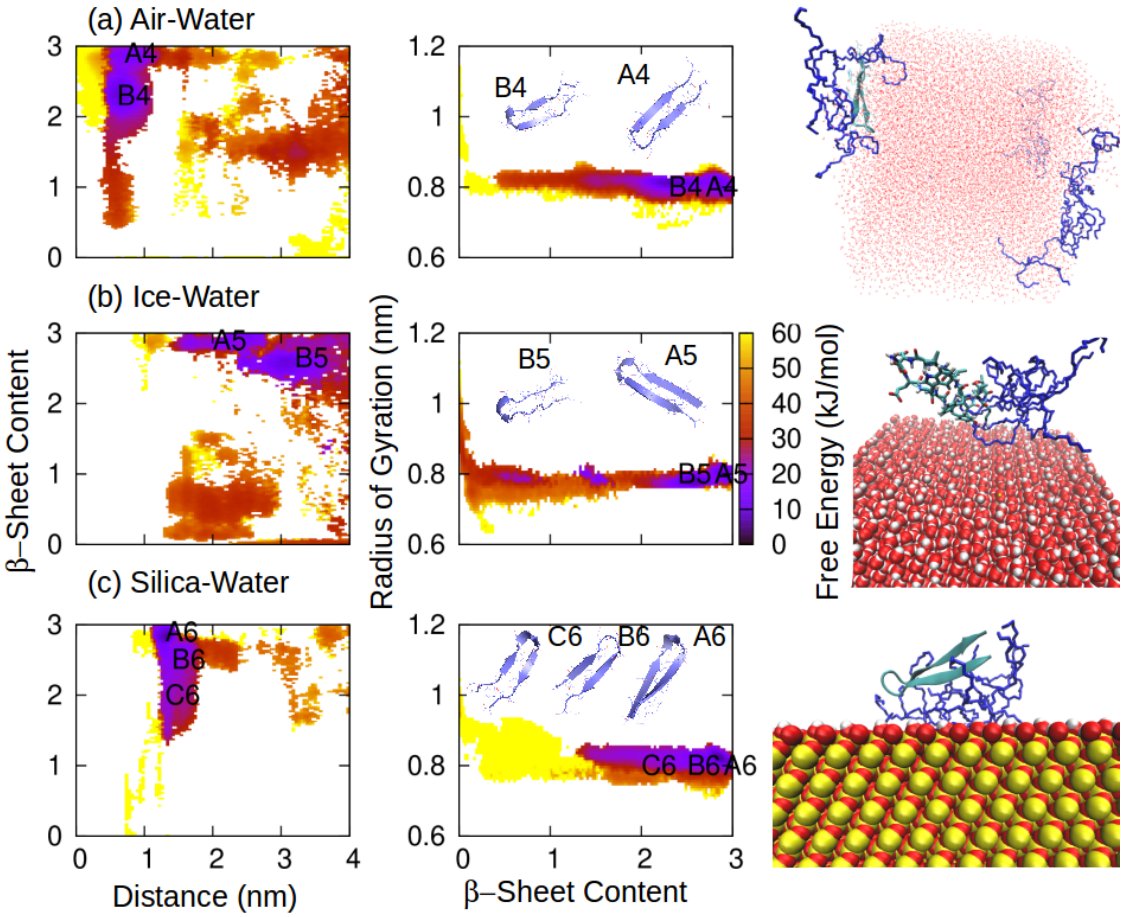

Figure 3.15: Free energy surface (FES) as a function of distance and β-sheet content (left) or β-sheet content and radius of gyration (center) for the (a) air-water, (b) ice-water and (c) silica-water interface. On each line, a snapshot of the system being investigated is also displayed (right), where the Tween 80 molecules are represented in blue. The letters on the FES identify the most sampled protein conformations, and a cartoon of each structure is also shown in the central panel. The results refer to simulations 4, 5 and 6 in Table 3.4, where 10 Tween 80 molecules were present in each simulation box.

for structures A4 and B4, respectively), and a compact radius of gyration R_g. The stabilizing effect of Tween 80 was remarkable, especially considering that structures A4 and B4 were more folded than the structure C9 (Figure 3.12a) sampled in bulk and in the absence of surfactants. It is also interesting to observe that conformation B4 showed the same radius of gyration ($R_g = 0.8$ nm) and β-sheet content ($\beta_s = 2.3$) as structure A10 (Figure 3.12b), which was the most populated in bulk upon addition of surfactants.

The surfactants bind to the surface (as shown by Figure 3.16a) thereby reducing adsorption of the protein at the air-water interface, as evident from the larger protein-surface distance in Figure 3.15a (0.5-1 nm) than in Figure 3.13a (structure C1 was mostly sampled at $d \approx 0.3$ nm). The protein was therefore no longer as close to the surface as in the absence of surfactants, because the Tween 80 molecules coated the interface, sterically preventing the hairpin adsorption. This significantly limited the range of d values explored by the protein, that generally remained confined between 0.5 and 1 nm from the interface (Figure 3.15a). In the absence of Tween 80, the protein denatured at the air-water interface because of direct interactions between the protein's hydrophobic core and the air; this process is prevented in the presence of the surfactants that preferentially bind to the interface over the protein.

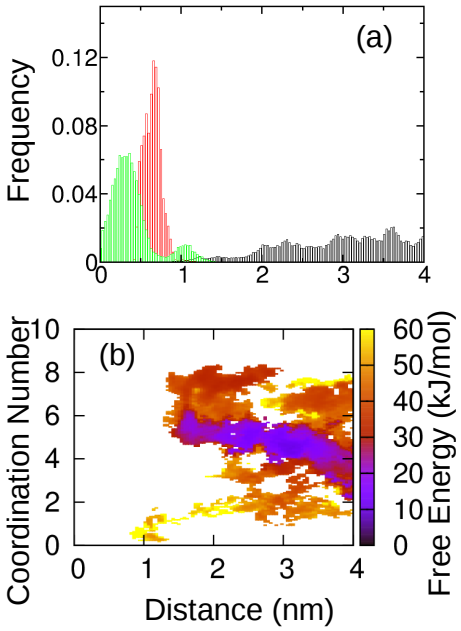

Figure 3.16: (a) Histogram showing the distribution of the distance between the Tween 80 molecules COM and the air-water (green bars, simulation 4), silica-water (red bars, simulation 6) and ice-water (black bars, simulation 5) interface. (b) FES as function of the coordination number between the Tween 80 molecules and the GB1 hairpin, and the distance between the protein COM and the ice-water interface (simulation 5).

Ice-Water Interface

The Tween 80 molecules showed a similar stabilizing effect in the presence of ice, as shown in Figure 3.15b. In this case, only two very compact conformations were observed, with a native-like secondary structure ($\beta_s = 3$ or 2.65 for conformations A5 or B5, respectively). The mechanism of stabilization appears to be quite different than in the air-water case, as the surfactants do not coat the interface (Figure 3.16a) but rather bind directly to the protein, as evident from the high (around 5) coordination number between the Tween 80 molecules and the protein (Figure 3.16b). The protein is further away from the surface than in the surfactant-free case, with values of d smaller than 1.5 nm rarely sampled (Figure 3.15b).

To further probe the nature of the binding of the surfactants to the protein, in particular the orientation (head or tail) of the surfactants with respect to the protein, an orientation parameter was defined as the ratio between the coordination number of the surfactant tails (CN_t) and surfactant heads (CN_h) around the protein,

$$\text{Orientation Parameter} = \frac{CN_t}{CN_h} \tag{3.12}$$

A 1 nm distance was chosen as cut-off for both coordination numbers. An orientation parameter larger than 1 indicates that the hydrophobic tails are closer to the protein surface than the hydrophilic heads, and vice versa.

The FES showing the β-sheet content as function of the orientation parameter for the ice-water simulations in the presence of surfactants is shown in Figure 3.17a. Native-like, compact structures (A5 and B5) are stable for values of the orientation parameter smaller than 1 (hydrophilic heads facing the peptide), while the few structures with smaller β-sheet content (primarily energetically unfavorable) were observed only when the orientation parameter was larger than 1 (hydrophobic tails facing the protein).

Therefore, this FES suggests the presence of an orientation-dependent mechanism in which the native-like structures are stabilized when the surfactant molecules orient with their hydrophilic heads towards the protein, with the opposite orientation promoting loss of structure.

To confirm the orientation-dependent mechanism, 2 additional simulations were performed (simulations 7-8 in Table 3.4), in which the coordination number between the surfactant hydrophobic tails (CN_t) or hydrophilic heads (CN_h) and the protein was added as a collective variable in the metadynamics protocol. By biasing the coordination number CN_t, the orientation of the surfactant molecules towards the protein was also implicitly biased, promoting, in simulation 7, the tails-toward-the-protein configuration. This biasing led to the population of a large number of unfolded structures, as shown in Figure 3.17b (snapshots of the protein structures A7-E7 are also illustrated in Figure 3.17c). Structures A7 and B7 were folded

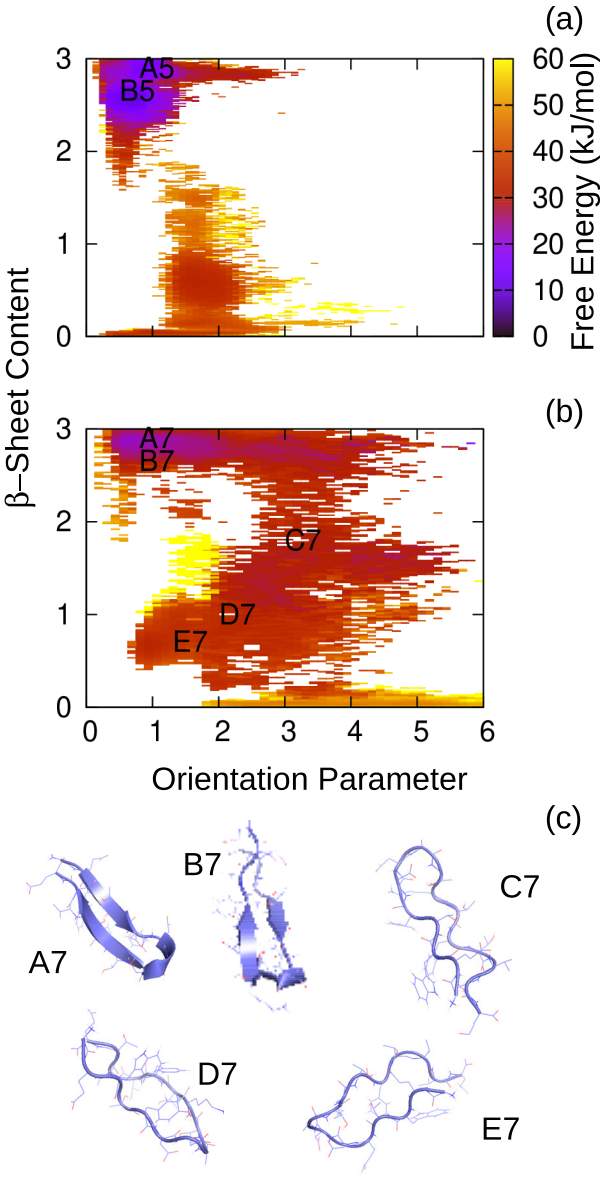

Figure 3.17: FES as function of the hairpin β-sheet content, and the orientation parameter during (a) simulation 5 (ice-water interface, presence of surfactants, unbiased coordination number) and (b) simulation 7 (ice-water interface, presence of surfactants, biased coordination number of surfactants tails). (c) Cartoon representations of the protein structures sampled during simulation 7.

(β_s = 3 and β_s = 2.65, respectively) , while structures C7 (β_s = 1.8), D7 (β_s=1) and E7 (β_s=0.7) showed a decreased β-sheet content. Analysis of the surfactant

orientation around the protein confirmed that these unfolded structures, characterized by an almost complete loss of secondary structure, had values of the orientation parameter larger than 1 (Figure 3.17b). In contrast, when the coordination number of the Tween hydrophilic heads CN_h was biased, as in simulation 8, the FES was restricted towards more native conformations (Figure 3.18). In fact, the two most sampled conformations, A8 and B8 (Figure 3.18b), showed a high β-sheet content ($\beta_s = 3$ and 2.5, respectively). In addition to supporting an orientation-dependent mechanism, simulations 7 and 8 also show that the coordination number is not a good collective variable when the molecules involved have an amphipathic nature as the choice of CN (tail or head) influences the outcome of the simulation.

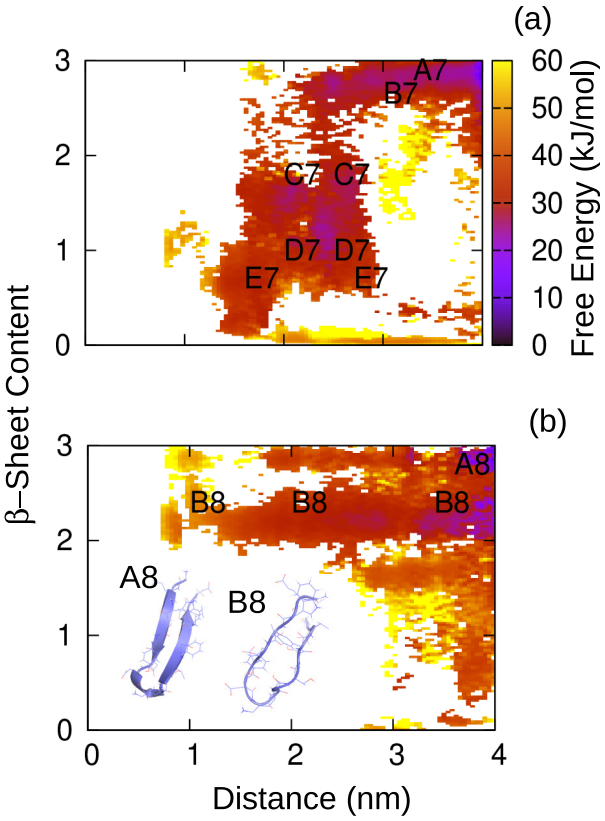

Figure 3.18: Free energy surface (FES) as function of distance between the protein COM and the ice surface d and β-sheet content for (a) simulation 7 (ice-water surface, presence of surfactants, biased coordination number of the surfactant hydrophobic tails) and (b) simulation 8 (ice-water surface, presence of surfactants, biased coordination number of the surfactant hydrophilic heads). The letters on the FES identify the most sampled protein conformations, and a cartoon of structures A8 and B8 is also shown in panel (b).

Due to their amphiphilic nature, surfactants gravitate to interfaces where there is a sharp gradient in hydrophobicity/hydrophilicity. In the case of air-water, the surfactants are strongly attracted to the interface, with the hydrophobic tails interacting with the air, and the hydrophilic heads exposed to water. Due to this preferential binding to the surface, the surfactants interact more strongly with the surface than the protein. In contrast, at the ice-water interface, while a gradient in hydrophobicity/hydrophilicity does exist, it is mild. The ice water can be considered more hydrophobic than liquid water, [223] with the ice-water molecules having diminished mobility, and reduced propensity to hydrogen bond. In this case, the surfactants are not strongly drawn to the interface, and preferentially coat the protein rather than the ice-water surface. Therefore, the hydrophilic heads of the surfactant face towards the protein, that exposes its hydrophilic residues in its native state.

The coordination number of the Tween 80 molecules around the peptide was probed further, as it has been reported that the surfactant concentration required for protein stabilization should not correlate with the CMC when the surfactants bind directly to the protein. Rather, the degree of protection should be maximized at the molar binding stoichiometry between the surfactant and the protein [48, 50, 52]. For instance, Chou *et al.* [176] studied the binding of both Tween 20 and Tween 80 with Albutropin, and they reported a molar binding stoichiometry of 10:1 and 9:1 (surfactant:protein), respectively. The value of 5 that was found in this work to be the most stable coordination number between the Tween 80 molecules and the β-hairpin in Figure 3.16b likely corresponds to the molar binding stoichiometry for the protein being investigated, and lies in the range of values which is experimentally observed.

Silica-Water Interface

In contrast, the surfactants slightly destabilize the protein at the hydrophilic silica-water interface (Figure 3.15c), promoting structures with smaller values of β-sheet content, with the C6 conformation, for example, having $\beta_s = 2$. In the absence of surfactants, as shown in Figure 3.13c, only the native-like conformation, characterized by $\beta_s = 3$, was sampled at the silica-water interface. Hence, in the case of the hydrophilic silica surface, Tween 80 promoted a partial loss of secondary structure.

The surfactants were strongly attracted to the silica-water interface, where they formed a layer of adsorbed molecules (Figure 3.16a). This hindered protein adsorption and increased the distance between the hairpin and the surface, as can be seen by comparing Figures 3.13c and 3.15c. A shown in the FES in Figure 3.19, the surfactants are oriented with their hydrophobic tails towards the protein and hydrophilic heads near the silica surface (orientation parameter larger than 1). Based on the surfactant-biased simulations at the ice-water interface, this orientation of

the Tween 80 molecules should promote the sampling of less structured conformations. The reason that the surfactants orient in this protein-unfavorable manner is because the hydrophilic heads have high affinity to the hydrophilic silica surface. As a result, only the hydrophobic tails remain available for interaction with the protein, and mild destabilization ensues.

Bulk Solution

A similar (but even milder) destabilization is observed when Tween 80 is added to the protein in the bulk (a hydrophilic aqueous environment), through the same orientation-dependent mechanism (as can be seen by comparing Figures 3.12a and 3.12b). This may also explain the reduction in secondary structure similarity to the native state upon addition of Tween 20 observed in the case of folded hGH as model protein (simulation 6 in Table 3.1). The FES in the presence of Tween 80 populates fewer conformations with β_s near 3 and favors conformations (A10) with slightly lower β-sheet content ($\beta_s = 2.3$). The reason for this mild destabilization is the orientation of the hydrophobic tails towards the protein (orientation parameter larger than 1 in Figure 3.12c). As in the case of silica, the GB1 hairpin is surrounded by the hydrophobic chains of the surfactant molecules, and this promotes an increased exposure of the non-polar residues. However, the protein is more stable in the presence of the surface than in the bulk because of the confinement effect imparted by the hydrophilic surface.

Therefore, the orientation-dependent mechanism of Tween 80 can explain both the protein stabilization at the ice-water interface, and the loss of structure induced at the hydrophilic silica-water surface and in the bulk. The amphipathic nature of Tween 80 also explains the prevention of protein unfolding at the air-water surface, as the tails favorably interact with the interface, reducing protein adsorption and its deleterious effects.

3.5 Conclusions

Overall, the results for hGH in presence of Tween 20 suggest that surfactants can effectively prevent aggregation, and may even promote refolding. The formation of a surfactant-protein complex is the preliminary step for both these phenomena. Moreover, it seems that the specific orientation of surfactant molecules toward the protein surface is at the basis of the observed mechanism, with the heads-toward-the-protein orientation fostering refolding. These orientation-related processes are possible because of the amphiphilic nature of surfactants. Other protectants, such as the disaccharides, show different effects, because of their lack of a surface-active nature.

It was also observed that GB1 was destabilized at the air-water and ice-water

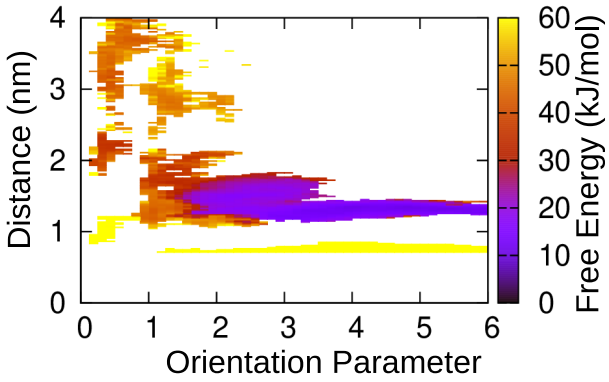

Figure 3.19: FES as function of the orientation parameter and the distance between the protein COM and the silica-water interface (simulation 6 in Table 3.4).

interfaces, but stabilized at the silica surface. In the case of the air-water interface, destabilization was caused by direct interaction of the hairpin turn and of the hydrophobic pair formed by residues 43W-54V with the vacuum. This strong interaction promoted a large exposure of non-polar surface area, and a resulting loss of structure. In the case of the ice-water surface, the interaction of residues 43W-54V with the surface promoted a partial loss of structure, but the turn structure was retained and complete unfolding was not observed. By contrast, the GB1 hairpin was stabilized at the silica surface, because of confinement effects and absence of strong protein-surface interactions.

It was also found that the surfactants stabilize GB1 when the denaturing air-water and ice-water surfaces are present, but slightly destabilize the peptide at the otherwise stabilizing silica interface. It was observed that the surfactant molecules bind to the surface in the presence of air and silica, while they preferentially cluster around the protein in the case of ice. At the silica surface, the surfactant molecules coating the interface also directly interact with the protein, leading to mild destabilization. In line with the results obtained for hGH, an orientation-dependent mechanism of the surfactants was also identified, in which the GB1 hairpin was stabilized when the hydrophilic heads of the surfactant were oriented towards the protein, and mildly destabilized when the hydrophobic tails pointed towards the peptide. The latter orientation stabilized partially unfolded states of GB1, characterized by a larger non-polar surface area. It was also found that the tails-toward-the-protein configuration is favored in a hydrophilic environment, explaining the mild destabilization observed at the silica-water interface. A hydrophobic environment, on the other hand, promotes the heads-toward-the-protein arrangement, which is particularly efficient in stabilizing the protein native structure. Finally, in the case of the air-water surface, the coating of the interface by the surfactant molecules, and the

resulting inhibition of protein adsorption, accounts for the observed stabilization of the protein native structure. In this case, the amphipathic nature of Tween 80 plays a major role, with the favorable interaction between the hydrophobic tails and the vacuum leading to the coating of the surface by the surfactants.

Therefore, the simulations discussed in this Chapter suggest that the action of surfactants is complex, as it can either stabilize or mildly destabilize the protein. The amphiphilic nature of the surfactant, and its relative affinity for the protein and the surface, eventually determines the effect on the protein structure.

List of Symbols

C	-	constant
CN_h	-	coordination number of the surfactant hydrophilic heads
CN_t	-	coordination number of the surfactant hydrophobic tails
d	m	distance
F	J mol^{-1}	free energy
$g(r)$	-	radial distribution function
$g(r_{dist})$	-	switching function
H_{all}	-	summation of non-polar area ratio over all residues (Equation 3.3)
H_{APRs}	-	summation of non-polar area ratio over residues in the aggregation prone regions (Equation 3.4)
k	N m^{-1} mol^{-1}	harmonic force constant
k_B	J K^{-1} mol^{-1}	Boltzmann constant
m	kg	mass
n	-	number of collective variables being biased
n_{all}	-	number of residues in the protein
n_{APRs}	-	number of residues in the aggregation prone regions of the protein
N_A	-	number of species in group A (Equation 3.11)
N_B	-	number of species in group B (Equation 3.11)
r	m	radial distance
r_0	m	cutoff distance
r_{dist}	m	distance RMSD with respect to a reference antiparallel β-sheet configuration
R	m	atomic coordinates, scalar
R^0	m	reference antiparalled β-sheet configuration
$\mathbf{R}$	m	atomic coordinates, vector
$\mathbf{R}_{COM}$	m	coordinates of the center of mass, vector

R_g	m	radius of gyration
R_{np}	-	non-polar area ratio (Equation 3.2)
s	-	collective variable
s_0	-	reference value of a collective variable
S	m^2	total surface area
S_{np}	m^2	non-polar surface area
t	s	time
T	K	temperature
ΔT	K	temperature difference related to the bias factor (Equation 3.7)
V	J mol^{-1}	potential

Greek Letters

$\alpha\beta$	-	AlphaBeta similarity collective variable
β_s	-	antiparallel β-sheet content
γ	-	bias factor
θ	-	ϕ/ψ dihedral angle
θ^{ref}	-	reference ϕ/ψ dihedral angle
σ	-	Gaussian width
ω	J mol^{-1}	Gaussian height
ω_0	J mol^{-1}	initial Gaussian height
Ω_μ	-	set of six residues of the protein

Chapter 4

Protein Denaturation During Freeze Drying: A Focus on the Ice-Water Interface

In Chapter 3, the folding behavior of the GB1 hairpin at different surfaces has been investigated. Among the others, the ice-water interface has been considered. It was observed that the GB1 protein partially unfolded when in contact with the ice surface, even though no strong binding was observed between the peptide and the interface. The present Chapter aims to provide further explanation to these observations.

It was mentioned already that the freezing process may have detrimental effects for the protein, leading to denaturation, aggregation and loss of biological activity. The conformational changes that may occur during freezing are related to cold denaturation [15, 16], cryo-concentration, and formation of the ice-water interface.

The reduced protein stability at extreme values of temperature is a well-known phenomenon, but the effects of high and low temperature conditions are remarkably different. Heat denaturation is entropically driven, as it is mainly dictated by the increase in conformational entropy upon unfolding. In contrast, cold unfolding is enthalpically driven [15–19,224]. More specifically, at low temperature the repulsive interaction between nonpolar residues and water is weaker, leading to a partial unfolding of the protein [225–227]. Because of the active participation of water in the cold denaturation process of a protein, an explicit description of the solvent is generally necessary in computational models to account for both heat and cold denaturation [228–230]. Using these water-explicit models, for instance, it was found that cold unfolding may be driven by the increased stability of hydrogen bonds at the protein-water interface at low temperature. At some point, shell water forms hydrogen bonds more favorably than bulk water [230], resulting in an overall enthalpic gain. Hence, low-temperature unfolding occurs with an associated release of heat from the hydration shell, whereas the opposite is true in the case of

heat-induced denaturation [19].

The sharp reduction in temperature is not the only phenomenon that may be harmful to protein stability during freezing, and the cryo-concentration process also plays a central role. Physical properties related to concentration, such as ionic strength and relative composition of solutes, may change. For instance, freeze-concentration-induced pH changes, due to selective crystallization of buffer components [231], may result in protein unfolding. This is particularly true when sodium or potassium phosphate buffers are used. Additionally, freezing polymer solutions may cause phase separation due to a change in polymer solubilities at low temperature [29, 232]. Phase separation creates a large interface that can promote protein denaturation. Furthermore, the protein will have a thermodynamic preference for one phase over the other and may partition into a phase with a low concentration of stabilizers.

However, it was shown [61] that ice formation represents the most critical destabilizing factor. Using LDH as model protein, a remarkable loss of activity was observed in frozen systems, while no degradation was detected in concentrated solutions at the same temperature and composition, but without ice. Similarly, solutions of the azurin protein exhibited a dramatic decrease in the average phosphorescence lifetime of the Trp-48 residue at the onset of ice formation, which is indicative of protein unfolding [20]. This means that prevention of ice-induced denaturation represents a key issue whenever a protein is subjected to a freezing process.

However, the reasons at the basis of the ice-induced denaturation of proteins are not completely understood, mainly because of the lack of appropriate experimental techniques to address this problem. Strambini and Gabellieri [20] used intrinsic phosphorescence emission to demonstrate that the formation of ice alters the native fold of proteins, and suggested that this perturbation may originate from the direct interaction between the protein and the ice surface. They also found that the addition of cryoprotectants such as glycerol and sucrose dramatically attenuates, or even eliminates, the loss of structure during freezing. It was therefore hypothesized that the stabilizing action of cosolutes should be regarded as a combined effect of lowering the freezing temperature and decreasing the adsorption affinity of the protein by coating the surface of ice. It was also suggested [20] that the preferential exclusion of the cosolutes from the protein, which is believed to stabilize the native fold of proteins in the bulk [32], may also contribute to an increased protein stability at the ice-water interface. However, the protective effect of these osmolytes against ice-induced denaturation is, at present, poorly understood.

Hence, the present Chapter aims to provide insight into the effect of the ice-water surface on protein stability, and to clarify the role of cryoprotectants. For this purpose, all-atom Molecular Dynamics (MD) simulations will be used. The effect of the interaction between the same model protein used in the previous analysis, i.e., human growth hormone (hGH), and the ice-water interface, will be studied both in

presence and absence of cryoprotectants. Some of the most common pharmaceutical excipients, namely, sucrose, trehalose, cellobiose, lactose, sorbitol and glycine have been selected for this investigation. The MD results will show that the interaction of the excipient molecules with specific amino acids sequences of the protein, rather than with the molecule as a whole, seems to be the key aspect for preservation of the protein structure.

Afterwards, metadynamics [208] will be employed to exhaustively sample the thermodynamic behavior of the 62 residue IgG-binding domain of peptostreptococcal protein L at the ice surface. Protein L has been widely studied both experimentally [233–238] and computationally [239, 240], and shows both a α-helix (res. ^{23}EKATSEAYAYADTL^{36}) and a β-sheet structure (res. ^{2}IKANLI^{7}, ^{13}TQTAEF^{18}, ^{43}WTV^{45} and ^{53}TLNIKF^{58}), which may be disrupted upon interaction with the ice-water surface (see Figure 4.1). Protein L was selected for this investigation because it is a smaller protein compared to hGH, and therefore more suitable for the application of enhanced sampling techniques. The presence of glucose will also be considered, and its behavior at the ice surface will be compared to its protective action in the bulk.

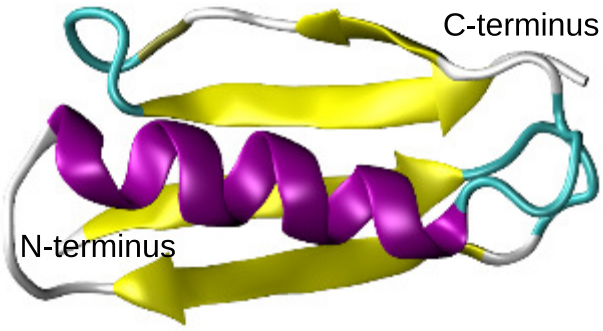

Figure 4.1: Cartoon structure of the IgG-binding domain of peptostreptococcal protein L, where the different secondary structure elements have been highlighted using different colors. Purple: α-helix, yellow: extended β-sheet, cyan: turn, white: coil.

This analysis will show that the unfolding process is characterized by a small free energy barrier, and a fast kinetics, in presence of ice. MD simulations also suggest that the observed destabilization of the protein structure at the ice surface may be due to an enhancement of cold denaturation phenomena, and is not mediated by direct adsorption. The addition of glucose will be shown to stabilize the protein structure, again because of preferential exclusion from specific patches on the protein surface. Hence, a new possible interpretation of the preferential exclusion

theory will be proposed to explain the results obtained for both hGH and protein L [1].

4.1 The Ice-Water Interface Promotes Unfolding of hGH

4.1.1 Simulation Setup

The molecular dynamics simulations were performed using the software GROMACS 5.0. The human growth hormone topology file was obtained from the RCSB Protein Data Bank (PDB 3hHR [120]). The simulations were carried out using periodic boundary conditions. For the protein the OPLS-AA force field [93] was used, while water molecules were modeled using either the TIP4P [242] or the TIP4P/Ice model [243] (see Appendix A).

The TIP4P/Ice water model was chosen because it fairly well predicts the melting temperature of ice and the ice-water phase diagram [244], as shown in Table 4.1.

Table 4.1: Melting temperature as predicted by some common water models

Model	Melting temperature, K
TIP4P/Ice	270±3
TIP4P/2005	250±3
TIP4P-EW	244±3
TIP4P	230±3
TIP5P	272±3
SPC/E	214±3

The TIP4P model was used in simulations where freezing of the system was not desired, as it makes it possible to reach very low values of temperature, preserving all water molecules in the liquid state.

For the simulations involving sugar-based excipients, the parameter set optimized for carbohydrates was used [245]. Cut-off radius for both Coulombic interactions (calculated using the PME method [126]) and Lennard-Jones interactions was 1.0 nm.

The simulated conditions aimed to reproduce the stresses encountered by the protein during freezing, and are summarized in Table 4.2. The box dimensions were in any case equal to 6.3 x 5.8 x 9.3 nm and each simulation box was filled with 1

[1] Part of this chapter is based on already published work [241]

native molecule of hGH. The GenIce algorithm [205] was used to obtain an initial configuration of hexagonal (Ih) ice with proton disorder and zero dipole moment. The starting simulation boxes contained Ih ice (the initial slab was 6.3 x 5.8 x 3.6 nm) and liquid water, togheter with hGH and, in some cases, excipient molecules (30 molecules in each box). Sucrose, trehalose, lactose, cellobiose, sorbitol and glycine were considered for this study (simulations 3-8 in Table 4.2). For one simulation only (simulation 1 in Table 4.2), no ice layer was introduced in the simulation box and all water molecules were described using the TIP4P water model [242]. In this way, no freezing would occur at the temperature used for the simulation (233 K, which is low enough to observe fast freezing when using the TIP4P/Ice model, but high enough to avoid freezing in the case of the TIP4P force field). This simulation was performed so as to make a comparison between the protein behavior in absence and presence of the ice-water interface.

The potential energy of the initial configurations was minimized with a steepest descent algorithm. A short simulation (1 ns) was thereafter sequentially performed at 233 K in an NVT ensemble imposing a position restraint on all ice-water oxygen atoms. Temperature was controlled by the velocity rescaling thermostat [125]. The system was then simulated for 10 ns at the previously chosen temperature in an NPT ensemble allowing the ice water molecules to oscillate freely. Pressure was set to 1 atm and controlled by the Parrinello-Rahman barostat [127]. Subsequently, a 100 ns simulation was performed in an NPT ensemble at 233 K storing the configurations every 10 ps. The temperature chosen is below the melting point of hexagonal ice obtained in previous studies for the TIP4P/Ice water model, but above the one for the TIP4P model. The melting of ice or the freezing of liquid water was monitored by looking at the total energy of the system [244]. If the total energy of the system decreases during the simulation time, it means that freezing has occurred, while an increase in the total energy indicates that at least part of the ice has melted. Moreover, visual inspection of the simulation box at the beginning and at the end of the simulation was used to confirm the behavior suggested by the total energy profile.

Table 4.2: Details of the simulations performed for hGH as model protein.

Sim. #	Water Force Field	Ice Surface	Excipient
1	TIP4P	No	-
2	TIP4P/Ice	Yes	-
3	TIP4P/Ice	Yes	sucrose
4	TIP4P/Ice	Yes	trehalose
5	TIP4P/Ice	Yes	lactose
6	TIP4P/Ice	Yes	cellobiose
7	TIP4P/Ice	Yes	sorbitol
8	TIP4P/Ice	Yes	glycine

4.1.2 Effect of the Ice-Water Interface in Absence of Excipients

In the case of hGH, two simulations were performed for its freezing in absence of excipients (simulations 1 and 2 in Table 4.2). In the first one, no ice-like water molecules were introduced in the box and a suitable water model, TIP4P [242], was used in order to prevent freezing at the simulated temperature of 233 K. The second simulation was performed at the same temperature, but an ice slab was introduced in the system and the water model, TIP4P/Ice [243], was chosen so as to ensure freezing. The ice layer progressively extended during this last simulation, eventually surrounding the protein molecule (see Figure 4.2).

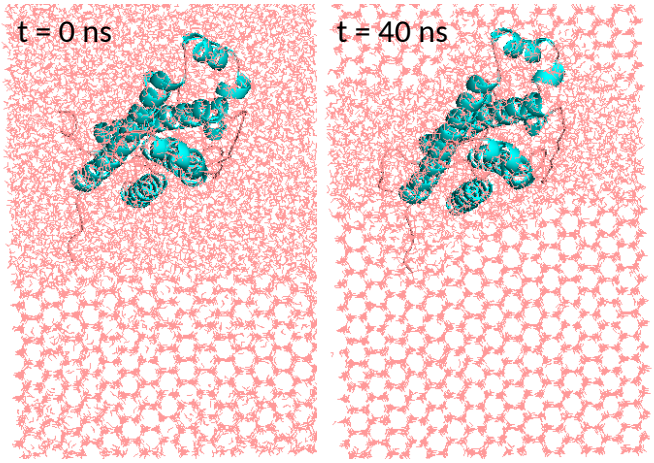

Figure 4.2: Snapshots illustrating the extension of the ice slab during simulation of the freezing of hGH, when the TIP4P/Ice water model was used.

Kasimova *et al.* [106] demonstrated that a large hydrophobic surface becomes exposed to the solvent in the unfolded conformation of hGH. This uncompensated exposure of non-polar surface leads to a highly unstable situation, and a large tendency to aggregate. Thus, to evaluate the protein conformational stability, the non-polar area ratio R_{np} was defined as,

$$R_{np}(t) = \frac{S_{np}(t)/S(t)}{S_{np}(0)/S(0)} \tag{4.1}$$

where t is time, and S_{np} and S are the non-polar and total surface area of the protein, respectively. At time 0, that is, for the native conformation of hGH, $S_{np}(0)$ and $S(0)$ are equal to about 50 and 105 nm^2, respectively. A value of R_{np} larger than 1 corresponds to an increase in non-polar surface area, and, hence, to occurence of the unfolding process.

The effect of the ice-water interface on protein conformation was determined by comparing the evolution of the non-polar area ratio R_{np} for the two simulations

(Figure 4.3a).

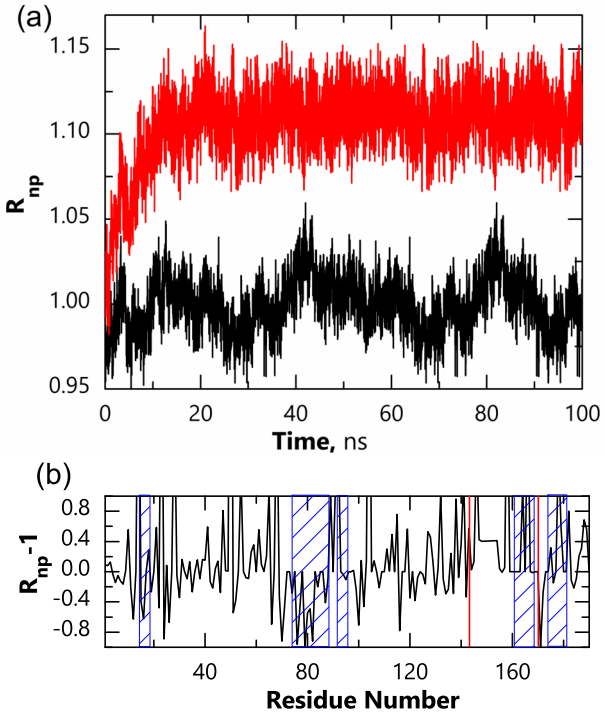

Figure 4.3: (a) Evolution of the non-polar area ratio R_{np} of the protein for the simulations employing the TIP4P water model (black curve) and the TIP4P/Ice water model (red curve). (b) Representation of the non-polar surface area per residue at the end of the simulation employing the TIP4P/Ice water model (the dominant conformation over the last 40 ns according to the Daura algorithm [187] was analyzed). Residues 143-170 are marked with a red line, while the APRs, as identified by the AMYLPRED server [197], are shaded in blue.

In presence of ice and absence of excipients there was a significant increase in the non-polar surface area of the protein. The same temperature was used in both simulations, but no unfolding was observed when the TIP4P water model was employed. Therefore, the change in protein conformation should be related to the presence of the ice-water interface, rather than to the low temperature employed.

The non-polar surface area per residue was also analyzed, for the simulation employing the TIP4P/Ice water model (Figure 4.3b). The group $R_{np} - 1$ was calculated as function of the residue number. A positive value of $R_{np}-1$ corresponds to an increase in the non-polar surface area of the residue, while a negative value indicates a reduction in the hydrophobic area exposed. It is possible to notice that residues 143-170 showed the largest increase in non-polar surface area (Figure 4.3b), with an average value of $R_{np} - 1 = 0.42$ for these residues, compared to

$R_{np} - 1 = 0.21$ on average for the rest of the protein.

This suggests that not all the amino-acid sequences of the protein are equally prone to unfolding, and that some of them are more sensitive than others to external stresses.

The idea that certain amino-acid sequences can be more important than others in determining protein behavior was already hypothesized in previous studies on protein aggregaton. For instance, it was mentioned in Chapter 3 that specific sequences of amino acids, called aggregation prone regions (APRs), tend to drive aggregation [193–196]. Because of their charge, aromaticity, hydrophobicity, average packing density, surface area and secondary structure characteristics, the APRs contribute significantly to the aggregation propensity and pathway of the protein.

Using the AMYLPRED [197] server (http://aias.biol.uoa.gr/AMYLPRED/) residues ^{14}MLRAH^{18}, ^{74}ELLRISLLLIQSWL^{87}, ^{91}QFLRSV^{96}, ^{161}YGLLYCFRK^{168} and ^{173}VETFLRIVQ^{181} were identified as APRs for hGH (shaded in blue in Figure 4.3b).

It is interesting to see that the protein sequence that unfolded during freezing (residues 143-170) contains one APR. For the other APRs, no clear trend was observed during freezing, and their non-polar surface area did not change dramatically.

In order to induce aggregation the APRs should be exposed onto the protein surface, or should become exposed upon conformational transitions. If these sequences are buried in the native state of the protein and this state is characterized by high thermodynamic stability, aggregation will be hindered. However, it seems that hGH exposes its APRs upon external stresses, such as contact with the ice-water surface. This explains the high aggregation propensity of this protein [102–105].

4.1.3 Effect of the Ice-Water Interface in Presence of Excipients

In this section a set of six simulations (number 3-8 in Table 4.2) were carried out, where freezing of hGH was investigated in presence of various excipients: sucrose, trehalose, lactose, cellobiose, sorbitol and glycine. These simulations were performed at 233 K, using the TIP4P/Ice water model to ensure freezing, and 30 molecules of excipient were inserted into the simulation box (the excipient concentration was approximately 0.24 M). The efficiency of excipients in preserving protein stability was evaluated by calculating the non-polar surface area of the protein during the simulation time.

It was observed that all the excipients under investigation prevented the increase in non-polar surface area which was observed during the simulation performed in absence of cryoprotectants. Moreover, the non-polar surface area per residue for the dominant conformation over the last 40 ns of trajectory [187] (see Figure 4.4) was analyzed. In particular, the parameter $R_{np} - 1$ was evaluated. Attention was focused mainly on residues 143-170, which were found to be the most prone to

unfold upon contact with the ice surface. These residues showed a large increase in non-polar surface area in pure water, while the increase was smaller in presence of excipients.

In order to quantify the efficiency of each excipient, the following parameter was evaluated,

$$\delta = \sum_{143}^{170}(R_{np} - 1) \tag{4.2}$$

where the sum runs over residues 143 to 170 (Table 4.3). The more δ is positive, the greater the protein tendency to unfold is. As the protein undergoes unfolding in absence of excipients, the value of δ observed in this condition (12.2) could be seen as the maximum theoretical estimate for this parameter.

According to this analysis, it seems that, in the case of hGH, sucrose should be an extremely efficient stabilizer during freezing, followed by lactose, cellobiose, glycine, sorbitol and trehalose.

In order to find an explanation to these observations, the preferential exclusion of each excipient from the protein surface was evaluated. In Figure 4.5, the parameter β, as defined in Equation 2.1, is plotted as function of the distance r from the surface of hGH (a) or from residues 143-170 (b). Also, the parameter Π, as defined in Equation 2.2, was computed, and the results are reported in Table 4.3.

Table 4.3: δ and Π parameter for each simulation

excipient	δ	Π	
		protein	res. 143-170
none	12.2	-	-
trehalose	7.5	0.81	1.68
sorbitol	5.4	1.17	0.84
glycine	3.0	1.26	1.21
cellobiose	1.6	0.74	0.44
lactose	1.0	0.76	0.71
sucrose	0.9	0.82	0.37

As regards preferential exclusion from the protein, the following "ranking" of excipients could be written, with the most excluded on the left and the least excluded on the right: cellobiose, lactose > sucrose, trehalose > sorbitol, glycine.
On the contrary, as concerns exclusion from residues 143-170, the ranking would be different: sucrose > cellobiose > lactose > sorbitol, glycine, trehalose.

This analysis suggests that the greatest stabilization of hGH (lowest δ parameter) occurred in presence of those excipients which were mostly excluded from

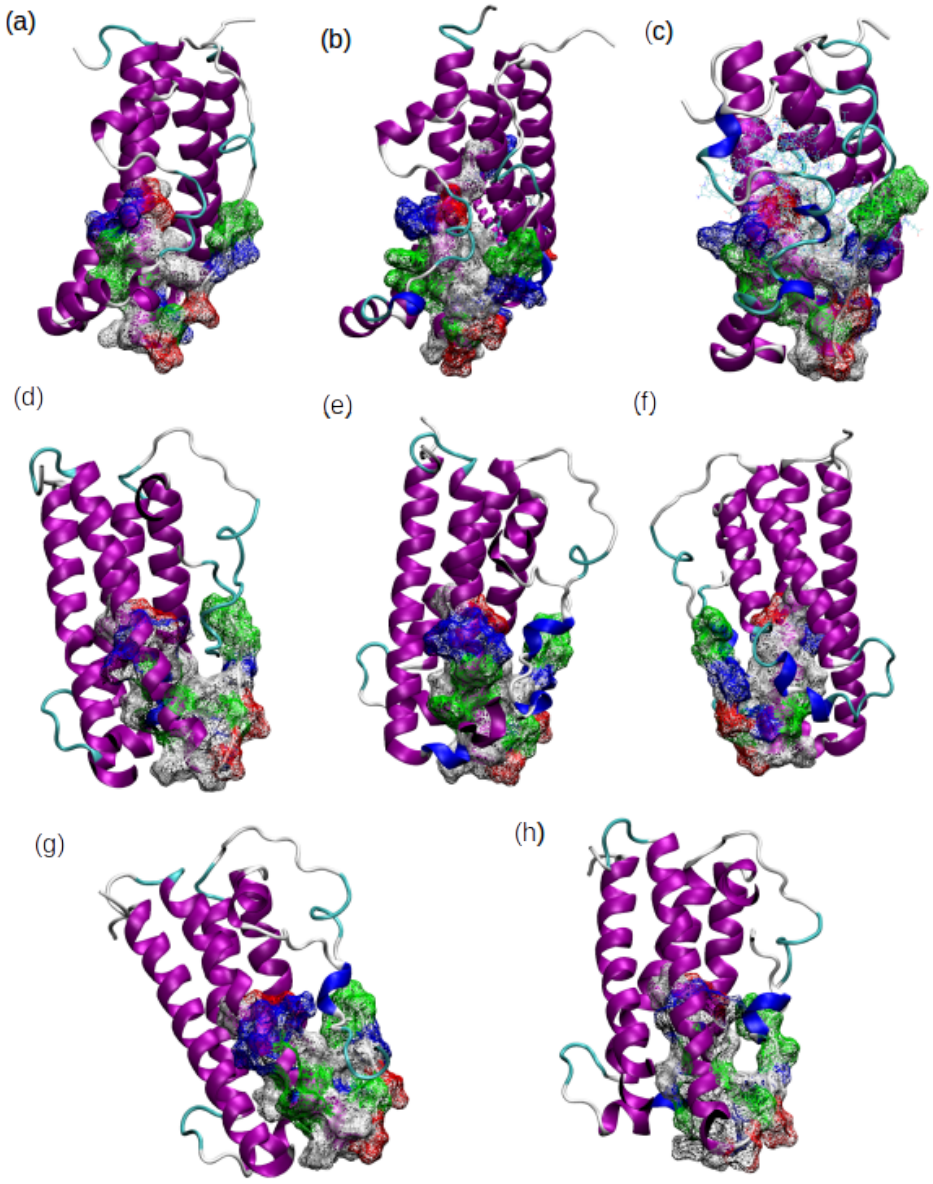

Figure 4.4: Native hGH (a) and dominant conformations over the last 40 ns [187] for simulations in ice with (b) no excipients, (c) sucrose, (d) trehalose, (e) lactose, (f) cellobiose, (g) sorbitol, (h) glycine. Residues 143-170 are illustrated in a surface representation, while the rest of the protein is displayed as a cartoon. In the surface representation, polar residues are colored in green, nonpolar residues in white, acidic residues in red and basic residues in blue. In the cartoon representation, α-helices are in purple, 3_{10}-helices in blue, turns in cyan and coils in white.

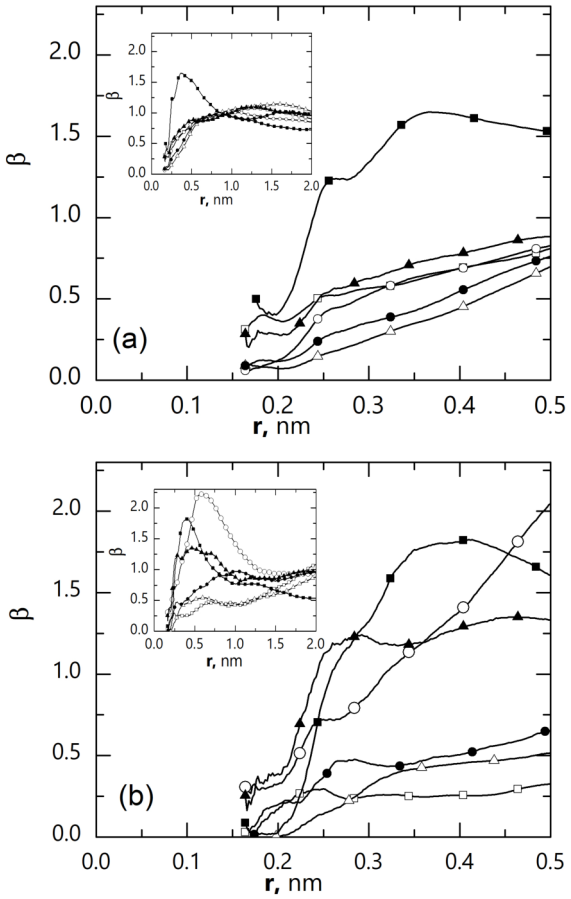

Figure 4.5: Preferential exclusion parameter β as function of the distance from the surface of hGH (a) or from residues 143-170 (b) for: □ sucrose, ○ trehalose, △ cellobiose, ■ glycine, ● lactose, ▲ sorbitol. In the main figure, the evolution of β for small values of r is shown, while the entire profile of β is illustrated in the insets.

residues 143-170, rather than from the protein as a whole. This is also evident from Figure 4.6, where the correlation between δ and Π as computed for the whole protein is very poor (R^2=0.083), while it improves significantly when Π is calculated for residues 143-170 (R^2=0.700). This result seems to indicate that the excipient interactions with specific regions of the protein may be more relevant for overall stability than the interaction with the whole molecule.

This suggests that it is of utmost importance to identify the amino-acid sequences of the protein that are more susceptible to unfolding. The interaction of the stabilizers with these sequences should then be specifically addressed to draw conclusions on the overall protein stability.

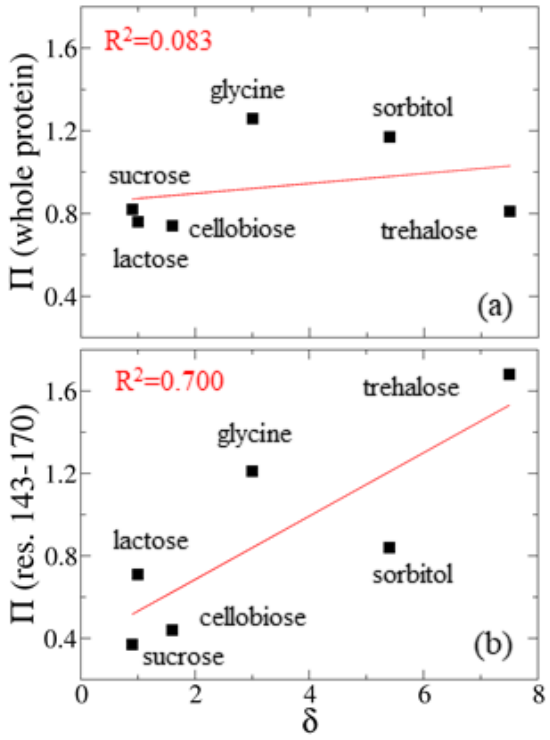

Figure 4.6: Π parameter for the whole protein (a) or for residues 143-170 (b) as function of the δ parameter. The coefficient of determination R^2 for a linear interpolation is also shown on the graphs.

In the following, a similar investigation will be performed for another model protein, namely, the IgG-binding domain of protein L.

4.2 The Ice-Water Interface Destabilizes Protein L by Enhancing Cold Denaturation Phenomena

4.2.1 Simulation Approach: Parallel Bias Metadynamics

The conformational stability of the IgG-binding domain of peptostreptococcal protein L was investigated both in bulk and at the ice-water interface, using 4-replica multiple walker [216] parallel bias metadynamics (PBMetaD) [209]. The effect of glucose on protein stability was also addressed.

The following collective variables were included,

- the distance d between the protein center of mass (COM) and the ice-water

interface. This CV was used to improve the sampling of the adsorption process at the ice surface.

- the protein α-helix (α_h) and antiparallel β-sheet (β_s) content. The native state of protein L contains both an α-helix and an antiparallel β-sheet, which may be disrupted as a consequence of unfolding. Therefore, the use of these CVs improves the sampling of partially or totally unfolded conformations. The α-helix or β-sheet content are defined as the number of 6 residue sections having an α-helical or antiparallel β-sheet configuration, according to [220]

$$\Gamma = \sum_{\mu} g[r_{dist}(\{R_i\}_{i\in\Omega_\mu}, \{R^0\})] \text{ with } \Gamma = \alpha_h,\ \beta_s \tag{4.3}$$

where the summation runs over all possible segments involved in the α-helix or antiparallel β-sheet structure, $\{R_i\}_{i\in\Omega_\mu}$ are the atomic coordinates of a set Ω_μ of 6 residues of the protein, and $g(r_{dist})$ is a switching function

$$g(r_{dist}) = \frac{1 - \left(\frac{r_{dist}}{r_0}\right)^8}{1 - \left(\frac{r_{dist}}{r_0}\right)^{12}} \tag{4.4}$$

A cutoff distance of $r_0 = 0.08$ nm was used, and r_{dist} is the distance RMSD with respect to a reference α-helix or antiparallel β-sheet configuration $\{R^0\}$.

- the distance root mean square deviation (dRMSD) for the backbone atoms with respect to a reference structure. The dRMSD was defined as function of the distances between all the pairs of atoms,

$$\text{dRMSD}(X^a, X^b) = \sqrt{\frac{1}{N(N-1)} \sum_{i\neq j} [d(x_i^a, x_j^a) - d(x_i^b, x_j^b)]^2} \tag{4.5}$$

where X^a and X^b are the two configurations to be compared, N is the number of atoms, and $d(x_i, x_j)$ represents the distance between atoms i and j. To reduce the computational time, only pairs of atoms within the range 0.1 - 3.0 nm were incorporated into the sum. The protein crystal structure was used as the reference structure.

The initial Gaussian height was set to 2 kJ/mol, the bias factor to 15, and the Gaussian deposition rate to 1 hill/ps. The Gaussian widths σ were 0.3 nm, 0.3, 0.3, and 0.02 nm for d, α_h, β_s, and dRMSD, respectively. Convergence was assessed by monitoring the fluctuation of the one-dimensional free energy profiles obtained for the different CVs during the last 10% of the simulation time (see Appendix B, Figures B.11-B.14).

More details about the simulations performed can be found in Table 4.4. In simulations 1 and 2 the bulk behavior of protein L was investigated, while simulations 3 and 4 were aimed at studying the effect of the ice-water interface. In simulations 2 and 4, the number of glucose molecules within the box was adjusted so as to give a 1 M concentration. Converged free energy surfaces are computed using the reweighting technique of Tiwary and Parrinello [222].

As a control, two unbiased simulations were performed in water and at the ice-water interface, in absence of glucose (sim. 5 and 6 in Table 4.4).

Table 4.4: Details of the PBMetaD simulations performed*.

Sim. #	Surface	# Glucose Mol.	# Liquid Water Mol.	CVs	Box Dim., nm
1	bulk	-	16733	α_h, β_s, dRMSD	8.0 x 8.0 x 8.0
2	bulk	308	14017	α_h, β_s, dRMSD	8.0 x 8.0 x 8.0
3	Ice-Water	-	18165	d, α_h, β_s, dRMSD	8.6 x 8.1 x 11
4	Ice-Water	335	15254	d, α_h, β_s, dRMSD	8.6 x 8.1 x 11
5	bulk	-	16733	unbiased	8.0 x 8.0 x 8.0
6	Ice-Water	-	18165	unbiased	8.6 x 8.1 x 11

*400 ns simulation time for sim. 1-4, 100 ns simulation time for sim. 5-6

4.2.2 Simulation Details

The simulations were performed using Gromacs 5.1.4 [198] patched with Plumed 2.4.1 [199]. The protein L topology file was obtained from the RCSB Protein Data Bank (PDB: 2PTL [246]), and modified using the software Pymol to obtain the Y43W point mutant already studied in previous experimental works [233, 234]. It was demonstrated that the Y43W mutation does not cause significant perturbations of the wild type structure [233], and was here introduced only to allow a direct comparison with experimental data. The OPLS-AA force field [93] was used, in combination with the TIP3P water model [242] (see Appendix A). For simulations of glucose, the OPLS-AA-SEI force field [247] was employed.

The GenIce algorithm [205] was used to obtain an initial configuration of hexagonal (Ih) ice with proton disorder and zero dipole moment, and the generated ice layer (8.6 x 8.1 x 2.7 nm) was oriented with the basal {0001} plane in the direction of the liquid phase (z-axis). Afterwards, the Ih ice water molecules were kept frozen in place during the simulations.

Periodic boundary conditions were used, and the cut-off radius for both Coulombic (calculated using the PME method [126]) and Lennard-Jones interactions was 1.0 nm. 1 native protein molecule was introduced into each simulation box, and its charge was neutralized using Na^+ ions. After energy minimization with the steepest descent algorithm, the system was equilibrated for 5 ns at 1 bar and 260 K in the NPT ensemble, using Berendsen pressure and temperature coupling [207]. The 4-replica multiple walker PBMetaD simulations were then performed at 260

K and 1 bar in the NPT ensemble, controlling temperature and pressure with the V-rescale thermostat [125] and Parrinello-Rahman barostat [127], respectively. 100 ns per replica were performed, and in all simulations a 2 fs time step was used.

Low Dimensional Reaction Coordinate

Recently, well-tempered metadynamics with an infrequent bias deposition rate [248, 249] was shown to give reliable residence time prediction of ligand-protein interactions [250], where the reaction coordinate was estimated using the spectral gap optimization of order parameters (SGOOP) method of Tiwary and Berne. [251]. Another study used a path CV to estimate the reaction coordinate of conformational transitions in T4 lysozyme mutants, obtaining rates again using metadynamics with an infrequent hill deposition rate [252]. These methods are based on the hyperdynamics formalism, which requires a boost potential that strictly vanishes at the transition state to ensure unbiased, ergodic sampling of the transition state region. Additionally, when the bias potential acts on a few low-dimensional order parameters as in metadynamics, the collective variables must be a good approximation of the reaction coordinate so that hidden slow degrees of freedom do not lead to orthogonal barriers which give rise to spurious rate predictions.

For obtaining kinetics a low dimensional representation of the reaction coordinate is needed, that can distinguish between meta-stable states. A path CV is constructed from representative frames extracted from the unfolding path observed during the PBMetaD simulations (sim. 1 and 3 in Table 4.4). The path CV [253] quantifies the progress s_{path} along a path connecting different conformations,

$$s_{path}(X) = \frac{\sum_{i=1}^{N_f} i \exp(-\lambda_{path} d[X - X_i])}{\sum_{i=1}^{N_f} \exp(-\lambda_{path} d[X - X_i])} \tag{4.6}$$

where N_f is the number of high-dimensional frames X_i used to describe the path, and $d[X - X_i]$ represents the RMSD distance between the istantaneous coordinate of the system, X, and the ith frame.

The parameter λ_{path} was computed as,

$$\lambda_{path} = \frac{2.3(N_f - 1)}{\sum_{i=1}^{N_f - 1} d(X_i - X_{i+1})} \tag{4.7}$$

In the case of protein L in bulk water, the path CV alone was not a sufficient approximation for the reaction coordinate as evidenced by the fact that simulations performed biasing this CV alone exhibited hysteresis. To find a better low-dimensional representation, the recent variational approach to conformational dynamics metadynamics (VAC-MetaD) approach [254] was used to find an optimal linear combination of candidate CVs. Recently, important intermediates along the conformational transformation path of the protein L99A T4 Lysozyme were identified using an optimized combination of order parameters derived from NMR and

short unbiased MD simulations [255]. Here, the path CV s_{path} defined above, the β-sheet content β_s, and the root mean square deviation dRMSD from the folded state were included. The VAC-MetaD method obtains an optimal linear combination of order parameters from an initial set of candidate order parameters $\mathbf{O}=\{O_k\}$ by solving the following generalized eigenvalue equation,

$$\mathbf{C}(\tau)\mathbf{b}_i = \mathbf{C}(0)\lambda_i(\tau)\mathbf{b}_i \tag{4.8}$$

where $\mathbf{C}(\tau)$ is the time-lagged matrix of the dynamical correlation functions $C_{j,k}(\tau) = \langle O_j(t)O_k(t+\tau)\rangle$. From a short metadynamics trajectory, the lag time τ is given by the rescaling formula $\tau = \sum_{t=t_0}^{t_f} e^{\frac{(V(\mathbf{s}(t)-c(t))}{k_B T}}\Delta t$, where $c(t)$ is a reweighting factor [222]. The computed eigenvalues $\lambda_i(\tau)$ provide a natural way to identify the slowest modes of the system, as a large eigenvalue corresponds to a slow relaxation time. Therefore, the eigenvectors of the highest eigenvalues, which can be approximated as $s_i = \sum_k b_{ik}O_k$, may be used as improved CVs. In our specific case, a gap was observed between the highest $\lambda(\tau)$ and the others, and the eigenvector associated with this highest eigenvalue $s = 0.0323\beta_s - 0.9942dRMSD + 0.1030s_{path}$ was used as optimized CV for the frequency adaptive metadynamics simulations.

In bulk water, a contact map CV Q was also used. Q is defined using a list of native contact pairs i, j. All pairs of Cα atoms i and j in the native fold are said to be in contact if the distance between i and j is less than 0.5 nm. Then, Q was defined as,

$$Q(X) = \sum_{(i,j)} \frac{1-\left(\frac{r_{i,j}(X)}{r_0}\right)^6}{1-\left(\frac{r_{i,j}(X)}{r_0}\right)^{12}} \tag{4.9}$$

A cutoff distance of $r_0 = 0.5$ nm was used and $r_{ij}(X)$ is the distance between i and j in configuration X.

Frequency Adaptive Metadynamics

Frequency adaptive metadynamics (FaMetaD) [256] is a variant of the metadynamics enhanced sampling method where the bias deposition frequency is adjusted during the simulation time. In FaMetaD the free energy basins are filled up quickly at the beginning of the simulation, but the bias deposition frequency is then reduced as the system moves close to the transition state region, so as to avoid biasing the free energy barriers. This makes it possible to efficiently speed up the transition, in a way that minimally corrupts its dynamics.

The deposition frequency τ_{dep} is adjusted according to the following scheme,

$$\tau_{dep}(t) = \min\{\tau_o \cdot \max\{\frac{\alpha(t)}{\theta}, 1\}, \tau_c\} \tag{4.10}$$

where τ_0 is the initial deposition time and τ_c is the infrequent deposition stride used as the system moves close to the transition state. Values of 2 ps and 1 ns, respectively, were selected in this work for the minimal and maximal bias deposition frequency. $\alpha(t)$ is the acceleration factor,

$$\alpha(t) = \langle e^{\frac{V(\mathbf{s},t)}{k_B T}} \rangle \tag{4.11}$$

where $V(\mathbf{s}, t)$ is the bias potential and the angular brackets indicate the time average until the simulation time t. To avoid excessive fluctuations in $\tau_{dep}(t)$ caused by changes in $\alpha(t)$, Eq. 4.10 can be rewritten to,

$$\tau_{dep}^{'}(t) = \max(\tau_{dep}((N_s - 1)\Delta t_{MD}), \tau_{dep}(N_s \Delta t_{MD})) \tag{4.12}$$

to yield a monotonously increasing function. Here, N_s is the current step in the metadynamics run with time step Δt_{MD}.

The threshold value θ in Eq. 4.10 determines the gradual transition from normal to infrequent metadynamics, and values of 10^5 and 10^7 were selected in this work for simulations at the ice surface or in bulk water, respectively.

For the case of protein L at the ice-water interface FaMetaD was performed biasing the path CV. In bulk water, protein L exhibits more stability and both the contact map CV Q and the optimized CV obtained from VAC-MetaD were biased. The Gaussian heights and bias factors were set to 0.4 kJ/mol and 15, respectively. The σ (Gaussian width) values used were 0.6 for the path CV at the ice interface, or 0.05 and 1 for the optimized and contact map CVs in bulk water, respectively. 25 independent trajectories for each system (bulk water and ice surface), with different initial velocities, were run in order to get some statistics on the transition event. Gromacs 2018.6 and Plumed 2.5.1 were used for these FaMetaD simulations. The transition times obtained were then reweighted to correct for the bias potential, according to

$$t^* = \Delta t_{MD} \sum_i e^{\frac{V(\mathbf{s},t)}{k_B T}} \tag{4.13}$$

where t^* is the "real" time in an unbiased simulation.

4.2.3 Protein L is Destabilized by the Presence of Ice, while Glucose Has a Cryoprotective Effect Both in the Bulk and at the Ice Surface

Using PBMetaD simulations, the free energy surfaces (FES) for protein L were computed, both in the bulk and in the presence of the ice-water interface. The presence of glucose as model cryoprotectant was also considered. This particular excipient was selected because experimental data are available on its effect on

protein L conformational stability [235]. The results of this analysis are shown in Figure 4.7, where the FES as function of antiparallel β-sheet content (β_s) and radius of gyration R_g or α-helix content (α_h) and dRMSD are shown in the left and middle panel, respectively.

The three most sampled conformations were identified in each FES, and labeled with a letter followed by a number. The letter N corresponds to the most folded structure, while letters A and B were used to identify two partially folded conformations, where either the α-helix content (conformation A) or the antiparallel β-sheet content (conformation B) have been lost to some extent. The number in each label is then used to distinguish between the different simulations, as listed in the first column of Table 4.4.

The relative contribution in the FES as a percentage of each of these conformations was also computed as,

$$P_X = \frac{\int_X ds e^{-\frac{F(s)}{k_B T}}}{\int ds e^{-\frac{F(s)}{k_B T}}} \tag{4.14}$$

where $F(s)$ is the free energy, and the integral in the numerator is over a subset of the volume that defines basin X = N, A or B. The calculated percentages are also displayed under the cartoon of each structure in the right panel of Figure 4.7.

This analysis showed that both the β-sheet and α-helix content could be lost upon unfolding in bulk water, even though the native fold remained by far the most probable ($> 99.9\%$). For instance, structure A1 ($\approx 0.065\%$ probability) showed values of α_h as low as 2, and conformation B1 ($\approx 0.0004\%$ probability) was characterized by $\beta_s = 4$, compared to values of 15 and 9.5 for α_h and β_s, respectively, in structure N1 (Figure 4.7a).

Upon addition of glucose (Figure 4.7b), the FES was restricted toward more folded conformations (the relative contribution of unfolded conformations was on the order of $10^{-6}\%$ or $10^{-10}\%$ for A2 and B2, respectively). In this case, the β-sheet structure was remarkably stabilized, with conformation B2 still showing a significant β-sheet content ($\beta_s = 7.6$). An almost complete loss of the α-helix content was still possible ($\alpha_h = 2$ for structure A2), but the free energy barrier for this unfolding process was significantly higher than in bulk water. The stabilizing effect of glucose observed in our simulations is in line with the experimental results of Plaxco and Baker [235], where the addition of 1M glucose was found to produce a 1 kcal/mol increase in the free energy of unfolding for protein L.

In contrast, the free energy barrier of unfolding was reduced by the presence of the ice-water interface, as shown in Figure 4.7c. The folded conformation N3 could in fact lose its β-sheet (β_s=6.5 in structure B3, having probability 8.27 %) or α-helix content (α_h=5 in structure A3, having probability 6.52 %) almost without free energy penalty.

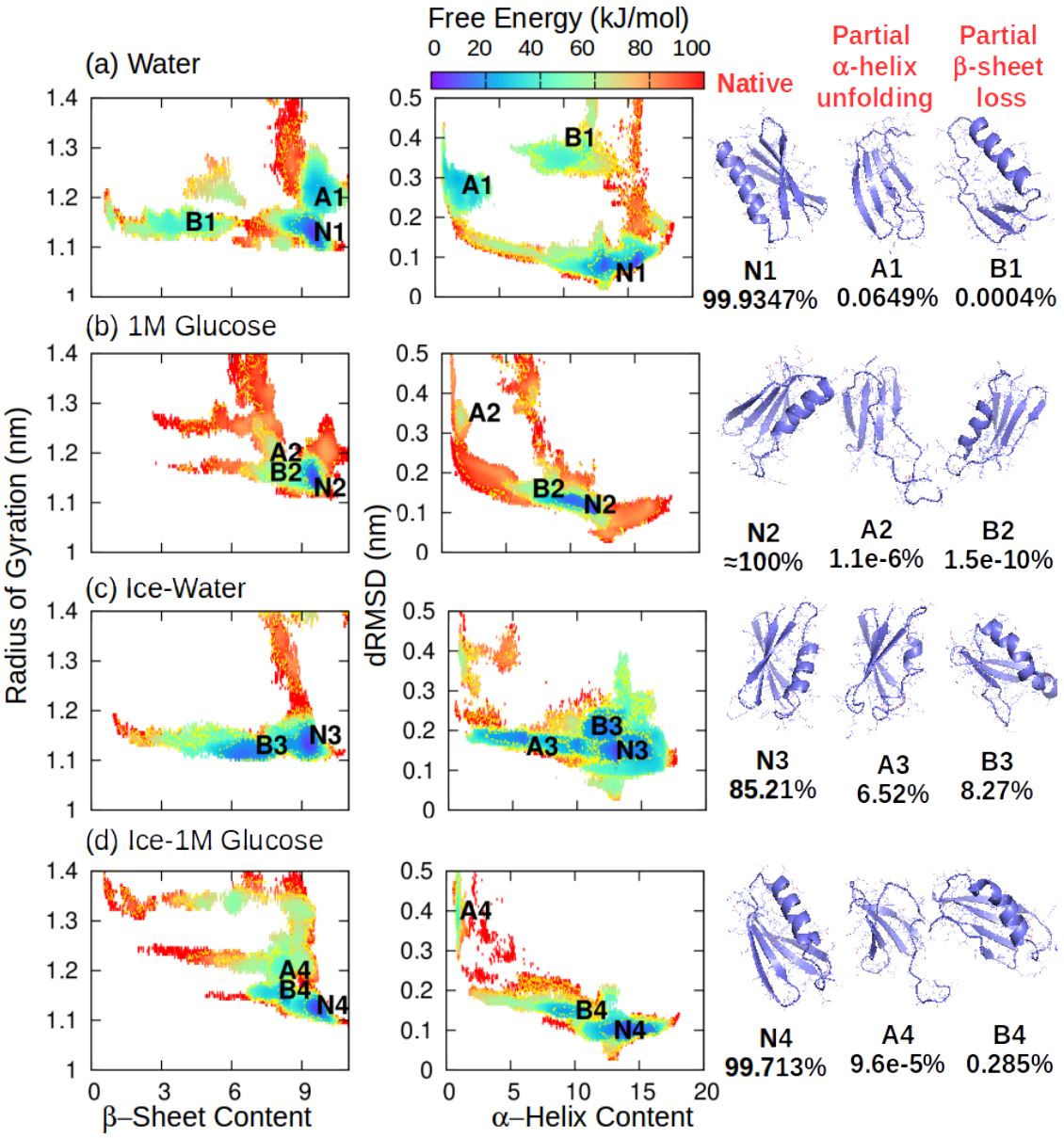

Figure 4.7: Free energy surface (FES) as a function of β-sheet content and radius of gyration (left) or α-helix content and dRMSD (center) for (a) bulk-water, (b) 1M glucose, (c) ice-water interface and (d) ice-1M glucose interface. The letters on the FES identify the most sampled protein conformations. A cartoon of each structure, with the relative contribution as a percentage of these conformations, is also shown on each line in the right-hand panel.

Finally, Figure 4.7d shows that glucose has a stabilizing effect also at the ice-water surface, again restricting the FES toward more folded conformations and

significantly hindering the loss of β-sheet content (β_s=8 in structure B4, characterized by 0.285 % probability). The loss of α-helix content is still possible in the presence of glucose (α_h=1 in structure A4, that shows a low probability, on the order of 10^{-4}%), but a quite large energy barrier characterizes this unfolding process. Interestingly, in bulk water, either in presence or absence or glucose, the unfolded structure having reduced α-helix content shows a higher relative contribution than the conformation with partial β-sheet loss, while the opposite occurs when the ice-water interface is considered. This suggests that the denaturing effect of the ice surface is mostly targeted at the β-sheet structure of protein L.

It is interesting and important to note that a completely unfolded structure was never sampled in our simulations. Each protein conformation still preserved some secondary structure, either α-helix or β-sheet, and a large increase in the radius of gyration was never observed. For instance, the folded structure was characterized by $R_g \approx 1.1$ nm, and values of R_g larger than 1.3 nm were rarely sampled. This last observation may be explained considering that the simulations were performed at low temperature (260 K) with a relatively low metadynamics bias factor relative to the energy barrier associated with unfolding. In these conditions, cold denaturation typically ensues. As previosuly discussed, heat denaturation is favored by the increase in conformational entropy as the protein unfolds to a largely extended conformation, while cold denaturation is enthalpically driven. As a result, cold-denatured proteins are more compact, partially unfolded conformations, showing a mild form of structural loss [19,225–227]. It is interesting to note that the presence of the ice-water interface does not alter the typical features of cold denaturation, leading to the formation of compact, partially folded conformations.

The protein segments, which were mostly involved in the unfolding process, were also investigated. In order to do this, the backbone RMSD of the protein conformations sampled during the simulations was computed, using N1 as reference structure (see Figure 4.8).

The largest RMSD was generally observed for the partially folded conformations of type A, suggesting that an extended protein conformation is originated when the α-helix content is lost. Moreover, this analysis showed that residues ^{13}TQTAEFKGTFEKATSEAY^{30} generally were the most prone to undergo the unfolding process. This segment includes one β-strand (residues ^{13}TQTAEF^{18}), and a large part of the α-helix (residues ^{23}EKATSEAY^{30}). However, in the presence of ice (conformation B3), the amino acid sequence ^{2}IKANLIFANGSTQTA^{16} was also significantly involved in the loss of secondary structure. Therefore, in this case the β-strand ^{2}IKANLI^{7} was mainly disrupted, in line with the previous observation that the effect of the ice-water interface is mainly targeted at the β-sheet structure of the protein. Finally, residues ^{33}ADTLKKDNGEWT^{44} showed a quite large RMSD when glucose was added to the simulation box. This was true not only for the unfolded conformations B2 and B4, but also for the folded structures N2 and N4.

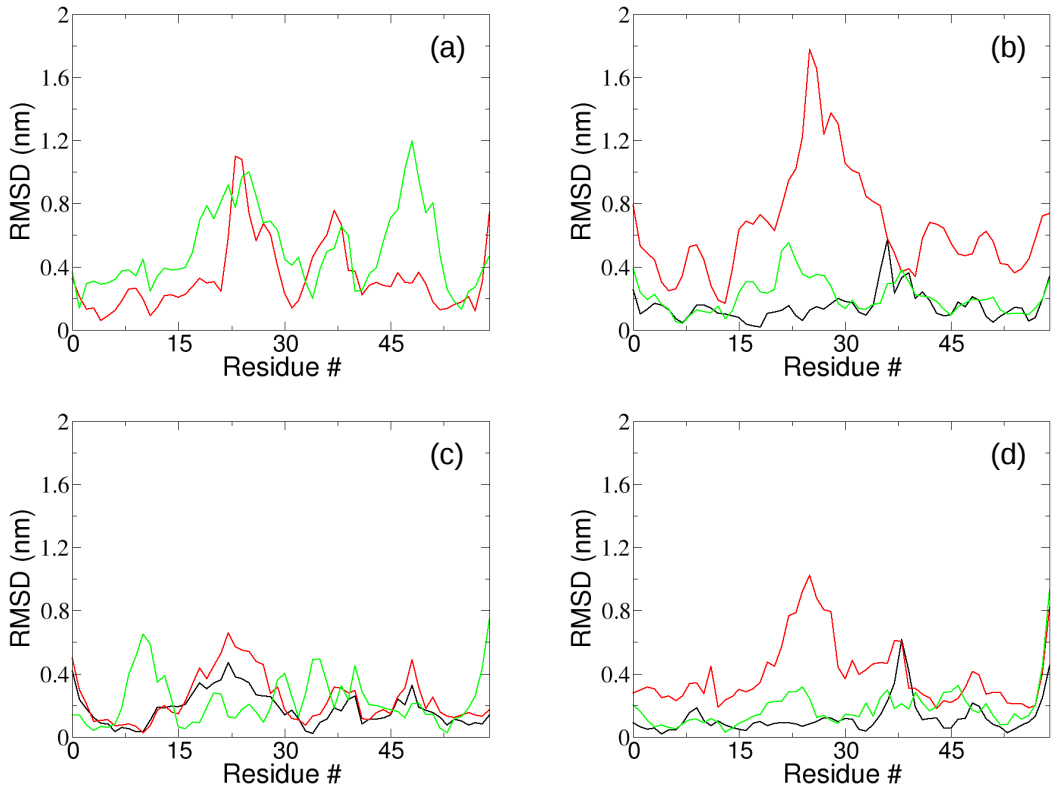

Figure 4.8: Backbone RMSD for the protein conformations N (black line), A (red line) and B (green line) sampled (a) in bulk water (sim. 1 in Table 4.4), (b) in 1M glucose (sim. 2 in Table 4.4), (c) at the ice-water interface (sim. 3 in Table 4.4), and (d) at the ice-water interface in presence of 1M glucose (sim. 4 in Table 4.4). The protein structure N1 (native state in bulk water) was used as reference.

This may indicate that the presence of glucose promotes the sampling of protein structures where this amino acid sequence is more expanded than in pure water. To further confirm this observation, structures N1, N2 and N4 were aligned using the STAMP (Structural Alignment of Multiple Proteins) program [257]. STAMP alignes protein structures by applying optimal rigid-body rotations and translations in order to minimize the C_α distance between corresponding residues of each conformation. In Figure 4.9a the 3 superimposed structures are identified by different colors, and residues 33-44 are highlighted in yellow. In Figure 4.9b the same aligned structures are shown, but the coloring method allows identification of the zones (displayed in red) where the conformations are structurally different. Comparison of Figures 4.9a and 4.9b confirms that residues 33-44 are the most interested in expansion in presence of glucose, even though the structural modification is not

dramatical.

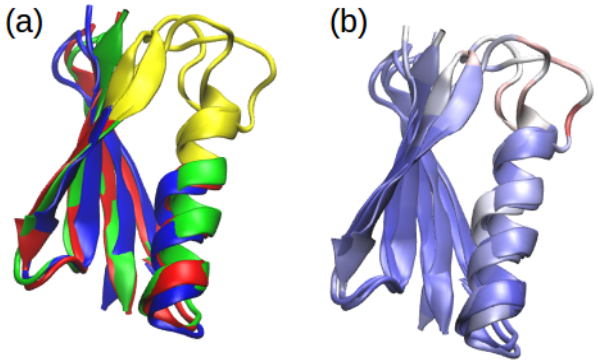

Figure 4.9: (a) N1 (red), N2 (blue) and N4 (green) structures after alignment with the STAMP program [257]. Residues 33-44 have been highlighted in yellow. In (b), the same image is shown, but a different coloring is used, where the blue areas correspond to structurally similar zones, and the red color indicates poor alignment.

4.2.4 The Ice Surface Promotes Loss of Protein Structure by Enhancing Cold Denaturation Phenomena

Having shown that the ice surface has a destabilizing effect on the protein structure, lowering the free energy barrier for the unfolding process, the interaction of protein L with the ice surface was further analyzed, with the aim of unveiling the presence of adsorption phenomena.

Interestingly, this analysis revealed that no direct interaction occurred between protein L and the ice-water surface. For instance, Figure 4.10a shows that the minimum distance between the protein and the ice interface always remained quite large. This was true for both the whole protein (black bars), and those amino-acid sequences which were the most prone to unfold according to the RMSD analysis (residues 13-30 in red, and residues 2-16 in green).

Also the number of protein-ice hydrogen bonds (Figure 4.10b) remained low during the simulation, with an average value of 0.163. This means that the direct interaction between the protein and the ice surface is not responsible for the observed destabilization of the protein structure. Strambini and Gabellieri [20] hypothesized that the denaturing effect of ice may be related to direct adsorption of the protein onto the ice surface. However, solid-state NMR studies suggest that the hydration shell of soluble proteins does not freeze below the freezing temperature of the bulk solution [258, 259]. Siemer *et al.* [260] observed that ubiquitin at -35 °C keeps its entire hydration shell, which prevents interaction with the ice lattice. The authors further suggest that most soluble proteins are likely to behave

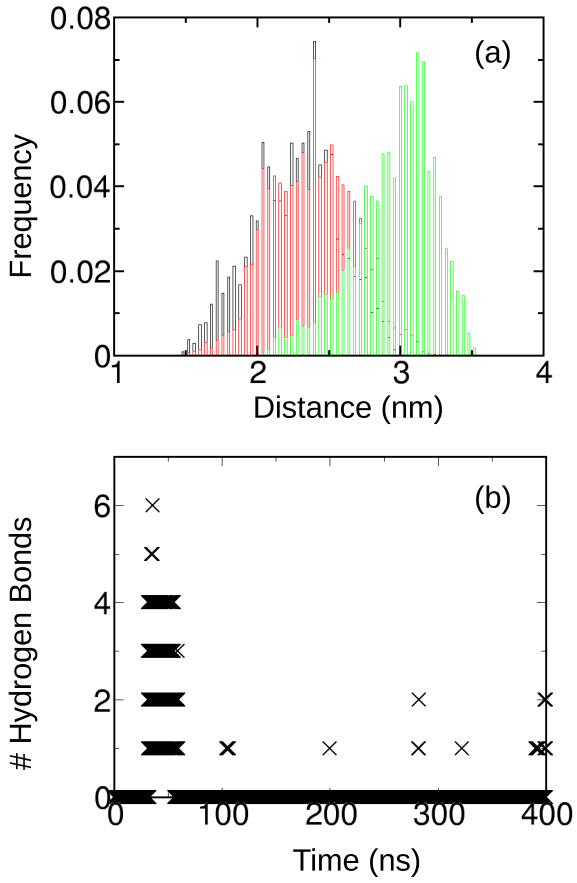

Figure 4.10: (a) Histogram showing the distribution of the minimum distance between the ice-water interface and the whole protein (black bars), residues 13-30 (red bars) and residues 2-16 (green bars). (b) Number of proten-ice water hydrogen bonds as function of the simulation time. The results refer to simulation 3 in Table 4.4.

like ubiquitin, and that their hydration shell does not freeze until a temperature which is much lower than the equilibrium freezing value [261]. Above this temperature, according to NMR measurements, no direct interaction is possible between the protein and the ice surface. More recently [262], X-ray diffraction studies of protein/ice interaction further suggested that two typical pharmaceutical proteins, recombinant human albumin and a monoclonal antibody, interact with ice crystals indirectly, by accumulating in the liquid-like layer above the ice surface, rather than by direct adsorption, in line with our simulations. A different behavior is generally observed only in presence of antifreeze proteins (AFPs), that can directly bind to ice nuclei and prevent them from growing [262–264].

The diffusion coefficient of the liquid water molecules in simulations 5 and 6 was also calculated, by least squares fitting a straight line through the mean square displacement as function of time. According to the Einstein relation, the slope of this line should be directly related to the diffusion coefficient. It was found that the diffusion coefficient was $2.9138 \pm 0.0123 \cdot 10^{-5}$ cm^2/s in absence of ice (simulation 5), while it decreased to $1.8733 \pm 0.0084 \cdot 10^{-5}$ cm^2/s when an ice layer was added to the simulation box (simulation 6). Therefore, the presence of the ice surface slows down the nearby layers of liquid water molecules. It was further verified whether this also translated into increased ordering of the water molecules. The tetrahedral order parameters S_g and S_k [265] were computed for simulations 5 and 6,

$$S_g = \frac{3}{32} \sum_{j=1}^{3} \sum_{k=j+1}^{4} (\cos \psi_{j,k} + 1/3)^2 \tag{4.15}$$

$$S_k = \frac{1}{3} \sum_{k=1}^{4} \frac{(r_k - \bar{r})^2}{4\bar{r}^2} \tag{4.16}$$

where $\psi_{j,k}$ is the angle between the jth and kth bonds in the tetrahedral arrangement, r_k the distance between the central atom and the kth atom and $\bar{r}$ the arithmetic mean of the four radial distances. Both 3/32 and 1/3 are normalization factors that guarantee $0 \leq S_g, S_k \leq 1$. For a perfect tetrahedron, S_g and S_k equal 0, while their value increases as the configuration deviates from tetrahedrality. In Figure 4.11 the average value of these order parameters as function of the z coordinate over the equilibrated trajectory (last 50 ns) is shown. In the case of simulation 5 (bulk water, red curve) both order parameters are significantly different from 0, indicating, as expected, absence of any ordering. For instance, S_g is close to 0.25, which is the expected value for randomly arranged bonds. In the case of simulation 6 (presence of an ice layer, black curve) both S_g and S_k are 0 in correspondence of the ice layer ($0 < z < 2.7$ nm), but their value in the liquid layers above the ice surface is similar to the case of bulk water (simulation 5), indicating that these layers do not show any significant tetrahedral arrangement.

Additionally, we computed the average number of hydrogen bonds between the liquid water molecules and the whole protein, or those regions that were more prone to unfold (residues 13-30 and 2-16). As illustrated in Figure 4.12, it was found that the degree of protein-solvent interaction increased in presence of ice, with the number of hydrogen bonds moving from 155 ± 6 (sim. 5) to 164 ± 7 (sim. 6) for the whole protein, from 51 ± 3 (sim. 5) to 57 ± 4 (sim. 6) for residues 13-30, and from 24 ± 2 (sim. 5) to 31 ± 3 (sim. 6) for residues 2-16. This last region of the protein includes the β-strand ^{2}IKANLI^7, and was significantly involved in the loss of secondary structure only in presence of ice. Correspondingly, it was observed that the number of hydrogen bonds formed by these residues with liquid water not only increased upon ice formation, but these bonds were also significantly

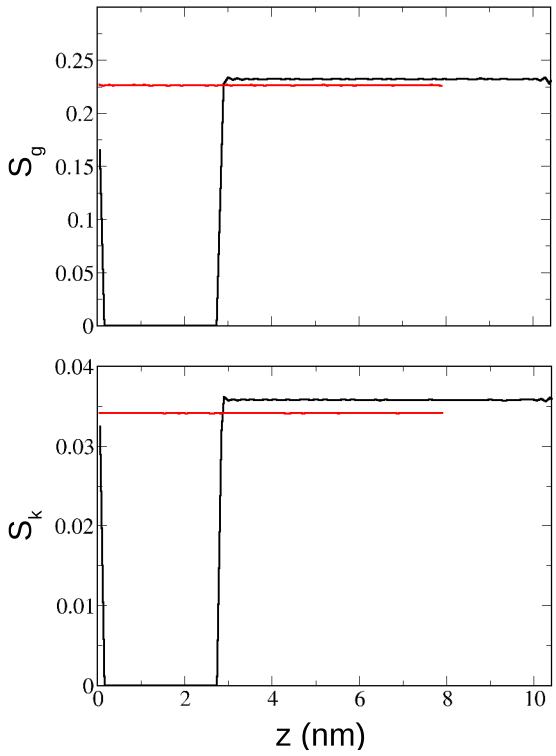

Figure 4.11: Values of the tetrahedral order parameters for water molecules S_g and S_k as function of the z coordinate in the simulation box. The red and black curves correspond to simulations 5 (bulk water) and 6 (presence of an ice layer) in Table 4.4, respectively.

stronger. The average hydrogen bonds lifetime [266–268], that is a good indicator of the hydrogen bond strength, was also computed. It was observed that the liquid water molecules in simulation 6, slowed down by the presence of ice, could form hydrogen bonds with residues 2-16 that lasted, on average, 25 ps, i.e., around 13 times larger than in pure water (about 1.9 ps in simulation 5).

This analysis suggests that the destabilizing effect of the ice interface should not be related to direct adsorption onto the surface, but should rather be mediated by a modification of the nearby liquid water behavior toward the protein. Within this framework, the effect of the ice surface could be explained as an enhancement of the cold denaturation phenomena. As previously mentioned, cold denaturation is the result of an increased tendency for nonpolar group hydration [15, 18]. The free energy penalty for the interaction between water and the hydrophobic patches of the protein becomes smaller as the temperature is decreased, leading to the observed loss of structure. In this work it was observed that, in the presence of

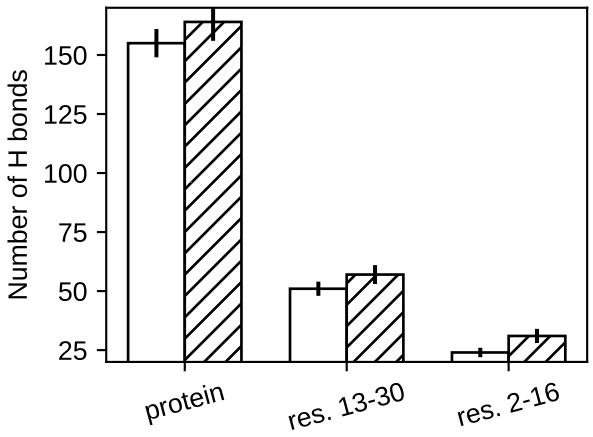

Figure 4.12: Bar graph showing the number of hydrogen bonds between liquid water molecules and the whole protein, residues 13-30 or residues 2-16 in the case of bulk water (plain bars) or at the ice-water interface (striped bars). The results refer to simulations 5 and 6 in Table 4.4.

ice, the nearby liquid water molecules did not arrange in an ordered structure, but were anyway remarkably slowed down, and could therefore form a large number of strong hydrogen bonds with specific regions of the protein. These regions show a slightly larger nonpolar surface area (the ratio of the nonpolar to total surface area S_{np}/S is 0.53 and 0.52 for residues 13-30 and 2-16, respectively) if compared to the protein average ($S_{np}/S = 0.49$). As a result, these amino-acid sequences, poorly hydrated in the folded structure, interact more favorably with water in the presence of ice, assuming an extended conformation. The ice surface therefore promotes the solvent penetration of poorly hydrated regions in the folded structure, which is a characteristic feature of the cold denaturation process.

4.2.5 Glucose Stabilizes the Native Structure by Being Preferentially Excluded from Specific Regions of the Protein

The mechanism of protein stabilization by glucose was further analyzed. According to the preferential exclusion mechanism [32, 33, 129], the protective osmolytes should stabilize the native state by being preferentially excluded from the protein surface. A possible way to quantify differences in the degree of preferential exclusion from specific patches on the protein surface is to compute the following relative distribution of glucose molecules,

$$\text{Relative distribution of glucose molecules} = \frac{(n_g(1\text{ nm})/n_{all}(1\text{ nm}))_1}{(n_g(1\text{ nm})/n_{all}(1\text{ nm}))_2} \qquad (4.17)$$

where $n_g(1\text{ nm})$ and $n_{all}(1\text{ nm})$ are the coordination number of glucose and water + glucose molecules, respectively, at 1 nm from the surface of patch 1 (numerator) or 2 (denominator). A value of the relative distribution parameter larger than 1 indicates that glucose interacts more with region 1 than with region 2, and vice versa.

The relative distribution parameter of glucose molecules for residues 33-44 (patch 1) over residues 2-30 (patch 2) is shown in Figure 4.13a. Figure 4.13a reveals that glucose interacted with different regions of the protein to a different extent. The relative distribution parameter showed an average value of 1.134, indicating that glucose was attracted by residues 33-44, while it interacted less favorably with residues 2-30. Interestingly, the protein region including amino acids 33-44 was more expanded upon addition of glucose than in bulk water, as evidenced by our RMSD and alignment analyses (Figures 4.8 and 4.9). In contrast, a 1M glucose concentration promoted a more collapsed conformation of residues 2-30, which were the most prone to undergo cold denaturation.

According to these results, it is possible to conclude that glucose stabilizes the native fold by being excluded from the hydrophobic, unfolding-prone regions of the protein. However, glucose promotes a slightly different native state, because of preferential interaction with residues 33-44. The glucose molecules may be particularly attracted by this amino-acid sequence because of its hydrophilicity ($S_{np}/S = 0.43$ for residues 33-44, compared to an average value $S_{np}/S = 0.49$ for the protein and $S_{np}/S = 0.53$ for residues 2-30).

It was also suggested [20] that polyols and sugars may stabilize the native fold by coating the surface of ice and decreasing the adsorption affinity of the protein. However, a direct adsorption of protein L to the ice-water interface was not observed in our simulations (Figure 4.10). Moreover, Figure 4.13b indicates that glucose was not attracted to the ice surface, and could therefore not form a coating of adsorbed molecules at the ice-water interface. The stabilizing mechanism of glucose is therefore not related to a decreased protein adsorption, at least in the case of protein L. In contrast, the protein approached the ice-water interface more closely in 1M glucose (Figure 4.13c) than in pure water (Figure 4.10a).

The observation that the preferential exclusion from specific amino-acid sequences is key for protein structural preservation is in line with previous observations for hGH discussed in this Chapter, and will be further investigated in the last section of this chapter, where the preferential exclusion mechanism will be reinterpreted as a variant of the bad solvent theory for polymers. However, the kinetic behavior of protein L in bulk water, or at the ice surface, will first be investigated.

4.2.6 Kinetic Analysis

The PBMetaD simulations revealed that the unfolding process was thermodynamically promoted by the ice-water interface. In order to characterize the effect of the

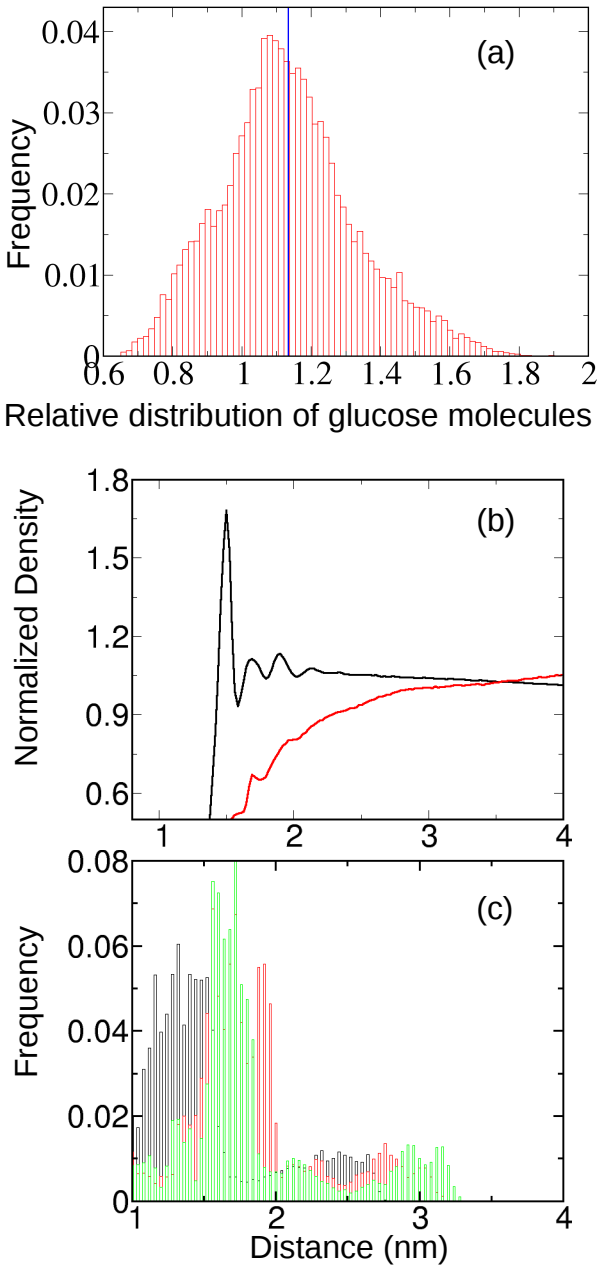

Figure 4.13: Relative distribution parameter of glucose molecules for residues 33-44 (patch 1) over residues 2-30 (patch 2). The average of the distribution is displayed as a vertical blue line in the graph. (b) Density of water (black line) and glucose (red line) as function of the distance from the ice-water interface. The density was normalized to the bulk value. (c) Histogram showing the distribution of the minimum distance between the ice-water interface and the whole protein (black bars), residues 13-30 (red bars) and residues 2-16 (green bars). The results refer to simulation 4 in Table 4.4.

ice surface also from a kinetic point of view, the results from the frequency adaptive metadynamics (FaMetaD) simulations described above were analyzed.

An illustration of the unfolding path followed in the case of bulk water, and at the ice water surface, is shown in Figure 4.14.

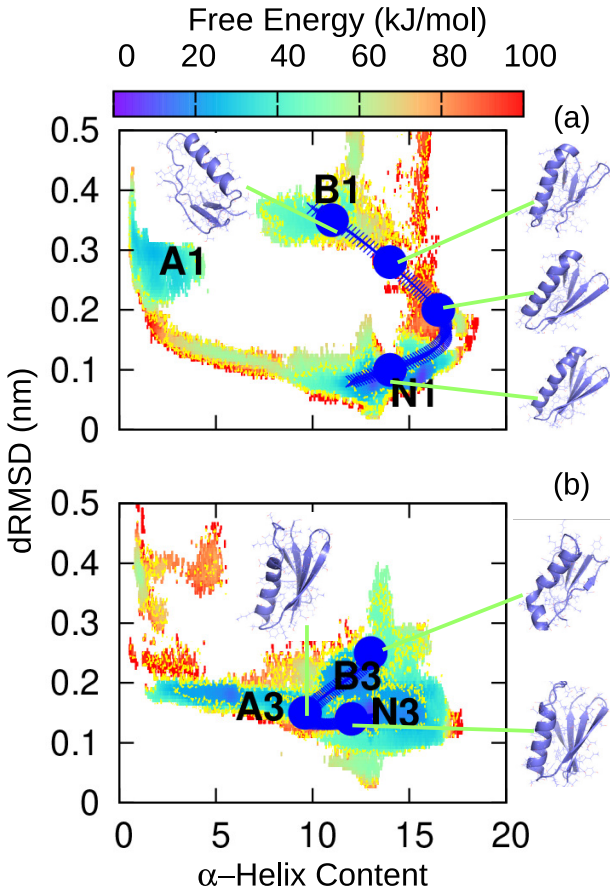

Figure 4.14: Representation of the unfolding path used for the kinetic analysis in the case of (a) bulk water and (b) the ice-water surface. Cartoon representations of selected conformations sampled during the unfolding process are also shown in each panel.

In both cases, the unfolding path connected the most folded structure (N1 and N3, respectively) to the partially folded conformations with reduced β-sheet content (B1 and B3, respectively). The set of intermediate structures connecting these end points was extracted from the unfolding path observed during the PBMetaD simulations (sim. 1 and 3 in Table 4.4). In bulk water (Figure 4.14a), the unfolding transition was characterized by an initial increase in α-helix content and decrease in β-sheet structure, followed by a second phase in which both secondary structures

were partially disrupted. For protein L in bulk water, we used both a native contact map CV and an optimized CV consisting of a linear combination of β-sheet content, dRMSD and path CV, selected using the VAC-MetaD approach, to speed up the crossing of energy barriers. In contrast, at the ice surface the path CV alone (Figure 4.14b) was found to efficiently promote the desired conformational changes, where the α-helix structure was first partially disrupted, and subsequently recovered, while the β-sheet content monotonically decreased during the whole unfolding process.

The average time required for the unfolding process in the presence of ice was computed, and compared to the case of bulk water. In order to obtain some statistics on the unfolding transition, 25 independent simulations, with different initial velocities, were performed for each system. The computational cost was 0.875 μs for simulations in bulk water, and 0.512 μs for simulations at the ice surface. A bootstrap analysis with 10000 subsamples was used to estimate the error, expressed as 95 % confidence interval.

Among the collected trajectories, one outlier, showing an unfolding time significantly longer than average, was observed both in bulk water and at the ice surface. In the outlying trajectory observed in bulk water, protein L sampled intermediate states with reduced β-sheet content several times before finally unfolding, and the first part of this last transition was corrupted by the deposition of the metadynamics bias. At the ice-water interface, some bias was deposited on the outlying trajectory during the last part of the unfolding process (the transition from dRMSD = 0.15 nm to dRMSD = 0.21 nm in Figure 4.14b). The transition dynamics was probably perturbed by the addition of this bias, resulting in an over-estimate of the unfolding time from these trajectories. Neglecting these outliers, i.e. considering the remaining 24 trajectories, mean unfolding times of 4 $\pm$ 2 s and 0.02 $\pm$ 0.01 s in bulk water and at the ice-water interface were obtained, respectively. The Kolmogorov-Smirnov (KS) test was used to verify the reliability of the computed unfolding times, by fitting the obtained values to the theoretical cumulative distribution function for a Poisson process [249]. When the selected bias is too high, or the choice of CVs is not optimal, the transition state is perturbed during the simulations, and a low p-value (typically < 0.05) is generally observed. Here, p-values of 0.18 and 0.37 were observed in bulk water and at the ice surface, respectively. If the outliers are also considered, the mean unfolding times are not dramatically affected (7 $\pm$ 5 s or 0.03 $\pm$ 0.02 s in bulk water and at the ice-water interface, respectively), but the p-values become significantly lower than 0.05 (3×10^{-5} in bulk water and 0.002 in presence of ice).

According to our results, the unfolding process is faster at the ice surface than in bulk water by two orders of magnitude (see Figure 4.15), even though the presence of ice slows down the motion of the surrounding water molecules. It was previously observed that the ice surface promotes a significant reduction in the free energy of unfolding (Figure 4.7). The FaMetaD simulations further show that this lowering of the energy barrier translates into an extremely fast denaturation process. Hence,

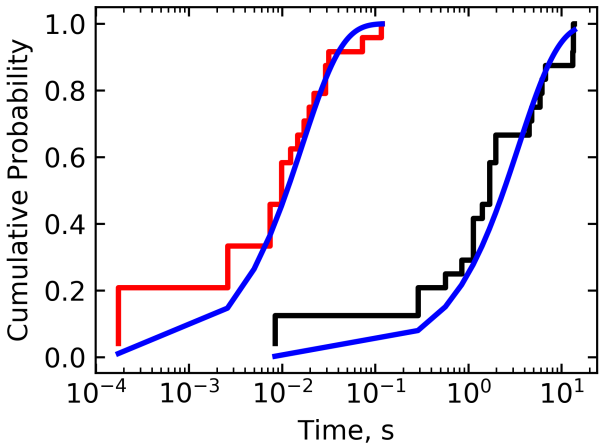

Figure 4.15: Cumulative distribution of unfolding times for protein L in bulk water (black curve) and at the ice-water interface (red curve), as obtained neglecting the outlying trajectory. The fits to an exponential distribution are shown in blue.

the process of ice formation represents a critical destabilizing factor for protein stability, from both a thermodynamic and a kinetic point of view.

Overall, a significant reduction in the free energy of unfolding at the ice surface was observed for protein L, which translated into a fast denaturation process. The simulations suggest that the ice interface modifies the properties of the nearby liquid water molecules, slowing down their motions and promoting the hydration of the nonpolar groups of the protein. The solvent penetration of nonpolar regions is a characteristic feature of cold denaturation, which seems to be dramatically enhanced in the presence of the ice-water interface. In this framework, glucose seems to counteract the ice-induced unfolding process by being preferentially excluded from the hydrophobic patches of the protein.

4.3 Preferential Exclusion Mechanism or "Bad Solvent Theory"?

As mentioned, the preferential exclusion mechanism is the most well-known theory to explain protein stabilization by osmolytes in the liquid state. According to this theory, the addition of excipients which are excluded from the protein surface should stabilize the native state by increasing the free energy of unfolding, i.e., the energy required for the unfolding transition. More specifically, the preferential exclusion hypothesis claims that a good stabilizer should be more excluded from the unfolded state than from the native fold.

Let us suppose that we have a protein in water and let us imagine an ideal experiment in which a stabilizing osmolyte, that is excluded from the protein surface, is added to the protein solution (see Figure 4.16).

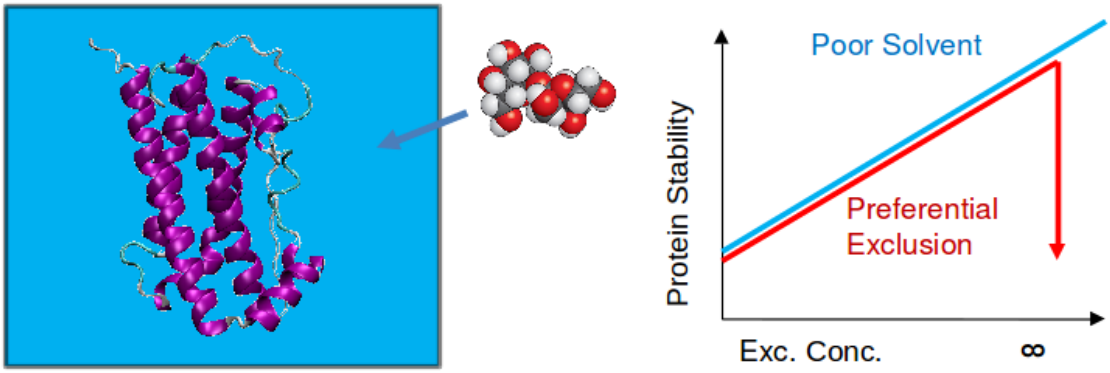

Figure 4.16: Schematic of an ideal experiment, where a protective osmolyte is added to a protein aqueous solution (left), and the resulting protein stability is monitored as function of the excipient concentration (right). The protein stability curves as predicted by the preferential exclusion theory (red line) or the poor solvent hypothesis (blue line) are shown.

As a result, the free energy of protein unfolding should increase, and the native fold should be stabilized. Let us also imagine that the selected osmolyte is soluble in water in all proportions, and that it is added to the protein solution until the percentage of water is negligible. In the end, we therefore have an osmolyte-only, homogeneous solution. At this point, the preferential exclusion theory would predict absence of stabilization, because it would be impossible to have preferential exclusion in a homogeneous solution. Therefore, the free energy of unfolding, which increased upon addition of the osmolyte, should drop again in the pure-osmolyte solution (red line in Figure 4.16).

This consideration, which suggests a discontinuity, or at least a non-monotonic behaviour, in the evolution of protein stability upon addition of excipients, seems unphysical and not realistic. This paradox could, however, be solved if the preferential exclusion mechanism was replaced with a "bad solvent theory", as already discussed in [117]. According to this bad solvent theory, a protein is stabilized in a poor solvent, which favors the collapsed, native fold against the expanded, unfolded states. In the previous example, for instance, the osmolyte would be a worse solvent than water, and its addition to the protein solution would prevent the unfolding transition. This bad solvent theory would predict that the protein stability monotonically increases when moving from pure water, i.e., the good solvent, to pure osmolyte, i.e., the bad solvent, without discontinuities (blue line in Figure

4.16). In this context, preferential exclusion would be a natural consequence of the unfavorable protein-osmolyte interactions. Moreover, this bad solvent theory has the advantage of unifying the description of protein folding with the stability of polymers in solutions. It is well known that, in a good solvent, the random coil of a polymer adopts an unfolded conformation, while in a poor solvent, where the polymer-solvent interactions are not favored, a more tight and contracted conformation is sampled.

However, proteins represent a very peculiar and heterogeneous class of polymers, formed by the assembly of the 20 amino acids. A huge number of possible sequences can result from the combination of these 20 residues, leading to a very large heterogeneity. For instance, the surface of a protein often shows patches which have completely different characteristics, in terms of hydrophilicity/hydrophobicity or surface charge. As a result of this heterogeneity, it is not easy to define a "bad solvent" for a protein polymer as a whole. In contrast, as shown in this Chapter, it is easier to observe preferential exclusion from specific regions of the protein. That is, a good solvent for specific patches on the protein surface could be, at the same time, a poor solvent for other side-chain sequences of the same protein. In this framework, for instance, glucose proved to be a good solvent for residues 33-44 of protein L, but a poor solvent for the hydrophobic region formed by amino acids 2-30.

The observation that certain amino acid sequences are more important than others in determining the protein behavior seems to suggest a sequence-dependent mechanism of protein folding and aggregation. However, it is also true that the osmolytes affect different proteins in similar ways. For instance, urea is a destabilizing osmolyte for any protein, while sugars, many polyols, amino acids and trimethylamine N-oxide (TMAO) always show a protective effect. The osmolyte effect is, therefore, universal, and operates on proteins in general. As already suggested [269], this universality implies a backbone-based mechanism of protein folding.

We could therefore imagine the folding transition as a multistep process. At the beginning, the unfolded state of the protein exposes its backbone to the solvent. If the solvent is a bad solvent for the backbone, the unfavorable backbone-solvent interactions start the refolding process. The folding transition is then guided by the side chains, which drive the protein conformation toward the specific and unique native fold [269]. According to this theory, exclusion of the osmolytes from the protein backbone is crucial for starting the folding transition. A protective osmolyte could even favorably interact with specific side chains, provided that it is a bad solvent for the backbone. For instance, it was shown in Chapter 3 that the sugars can reduce the risk of unfolding, even though they are attracted by the polar side chains on the surface of hGH. However, also the preferential exclusion from specific side chains sequences on the protein surface may be important to prevent undesired conformational changes. Exclusion from the APRs, for instance, can minimize the risk of aggregation.

List of Symbols

b	-	eigenvector component (Equation 4.8)
$\mathbf{b}$	-	eigenvector (Equation 4.8)
c	-	reweighting factor
$\mathbf{C}$	-	time-lagged matrix of the dynamical correlation functions (Equation 4.8)
$C_{j,k}$	-	dynamical correlation functions
d	m	distance
dRMSD	m	distance root mean square deviation
F	J mol^{-1}	free energy
$g(r_{dist})$	-	switching function
k_B	J K^{-1} mol^{-1}	Boltzmann constant
n_{all}	-	coordination number of water and glucose molecules
n_g	-	coordination number of glucose molecules
N	-	number of atoms in the protein
N_f	-	number of frames used to describe the path collective variable (Equation 4.6)
N_s	-	current step in the metadynamics run
O	-	candidate order parameter for the VAC-MetaD method
$\mathbf{O}$	-	set of candidate order parameters for the VAC-MetaD method
P_X	-	percentage of a conformation X in a free energy surface
Q	-	contact map CV
r	m	radial distance
$\bar{r}$	m	arithmetic mean of the distance
r_0	m	cutoff distance
R	m	atomic coordinates, scalar
R^0	m	reference α-helix or antiparallel β-sheet configuration
R_g	m	radius of gyration
R_{np}	-	non-polar area ratio (Equation 4.1)
r_{dist}	m	distance RMSD with respect to a reference α-helix or antiparallel β-sheet configuration
s	-	collective variable
$\mathbf{s}$	-	set of collective variables
s_{path}	-	path collective variable
S	m^2	total surface area

S_g	-	tetrahedral order parameter
S_k	-	tetrahedral order parameter
S_{np}	m^2	non-polar surface area
t	s	time
t_f	s	final time for the computation of the lag time τ
t_0	s	initial time for the computation of the lag time τ
t^*	s	"real" time in an unbiased simulation
Δt	s	time difference associated with the the lag time τ
Δt_{MD}	s	time step in the metadynamics run
T	K	temperature
V	J mol^{-1}	potential
x	-	atom in a protein configuration
X	-	protein configuration

Greek Letters

α	-	acceleration factor (Equation 4.11)
α_h	-	α-helix content
β	-	preferential exclusion parameter (Equation 2.1)
β_s	-	antiparallel β-sheet content
Γ	-	α-helix or antiparallel β-sheet content (Equation 4.3)
δ	-	sum of non-polar area ratios (Equation 4.2)
θ	-	threshold value for the deposition frequency in Frequency Adaptive Metadynamics
λ	-	eigenvalue (Equation 4.8)
λ_{path}	-	parameter for calculation of the path collective variable (Equation 4.7)
Π	-	integral preferential exclusion parameter (Equation 2.2)
σ	-	Gaussian width
τ	s	lag time (Equation 4.8)
τ_c	s	infrequent deposition frequency in Frequency Adaptive Metadynamics
τ_{dep}	s	deposition frequency in Frequency Adaptive Metadynamics
τ'_{dep}	s	corrected deposition frequency in Frequency Adaptive Metadynamics (Equation 4.12)
τ_0	s	initial deposition frequency in Frequency Adaptive Metadynamics
ψ	-	tetrahedral angle
Ω_μ	-	set of six residues of the protein

Chapter 5

Vacuum Induced Surface Freezing: Does Nucleation Control Affect Protein Stability?

It is clear from the previous Chapters that the possibility to control the morphology of a frozen product, and hence the extension of the ice interface, would be extremely beneficial. In principle, adsorption and denaturation phenomena may be minimized by promoting the formation of a small ice-water surface area. This could be achieved reducing the cooling rate, and inducing nucleation at a high temperature. However, the nucleation temperature is generally a stochastic variable, and its value is randomly distributed. This results in huge heterogeneity, and limited possibility to control product morphology.

This inhomogeneity contrasts with the stringent requirements of the pharmaceutical industry in terms of process control and product quality. To overcome this problem, various techniques have been developed over the years [58, 62–64]. In this Chapter, the effects of a technique for the control of nucleation, namely vacuum induced surface freezing (VISF), on protein stability will be investigated [1].

5.1 Control of the Freezing Process: A Literature Overview

A common practice to control crystal size in a freeze dried sample is annealing, which consists in holding the frozen product at a temperature above the glass transition value for a given amount of time. As a result, large ice crystals grow at the expense of smaller ones, improving homogeneity [76]. However, annealing

[1] Part of this chapter is based on already published papers [64, 270, 271]

introduces an additional step in the freeze-drying process, and the reduction in drying time due to increased pore size could be offset by the extra time required for annealing.

Another technique which could be used is the addition of impurities to the solution undergoing the freezing process, which may serve as nucleating agents [59, 272–275]. These agents, which often consist of silver iodide and bacteria such as *Pseudomonas syringae*, promote the formation of nuclei at high temperatures. However, they cannot be used to regulate the nucleation temperature and would not be accepted by regulatory agencies. Thus, a true controlled nucleation technology would provide superior benefits, and different techniques have been proposed over the years to address this problem.

The first ones were electrofreezing [276, 277], and ultrasound-induced ice nucleation (UIIN) [278–283]. Electrofreezing uses a high voltage pulse to initiate nucleation in supercooled water. However, the influence of electrofreezing on bubble formation, and the inhibitory effect observed in the case of a high saline concentration need further investigation. In addition, the need for individual electrodes in each sample and, most importantly, the presence of an electrode in direct contact with the product is not practical for GMP (Good Manufacturing Practice) product manufacture. In UIIN, ultrasounds are used to trigger ice nucleation in a chemically noninvasive way, but some concerns still exist, because localized high temperatures generated by cavitation, and the bubble themselves, may potentially damage or favor aggregation of very sensitive products.

Another possible solution may be the ice-fog technique [284–288]. In this case, small ice particles, generated by the release of cold nitrogen within the freezing chamber, penetrate into the vials, thus inducing ice nucleation. This ice fog could also be generated within an external condenser, or, alternatively, a gas flow cooled down within the condenser could be used to generate the ice particles inside the drying chamber. However, in the case of the ice-fog technique, it is not clear if it can be guaranteed that very large batches of vials, like those used in manufacturing, may be simultaneously reached by the ice particles, assuring a uniform batch behavior.

Finally, another freezing protocol which found application in lyophilization is the high-pressure-shift or depressurization method [289–291]. According to this technique, the drying chamber is first pressurized to 1.5-4.5 bar, and then rapidly depressurized. This initiates nucleation in the samples, previously equilibrated at a temperature below the equilibrium freezing point. The application of this approach, however, requires that the freeze-dryer can manage the 0.5-3.5 bar overpressure which is required to trigger the freezing process.

A method which could easily be scaled-up, and which will be used in this work is vacuum induced surface freezing (VISF), also known as vacuum induced nucleation (VIN), originally proposed by Kramer *et al.* [292], and later used by Liu *et al.* [272]. Substantial improvements in the VISF method were achieved by Oddone *et al.* [82–84], who proposed two new strategies, namely, VISF-1 and VISF-2. The

VISF method starts by stabilizing the product at a given temperature (T_n), followed by lowering the pressure inside the chamber to a formulation-specific value. This reduction in pressure promotes endothermic evaporation of water, thus inducing ice nucleation. After nucleation has started in all vials, atmospheric pressure is re-established and the fluid temperature decreased to a value (T_m) below the onset of ice melting to promote the formation of large ice crystals. Finally, the product is cooled down to, for example, -40/-50 °C to complete solidification of the solution. The typical evolution of temperature and pressure during VISF is shown in Fig. 5.1.

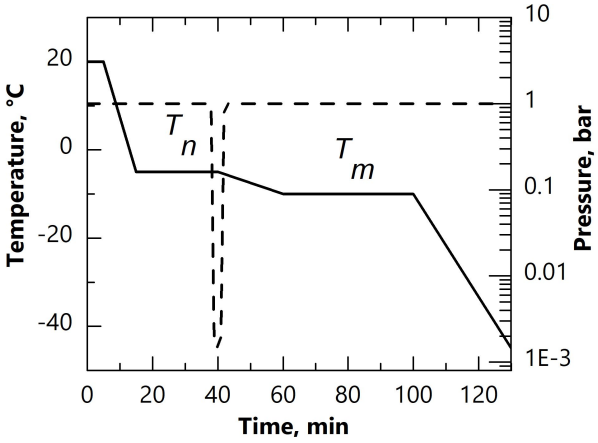

Figure 5.1: Scheme of the evolution of temperature and pressure during VISF

With respect to the original approach by Kramer *et al.* [292], in the VISF-1 method the vacuum time is shorter, 1 min vs 5 min, and the temperature of the fluid is maintained at the set point value during this period, and decreased only when atmospheric pressure is re-established. In the VISF-2 method, once the desired value of vacuum is reached, the drying chamber is isolated from the condenser. These modifications were found to substantially improve the elegance of the final product.

5.2 Aim of the Study

Oddone *et al.* [82–84] showed that vacuum-induced nucleation can be highly beneficial for uniformity of the batch and optimization of the freeze-drying cycle. Here, further insight will be provided into the effects of controlled nucleation on protein stability, using the VISF-2 approach. Human growth hormone (hGH) and factor VIII (FVIII) will be used as model proteins for this analysis. The extent of hGH aggregation after freeze-drying, carried out using either spontaneous or vacuum-induced nucleation, and the bioactivity of both the lyophilized proteins have been

chosen as benchmarks to investigate the impact of the selected freezing protocol on the protein stability. Furthermore, the effects of controlled nucleation on the reconstitution time and the applicability of vacuum induced surface freezing for different types of containers will be investigated.

5.3 Experimental Approach

5.3.1 Freeze-Drying Cycles

Different freeze drying cycles were performed, using either spontaneous or controlled nucleation, and for different model proteins. Details of these cycles are listed in Table 5.1

Table 5.1: Details of freeze drying cycles performed to assess the effect of controlled nucleation on protein stability. Temperature T and pressure P conditions used during primary (PD) and secondary (SD) drying are also listed.

#	Formulation	Nucleation Type	T_{PD} °C/P_{PD} Pa	T_{SD} °C/P_{SD} Pa
1	hGH formulation	Spon.	-35/7	20/7
2	hGH formulation	VISF	-35/7	20/7
3	Human plasma	Spon	-12/10	25/10
4	Human plasma	VISF	-12/10	25/10

For cycles 1-2, the stability of human growth hormone (hGH) was evaluated in presence of either 6 mg/ml trehalose (T) or 6 mg/ml cellobiose (C). In both cases, mannitol (M) at 42 mg/ml concentration was used as bulking agent. Mannitol is among the most commonly used bulking agents, and trehalose is often selected as lyoprotectant in protein formulations. Cellobiose is not equally common in pharmaceutical formulations, and was used in this work with the aim to evaluate its potential as a protein stabilizer. The high mannitol content should ensure quantitative crystallization during freezing, whereas the protein remains immersed in the amorphous sugar phase. Here, the objective was to compare two freezing protocols (spontaneous and controlled nucleation), and two different stabilizers (cellobiose and trehalose), focusing on protein stability. The effect of the freezing protocol should mainly be related to the ice crystal size, and therefore specific surface area, of the cake. The ice crystal size is affected by several factors, encompassing both the freezing protocol (cooling rate, nucleation temperature, presence/absence of an annealing step), and the formulation. The addition of a bulking agent, such as mannitol, as dominant component (42 mg/ml vs 6 mg/ml) makes it possible to rule out the effect of having a different stabilizer (cellobiose vs trehalose). Therefore, the ice crystal size should be influenced solely by the freezing protocol. This allows

a clearer separation of the contributions of formulation and freezing protocol on protein stability.

All the formulations were prepared in 1.1 mM sodium phosphate buffer, pH 7.8. The excipient formulations were filtered prior to filling using 0.2 μm filters, and the protein was always added to a pre-filtered formulation. The hGH used was a commercial time-expired recombinant product supplied by NIBSC, and was dialysed for about 24 h against the selected formulations using a dialysis kit (Thermo Scientific Slide-A-Lyzer, 3.5 kDa cutoff). Dialysis was carried out under controlled temperature conditions (4 °C). At the end of the dialysis, the concentration of hGH was determined by UV spectrophotometry, and the dialysate was diluted so as to obtain a concentration of hGH of 0.25 mg/ml. As a first step, the protein formulation used for these two cycles was quench frozen to -70 °C. After thawing, an aliquot was freeze-dried, while the remaining part was directly analyzed by size exclusion chromatography, as will be described in the following.

Each vial (5 ml screw capped vial from Schott, internal diameter 14 mm) was filled with 1 ml of solution. The freeze-drying cycles were carried out using a Virtis Genesis 25EL (Biopharma Process Systems, Winchester, UK) freeze dryer. Each batch contained 100 vials, of which 38 contained hGH (21 in the T+M formulation and 17 in the C+M one), and 62 were protein-free formulations (31 T+M and 31 C+M). In any case, care was taken to have the active (hGH-containing) vials in the middle of the batch, and therefore shielded by other vials, so as to avoid edge effects. This should guarantee a more homogeneous behavior of active vials in different batches and during different cycles because of a similar exposure to external radiations. As regards spontaneous freezing, a 0.5 °C/min ramp to -45 °C was performed. The product was then kept at -45 °C for 2 hours. By contrast, for VISF the batch was equilibrated at -5 °C for 1 h. Pressure inside the chamber was then lowered until nucleation was observed (1 mbar). When all the vials nucleated ($<$ 1 min), pressure was quickly released to the atmospheric value. The product was then kept at -10 °C for 1 h. Finally, temperature was decreased to -45 °C and held for 2 hours. Primary drying was then carried out at -35 °C and 7 Pa (0.07 mbar), while during secondary drying temperature and pressure were 20 °C and 7 Pa, respectively. A 10 h ramp from the primary to secondary drying temperature was used, in order to reduce the risk of collapse.

This system is easy to be freeze dried, and is also representative of the low-density, partially crystalline cakes that are typical of many pharmaceutical products; it represents therefore a suitable model formulation. However, controlled nucleation may be extremely beneficial also for higher-concentrated systems, which are more difficult to process. The application of vacuum induced surface freezing to high-concentrated systems was investigated in cycles 3-4. In this case, the freeze dried product was screened human plasma (Blood Group: A+, National Blood & Transplant, Colindale, UK), with the addition of 40 mM Hepes (free acid H3375, Sigma Merck, Poole, UK). Batches of 1 ml aliquots were dispensed using a

Hamilton autodilutor (Hamilton M510B, supplied by Microlab Technologies Ltd, Westcliff-on-Sea, UK) into 5 ml ampoules (glass type I, Schott supplied by Adelphi Tubes, Haywards Health, UK) and fitted with 13 mm diameter halobutyl rubber lyo-closures (West Pharma, supplied by Adelphy) partially stoppered to allow sublimation to occur. A batch of 100 ampoules was prepared for both cycles. The ampoules were then loaded onto a freeze dryer (LyoBeta 15, Telstar Azbil SpA, Terrassa, Spain) and the freeze drying cycle begun. For the VISF protocol, nucleation was induced after equilibration at -5 °C for about 1 h, and the product was then equilibrated at -10 °C for about 45 min before the final ramp to -50 °C. As regards spontaneous freezing, a 0.5 °C/min ramp to -50 °C was performed. For both the spontaneous and the VISF cycle, the product was held at -50 °C for 3 h to complete freezing. The product was then held at -50 °C for 1 h at 20 Pa, and for 1 h at 10 Pa. The temperature was subsequently raised to -12 °C in 1 h and held at 10 Pa for 30 h. Secondary drying was eventually performed at 25 °C for 20 h. A 10 h ramp between primary and secondary drying was used.

After the cycle, the dryer was back-filled with dry nitrogen (from boil off of pure liquid nitrogen) and the closures stoppered down before ampoules were removed from the dryer. Ampoules were then flame-sealed using a manual ampoule sealer (Ampulmatic, Adelphi Tubes).

In each batch, the temperature profile inside 2-3 different vials/ampoules was monitored by means of T-type copper/constantan miniature thermocouples. Thermocouples were placed at the bottom center of the vial/ampoule, and touching the bottom. During drying, the pressure inside the drying chamber was monitored by means of a capacitance manometer (MKS Baratron, cycles 1-2), or using both a capacitance and a thermal conductivity (Pirani) manometers (cycles 3-4). The capacitance manometer always measures the correct value of pressure inside the chamber, while the Pirani gauge is calibrated in nitrogen and its readings are therefore shifted to higher values during the primary drying phase, when water vapor is present in the drying chamber as a result of sublimation. By comparison of the two pressure profiles, it is also possible to calculate the onset and offset times [293]. The onset corresponds to the time at which the Pirani gauge reading begins to decay/decline, and indicates that sublimation has ended in a not negligible number of vials. By contrast, at the offset time, the ratio between the capacitance and Pirani readings equals 1, meaning that sublimation has ended in all the vials. The difference between offset time and onset time (in the following referred to as onset-offset time) is an indicator of within-batch heterogeneity. The larger it is, the more the sublimation behavior of vials within the batch is different. The ramp to secondary drying was always started after the offset time, which can be considered as the end of the sublimation process. By contrast, the duration of secondary drying was set to 5 h for cycles 1 and 2, or 20 h for cycles 3 and 4.

During cycles 3-4, the pressure rise test (PRT) [294,295] was also used to monitor the end of primary drying. For doing this, the valve between the vacuum chamber

and the condenser was periodically closed for a given time interval, letting the pressure increase because of the accumulation of water vapor. This increase is fast at the beginning, but becomes slow towards the end of primary drying, providing useful information about the progress of the sublimation process.

For all cycles, the coefficient of variation (CV) of fill was assessed by measuring empty, filled and dried weights on three ampoules across the batch.

5.3.2 Characterization of the Formulations

Modulated differential scanning calorimetry (MDSC), dynamic mechanical analysis (DMA), and freeze-drying microscopy (FDM) analyses were performed on the 2 formulations under investigation in cycles 1 and 2.

MDSC was performed using a Q2000 calorimeter (TA Instruments, Elstree, UK). Large volume (100 μl) stainless steel pans were used and 80 μl aliquots were sealed inside them and analyzed against an empty pan as reference. Samples were frozen to -90 °C at 10 °C/min, and then heated to 25 °C at 1.5 °C/min, with modulation at 0.23 °C/min. Results were analyzed using Universal Analysis software (TA Instruments), and the critical thermal event temperature was assigned based on exotherms in the profiles (at peak maximum). These were detected in reversing, nonreversing, and total heat flow curves with a difference between them of 0.2 °C to 0.5 °C.

For the dynamic mechanical analyses, a DMA 800 (TA Instruments) was used [296]. A small volume (200 μl, approximately 200 mg) of the solution of interest was pipetted onto one surface of a steel sample tray which was then clamped and inserted into the DMA. Cooling of the system was provided by evaporation of liquid nitrogen. The storage modulus and $\tan\delta$ values were recorded and plotted against temperature (between -70 °C and +10 °C), at 1 Hz oscillation frequency. The dynamic amplitude and scan rates were 50 μm and 1 °C/min, respectively. The glass transition temperature was determined based on peaks in the $\tan\delta$ values.

For the FDM analyses, a small amount (5 μl) of each formulation was pipetted onto a special quartz crucible, using a metal shim and coverslip. The crucible was then mounted on a Linkam FDCS196 Cryostage and analyzed using an Olympus BX51 microscope with plane polarized light. Images were captured by a charge-coupled device camcorder. This process was programmed and controlled (Linkam Scientific Ltd., Tadworth, Surrey, UK). Samples were frozen to -50 °C at -5 °C/min. Subsequently, pressure was lowered to 10 Pa, and the sample was heated at 2 °C/min until collapse was observed.

The system was calibrated using 5 % w/v trehalose in water, and single determinations were made for each formulation.

5.3.3 Residual Moisture Analysis

For cycles 1-2, the residual moisture of dried samples was evaluated using automated Karl Fischer titration. Sample ampoules were opened in a low relative humidity environment using a glove bag (CaptairPyramid, Cole Parmer, London, UK) flushed with dry nitrogen to a RH level of 5% - 10%. Coulometric Karl Fischer titrations were performed using an automated Karl Fischer system, based upon the Aquafast system (A1-Envirosciences, Blyth, UK). Samples of lyophilized material were dispensed into septum-sealed autosampler vials in the glove bag environment and then reconstituted remotely with a known aliquot of anolyte. Following a period of time to allow insoluble material to settle, an aliquot of the supernatant was titrated into the coulometric Karl Fischer cell. The coulometer was calibrated using the Aquamicron P water standard solution (4 mg/g water content, A-1 Envirosciences) and titrations deemed valid only if a 50 μl injection gave a water result of 180 - 210 μg with a relative standard deviation less than 5% over three consecutive titrations. Three placebo-containing vials were analyzed for each formulation, and both the average value and relative standard deviation on the residual moisture could be calculated.

5.3.4 SEM Analysis

For all cycles, the pore dimension of the products obtained after freeze-drying was analysed using a Scanning Electron Microscope (JSM 7401F SEM, Jeol Ltd, Welwyn Garden City, UK), using secondary electron detection. The samples were cut along their vertical axis, and at least 3 SEM images were taken at the top, centre and bottom of the cake. Afterwards, approximately 100 pores were selected in each image, and each of them was approximated to an ellipse. The diameter of the circle having the same area to perimeter ratio of the approximating ellipse was then assumed as pore dimension, and the numerical average (with a corresponding standard deviation) of the obtained distribution was assumed as the average pore size, D_p, of the product.

5.3.5 X-Ray Diffractometry

For cycles 1-2, X-ray diffractometry was also performed on the placebo formulations, in order to identify the polymorphic state of mannitol in the dried samples. XRD patterns were obtained using a PANalytical X'Pert (Cu Kα radiation) diffractometer, within 5-65° as 2θ range. The data were acquired at each 0.026°. The resulting diffractograms were analysed using the HighScorePlus program (vers. 3.0.5), by comparison with the X-ray diffraction patterns of the reference materials (α, β and δ anhydrous mannitol, and the hemihydrate).

5.3.6 Evaluation of Protein Stability

In the case of cycles 1-2, the soluble aggregates of hGH were measured by size exclusion chromatography (SEC). HPLC-SEC analysis was performed on the initial frozen baseline, as well as on the dried samples. The freeze-dried samples were reconstituted to 1 ml using ultrapure water and then centrifuged for 5 min at 13000 rpm. A TSKgel G3000 SWXL column (Tosoh Bioscience, Reading, UK) was used on a Thermo Dionex U3000 HPLC system (ThermoFisher Scientific, Leicester UK) using a flow rate of 0.6 ml/min, a run time of 50 min and UV detection at 214 and 280 nm. The mobile phase used was 2-propanol R, 0.063 M phosphate buffer solution pH 7.0 R (3:97 V/V), as suggested in European Pharmacopoeia 9.0 [297]. 3 hGH-T+M vials and 3 hGH-C+M vials were analysed. A calibration curve for HPLC-SEC was also calculated, using a mixture of molecular weight markers ranging from 1350 to 670,000 Da (Bio-Rad's gel filtration standard, Bio-Rad, Watford, UK). Using this calibration curve, the retention time, as read from the chromatograms, was related to a corresponding molecular weight. Dimers and higher order soluble aggregates were observed and the relative peak area of these, as a percentage of the total peak area, was used to indicate protein preservation.

A bioassay was also used to quantify the bioactivity of the lyophilized hGH, by measuring cell proliferation in the lactogenic lymphoblastic rat cell line, Nb2-11 [298] (European Collection of Authenticated Cell Cultures, Public Health England, Salisbury, UK). Cells were routinely grown in suspension culture in Fischer's medium containing 10% (v/v) heat inactivated fetal bovine serum, 10% (v/v) horse serum, 0.075% (v/v) sodium bicarbonate, 50 μM 2-mercaptoethanol, 50 units/ml penicillin, and 0.05 mg/ml streptomycin. Fischer's medium, heat inactivated horse serum and 2-mercaptoethanol were obtained from Thermo Fisher Scientific. Fetal bovine serum was obtained from BioSera Europe (Nuaille, France) and sodium bicarbonate (7.5%), penicillin, and streptomycin from Sigma Chemical Co., Ltd. (Poole, UK). Cells were maintained by passaging every 4-5 days at a cell density of approximately 1×10^4 cells/ml and incubated at 37 °C in an atmosphere of 5% CO_2. At 16-20 h before assay, cells were prepared in culture medium without penicillin and streptomycin at a cell density of 2×10^5 cells/ml and incubated as aforementioned. The assay is based on the quantitation of ATP in metabolically active cells. The International Standard (WHO International Standard Somatropin, NIBSC code: 98/574) and the test samples were reconstituted, and diluted in assay medium which comprised Fischer's medium containing 1% (v/v) horse serum, 0.075% (v/v) sodium bicarbonate and 50 μM 2-mercaptoethanol. Nb2-11 cells, pretreated as described previously, were washed twice with phosphate buffered saline (pH 7.2), suspended in assay medium at a density of 1 x 10^5 cells/ml and transferred (50 μl/well) to a 96-well microtitre plate (Greiner Bio-One (Stonehouse, UK) code 655180). Dilutions of standard and test samples were added as an additional 50 μl volume to each well. The maximal cell response was assessed by stimulating

with a 10 ng/ml concentration of the International Standard, 98/574. The cells were incubated for 30h ± 2h at 37 °C in an atmosphere of 5% CO_2. Proliferation was then estimated by the addition of a 100 μl volume of CellTitre-Glo reagent to each well and incubating with shaking in the dark at room temperature. CellTiter-Glo® Luminescent Cell Viability Assay reagent was obtained from Promega UK (Southampton, UK). The resulting cell extracts were then transferred to a white 96-well plate (Greiner Bio-One code 655075), incubated at room temperature in the dark for 15 minutes and the luminescence measured using a SpectraMax M5 microplate reader (Molecular Devices (UK) Ltd., Wokingham, UK) controlled by software SoftMax Pro 6 (Version 6.4, 1992-2014 Molecular Devices LLC). Each test dilution was measured in triplicate and test samples were assessed in two independent bioassays. Relative bioactivity was assessed by fitting a sigmoidal curve using CombiStats, version 5.0 (1999-2013 EDQM/Council of Europe).

For cycles 3-4, the freeze dried ampoules were put down to degradation study at 45 °C, 37 °C , 20 °C, 4 °C and -20 °C (5 ampoules at each temperature, for both freezing protocols). A protein that is present in plasma, and which is particularly sensitive to denaturation stresses, is factor VIII (see Figure 5.2). FVIII is an essential clotting factor whose impaired function results in haemophilia A, a rare, sex-linked bleeding disorder that manifests with prolonged bleeding time, bruising, and joint bleeds. Assessment of the stability of FVIII was carried out through accelerated degradation studies which allow the prediction of degradation rates for ampoules stored at low temperatures (e.g. -20 °C) based on the observed loss in potency of samples stored at elevated temperatures (e.g. +4, +20, +37, +45 °C) [299]. This is an indirect method used to determine rate of loss based on the relationship between reaction rates and temperature given by an Arrhenius equation and where a first order reaction rate is assumed [300, 301]. The potency of factor VIII, in ampoules prepared by the two different freezing protocols, was estimated by a chromogenic method [302]. In this method, optimal amounts of calcium and phospholipids, and excess amounts of factors IXa and X, are added to the freeze dried sample. In these conditions, factor X is converted to factor Xa by factor IXa, and the rate of this activation is dependent on the amount of factor VIII. The factor Xa generated hydrolyses the chromogenic substrate and the amount of colour produced is read photometrically at 405 nm. The intensity of colour is therefore proportional to the amount of factor VIII in the test sample.
The Coatest SP4 FVIII kit (Chromogenix, Werfen Ltd., Birchwood, UK) was used for the analyses. At different time points after freeze drying, a reference standard (WHO 6^{th} IS FVIII/VWF Plasma 07/316, 0.68 IU/ml) and test samples were reconstituted with 1 ml sterile water, and allowed to stand until dissolved. The reconstitution time obtained for the two different freezing protocols was also measured by visual inspection. The reconstituted samples were then analyzed on the ACL TOP 550 analyzer (Werfen Ltd., Birchwood, UK). Results were analyzed using CombiStats, version 5.0 (1999-2013 EDQM/Council of Europe).

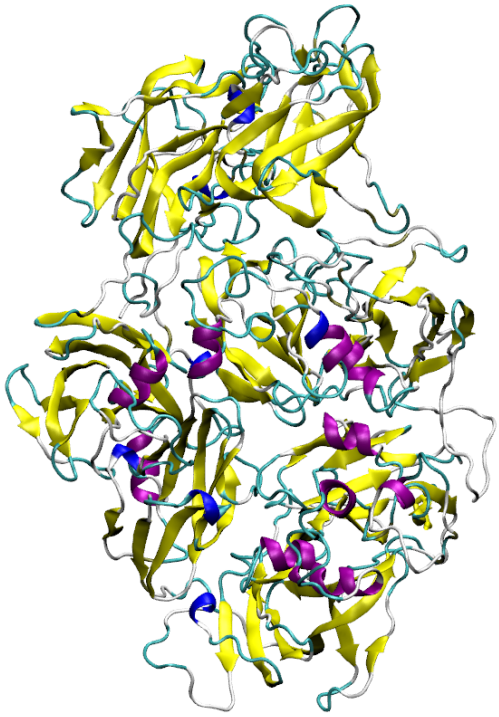

Figure 5.2: Snapshot of native FVIII. α-helices are in purple, β-sheets in yellow, coils in white, turns in cyan, and 3_{10}-helices in blue.

5.4 The Effect of Vacuum Induced Surface Freezing on Protein Formulations

5.4.1 Thermal Characterization of the Formulations

In Table 5.2, the glass-transition temperature T_g' (mid-point) as measured by MDSC and DMA, and the collapse temperature T_c as measured by FDM are shown, for the formulations used in cycles 1 and 2. The formulation with the lowest collapse temperature was the cellobiose-based one, which showed a first microcollapse at about -35 °C, and complete collapse around -29 °C. The collapse temperature observed for the trehalose-based formulation was about 3 °C higher (-26.5 °C). Results obtained by freeze-drying microscopy are correlated with the glass transition temperature, as measured by both MDSC and DMA. These 2 techniques showed good accordance and revealed a higher critical temperature in the case of trehalose as excipient.

Table 5.2: Results of MDSC, DMA, and FDM performed on the 2 formulations under investigation in cycles 1 and 2.

	T'_g, °C (mid-point)		T_c, °C
	MDSC	DMA	FDM
T+M	-27	-26	-26.5 (collapse)
C+M	-28	-28	-35 (microcollapse), -29 (collapse)

5.4.2 Effect of the Freezing Protocol on Product Morphology

The higher nucleation temperature in the case of the VISF technique resulted in the formation of larger pores, as shown in Figure 5.3 and Table 5.3 for cycles 1-4.

Table 5.3: Pore dimension D_p in the dried product, for freeze-drying cycles 1-4.

Cycle		D_p, μm		
#	nucleation	bottom	center	top
1 T+M	spon.	27 ± 14	42 ± 23	33 ± 14
1 C+M		21 ± 9	28 ± 12	25 ± 14
2 T+M	VISF	55 ± 23	83 ± 33	57 ± 25
2 C+M		53 ± 22	81 ± 30	74 ± 33
3 plasma	spon.	56 ± 17	66 ± 18	56 ± 14
4 plasma	VISF	114 ± 48	127 ± 58	112 ± 62

In all cases, larger pores were observed at the center of the dried cakes, while the smallest pores were either at the bottom or at the top. This can be explained considering that the larger temperature gradients at the bottom of the product, where the vial is in contact with the dryer shelf, and the presence of cryoconcentration effects at the top, promote the formation of smaller ice crystals and therefore smaller pores. It is also evident that the average pore size was significantly larger in the case of VISF, for all formulations considered. Specifically, the lower nucleation temperature in the spontaneously nucleated samples approximately halved the pore size compared to the case of the controlled freezing approach.

Furthermore, VISF also had a not negligible effect on the polymorphs formation in mannitol-containing samples (cycles 1 and 2). Figure 5.4 shows the XRD spectra for the lyophilized samples produced in cycles 1 and 2.
The various mannitol polymorphs (α, β, δ) have different stability, with the β form being thermodynamically stable at room temperature, and the other two being metastable [303]. In the case of the C+M formulation, both the δ and the β polymorphs were present after both uncontrolled and controlled nucleation. Only traces of α and hemihydrate were present after spontaneous nucleation, while the

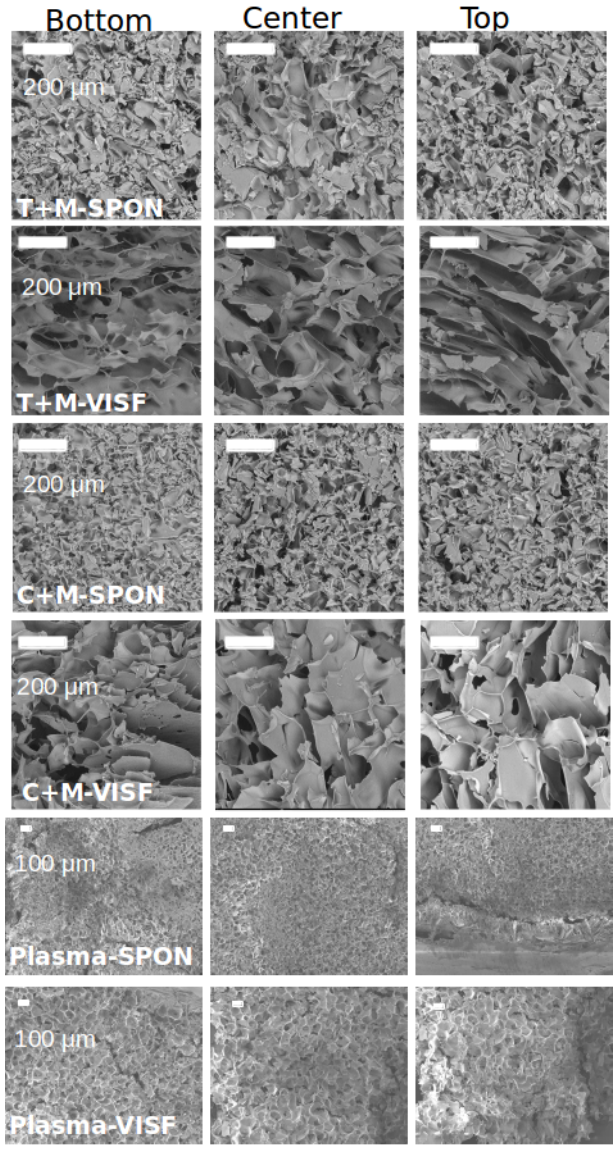

Figure 5.3: SEM images of the samples obtained at the end of cycles 1 (T+M and C+M spon.), 2 (T+M and C+M VISF), 3 (plasma spon.) and 4 (plasma VISF). T+M: trehalose + mannitol, C+M: cellobiose + mannitol

VISF resulted in a slightly higher content of these forms. Also in the case of the trehalose-based formulation, the β and δ forms were present, while the α polymorph was present in traces, and no clear peak for the hemihydrate was detected after spontaneous nucleation. However, the VISF again seemed to slightly increase the presence of α mannitol, and to induce formation of the hemihydrate. This XRD analysis suggests that the controlled nucleation approach may have an impact on

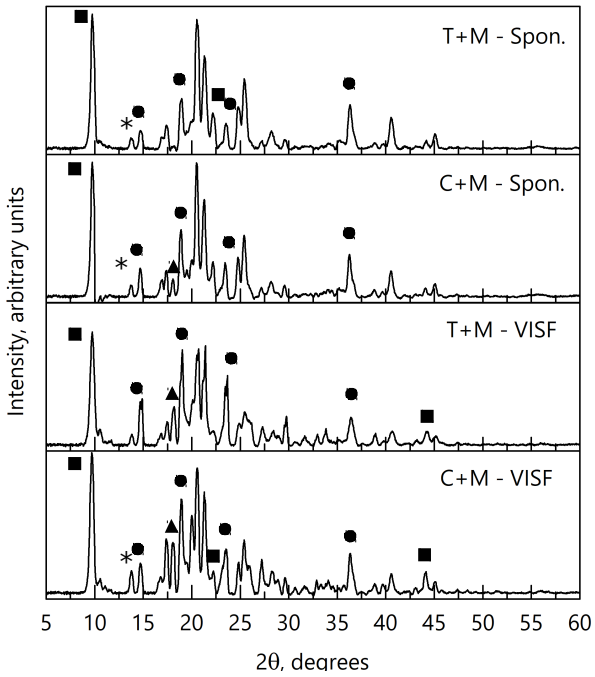

Figure 5.4: X-ray diffraction spectra for the samples produced in cycles 1 (spontaneous nucleation) and 2 (VISF). The labels identify the most representative peaks for the anhydrous α (*), β (●) and δ (■) polymorphs and for the (▲) hemihydrate.

the formation of mannitol polymorphs, in line with previous observations [83,304].

The XRD analysis also suggests that the polymorphs composition of mannitol-based formulations may be influenced by the addition of other excipients, such as trehalose and cellobiose, and by the freezing protocol. This aspect should be taken into account when designing a freeze-drying process for a biopharmaceutical, assuming mannitol as a formulant. Many factors contribute to the specific surface area of a lyophilized cake where mannitol is used as bulking agent. Some of them, such as the ice crystal size and polymorphs composition, are strongly related to the freezing protocol. A second class of factors, such as the surface roughness and crystal habit of the mannitol polymorphs, does not depend on the freezing procedure employed. For instance, Cares-Pacheco *et al.* [305] induced the recrystallization of pure β and δ polymorphs from an aqueous solution of D-mannitol using acetone as an antisolvent. For the α mannitol generation, a crystallization procedure by seeding and fast cooling was used. Afterwards, they measured the specific surface area of D-mannitol polymorphs using the IGC-surface energy analyser (IGC-SEA) from SMS, and found that α, β and δ mannitol have specific surface area of 8.54 m^2g^{-1}, 0.37 m^2g^{-1} and 1.01 m^2g^{-1}, respectively, indicating that the ice-freeze concentrate surface area, and potentially the protein stability, may be strongly influenced by

the polymorphs composition.

5.4.3 Influence of the Freezing Protocol on Drying Performance

The positive effects of VISF on primary drying time, already discussed in the literature [82, 83], were confirmed in this work for cycles 3 and 4 in Table 5.1. As shown in Figure 5.5a, primary drying lasted about 17.6 h for the cycle with spontaneous nucleation (difference between offset time and the beginning of the drying process). Moreover, the onset-offset time in this case amounted to about 4.2 h. In contrast, primary drying was shorter, about 13.1 h, when the VISF technique was applied (cycle 4, Figure 5.5b). This result is not negligible, as it corresponds to approximately 25.6 % reduction in sublimation time upon application of controlled nucleation. The onset-offset time, which is a measure of variability in sublimation behaviour, also decreased to about 3 h when controlled nucleation was used. This suggests that the application of VISF is beneficial when homogeneity is an issue, and this effect may be even more significant in the case of large industrial-scale batches. Cycles 1 and 2 were performed in a different freeze dryer (Virtis Genesis 25 EL) where it was not possible to accurately determine the end of primary drying because of the absence of a thermal conductivity manometer (Pirani). However, the pore size reported in Table 5.3 and other data on the same formulation [271] suggest that a similar reduction in drying time could be obtained upon application of the VISF protocol.

The observed difference in sublimation rate between spontaneous and controlled nucleation is related to the difference in pore size previously discussed. The VISF technique made it possible to induce nucleation in all samples at a high temperature (-5 °C), where formation of ice nuclei is still not observed in spontaneously-frozen ampoules. In turn, this high nucleation temperature translates into the formation of large ice crystals, which subsequently convert into equally large pores when ice is removed during sublimation. The removal of water vapour through these pores occurs with a significantly reduced resistance to mass transfer, boosting the sublimation process.

Another consequence of this difference in pore size is the temperature profile measured within the product (red curves in Figure 5.5). The information about product temperature given by thermocouples can be considered accurate only for the first part of primary drying, before a significant change in slope is observed. Considering the significant region of the temperature profile recorded, the maximum temperature reached by the plasma product during the spontaneous run was about 5 °C higher (-23 °C against -28 °C) than for the VISF cycle. This can be explained considering that the heat flux provided to the ampoules by the shelf is the same for both cycles, but the endothermic sublimation process is faster when controlled nucleation is used. Therefore, application of the VISF protocol also reduces the

risk of product collapse.

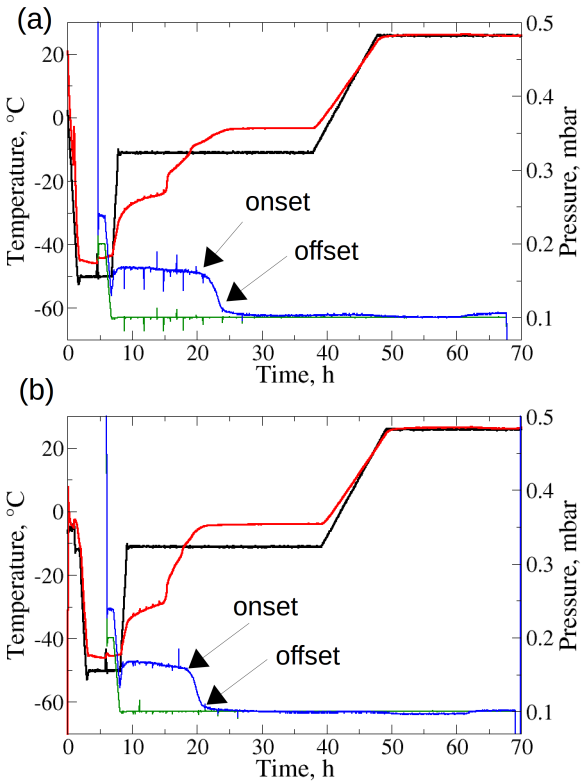

Figure 5.5: Temperature and pressure profiles during a) the spontaneous run (cycle 3) and b) the VISF run (cycle 4) for the plasma batch. Black line: fluid temperature, Red line: Product temperature, Green line: Baratron, Blue line: Pirani.

5.4.4 Effect of Controlled Nucleation on Protein Stability

HPLC-SEC Results for the Human Growth Hormone

The formation of hGH soluble aggregates in the frozen baseline (freezing to -70 °C) or after freeze-drying (cycles 1-2) was characterized using HPLC-SEC analysis. As can be observed in Figure 5.6, HPLC-SEC analysis of available vials suggests that the percentage of high molecular weight species is minimally affected by the use of controlled nucleation on the hGH formulation being investigated. It seems that most of the aggregates were formed after the quench freezing process (red bars), especially in the case of the trehalose-based formulation. This may suggest that the hGH should be extremely sensitive to freezing stresses, and in particular to the formation of a large ice-freeze concentrate interface, in line with the molecular

dynamics simulations presented in Chapter 4 and the experimental observations by Eckhardt *et al.* [101]. From the HPLC-SEC results here reported, it seems that no significant difference exists between trehalose and cellobiose as excipients. However, cellobiose is a reducing sugar unlike trehalose, therefore unlikely to be as good in long-term stability at room temperature or above due to Maillard reactions.

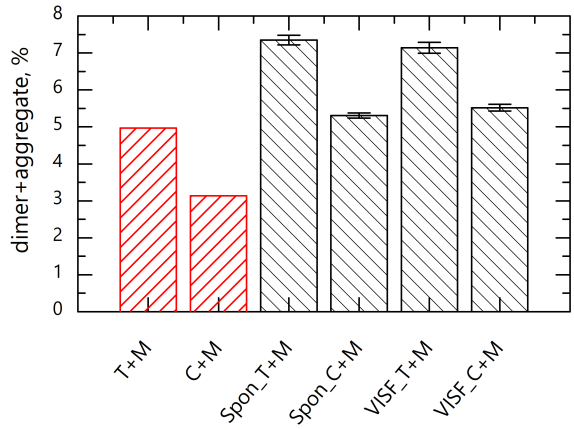

Figure 5.6: Percentage of dimers and higher order aggregates evidenced by HPLC-SEC in the frozen baseline (red bars) and after freeze-drying (black bars). The values for the two freezing protocols being used (spontaneous and controlled nucleation), and for the 2 formulations under investigation (T + M and C + M), are shown.

Cell-Based Potency Assay for hGH

The bioactivity of freeze-dried hGH was also evaluated, for both cycles 1 and 2 in Table 5.1. Nb2-11 cells were stimulated with concentrations of hGH ranging from 0.0078 to 1 ng/ml. Maximal fold change, calculated as the ratio of the mean response of the 10 ng/ml positive control ($n = 8$, per plate) to the mean response of the 0 ng/ml control ($n = 8$, per plate), was consistently greater than 2.8. Representative dose response curves for each sample are shown in Figure 5.7.

All samples were shown to stimulate proliferation of Nb2-11 cells and a sigmoidal (log dose) response was observed. The bioactivity of the vacuum-induced samples was comparable (within 6%) of the bioactivity measured in the spontaneously nucleated samples, and only the trehalose formulation showed a slightly decreased potency compared to the standard. Overall, this confirms that VISF does not affect hGH bioactivity as measured in Nb2-11 cells.

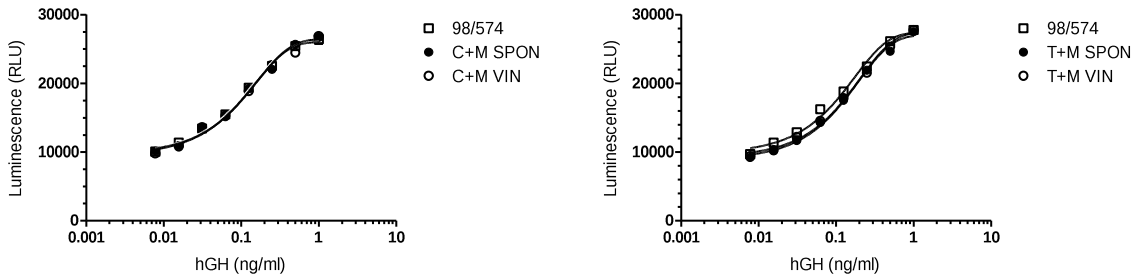

Figure 5.7: Representative dose-response curves of the stimulation of Nb2-11 cells with hGH. The response to the International Standards is shown as open squares (□), spontaneously nucleated samples are shown by closed circles (●), and vacuum-induced nucleated samples by open circles (○).

Comparison with Previous Works

Hence, according to the cell-based potency assays performed, VISF is not detrimental for the biological activity of hGH. This indicates that the reduction in chamber pressure used to induce nucleation within the vials should not have a negative impact on protein conformational stability. The HPLC analysis also seemed to evidence no significant difference in the aggregation behavior of hGH, no matter which of the two freezing protocols was used. The findings herein discussed compare fairly well with the report by Fang *et al.* [66], where a controlled nucleation approach was found to improve LDH stability after freeze-thawing. However, Fang *et al.* also observed that the controlled nucleation technique was not equally beneficial for LDH stability during drying. This observation, which remained unanswered in [66], may be related to the different level of residual moisture in samples produced by spontaneous or controlled nucleation. Oddone *et al.* [83,84] observed that the moisture content during drying was higher in the case of controlled nucleation, because of a lower desorption rate. This is also suggested by the slightly higher residual moisture content detected in this work (Table 5.4) in samples obtained by VISF.

It is also interesting that, in the case of both spontaneous and controlled nucleation, the trehalose-based formulation showed a slightly higher residual moisture content after freeze-drying than the cellobiose-based one, and the difference would have probably been more pronounced if measured immediately after primary drying [83,84].

The water content is a crucial parameter for the post-drying stability of proteins, even though a more dried cake is not always more stable. For instance,

Table 5.4: Results of coulometric Karl Fischer analysis performed on dried samples, in the case of cycles 1-2.

	Karl Fischer titration	
	Water content w/w ($n = 3$)	CV %
T+M spon.	0.49	12.2
C+M spon.	0.37	12.6
T+M VISF	0.76	25.7
C+M VISF	0.53	11

lyophilized BSA and bovine γ-globulin formulations were more stable at 10% water content than at <1% [81]. However, a higher moisture content will result in higher mobility, according to a Williams-Landel-Ferry relation [306]. More specifically, the protein mobility responsible for both unfolding and aggregation in the dried state is significantly enhanced in the presence of water. The absence of beneficial effects on LDH post-drying stability observed by Fang *et al.* [66] may therefore be at least partially explained by the higher water content during drying in the case of controlled nucleation.

If this hypothesis were true, the controlled nucleation approach may be beneficial for protein stability in the dried state, provided that a strict control of the moisture level is carried out. This is certainly possible because, on constant processing conditions, the maximum temperature reached during drying in samples produced by VISF is lower than in the case of conventional freezing. This means that if process conditions were adjusted so as to have equal product temperature, the rate of water removal could be enhanced for the VISF nucleated samples, thus minimizing the detrimental effects related to a high moisture content. Alternatively, a higher temperature could be used during secondary drying, or this step could be prolonged, in the case of controlled ice nucleation to guarantee post-drying stability. However, further investigation would be required to clarify whether this hypothesis is true.

As discussed in the previous Chapters, the thermodynamic mechanism of protein stabilization by preferential exclusion [32] is generally dominant during the first steps of the freezing process, because of the high availability of liquid water before cryoconcentration and drying. However, during primary and secondary drying, protein preservation should be driven by a kinetic mechanism, where the protein motions responsible for unfolding and aggregation are hindered by the formation of a viscous, glassy matrix [31]. A thermodynamic mechanism of protein stabilization in the dried state was also proposed, where the excipient should replace water molecules in hydrogen bonding with the protein, thus inhibiting conformational changes [40]. Preliminary results seem to suggest that there may be a difference in

the behavior of cellobiose and trehalose during freezing or drying (Figure 5.6). For instance, the C + M formulation showed only 3% aggregate content after freeze-thawing, whereas in the trehalose-based one, an aggregate concentration as high as 5% was detected. By contrast, the noticeable efficiency of trehalose in the dried state is not surprising because it was shown in the previous Chapters that trehalose can form highly structured matrices, characterized by a dense, compact hydrogen-bonding network, which kinetically prevent any protein motion. While the data herein presented are not enough to draw conclusions on this point, the theory that not all the excipients are equally effective during freezing and drying has been also supported by the *in silico* modeling presented in this work.

FVIII Potency

As regards cycles 3 and 4, the potency of FVIII was measured post-drying (n=2), and after 1 (n=1), 3 (n=1), 6 (n=2) or 9 (n=2) months storage at 45 °C, 37 °C , 20 °C, 4 °C or -20 °C (see Figure 5.8). As expected, the measured potency of FVIII decreased during storage especially at the highest temperature (37 °C and 45 °C) considered in this work. Moreover, Figure 5.8 suggests that the VISF technique generally resulted in a slightly improved FVIII stability, both at low temperature (-20 °C for the black curves in Figure 5.8a, 4 °C or 20 °C for the orange and blue series in Figure 5.8b), and during high temperature storage (37 °C or 45 °C for the green and red curves in Figure 5.8a).

Furthermore, the residual FVIII potencies for ampoules stored at +4 °C, +20 °C, +37 °C and +45 °C, for the two different freezing protocols, were expressed relative to ampoules stored at -20 °C using an arbitrary value 1.00 for the -20 °C ampoules. The Arrhenius model was then fitted to the data to obtain predictions of the expected loss in potency over time. The predicted mean % loss per year, based on above data after 9 months storage at the different elevated temperatures, for the VISF technique and the spontaneous nucleation, were calculated and are shown in Table 5.5. The predicted % loss in FVIII potency per year tended to be greater for the spontaneous nucleation compared to controlled nucleation (e.g. 0.198 vs 0.003 respectively for storage at +4 °C). These results indicate a greater stability of the FVIII molecule in human plasma, when freeze-dried under controlled (VISF) nucleation compared to spontaneous nucleation.

These results for the highly concentrated plasma system, combined with the minimal effect observed on the low-concentration hGH formulation of cycles 1-2 and with previously published data reporting a negligible effect of controlled nucleation on protein stability in low concentrated systems [307], seem to suggest that the benefits of controlled nucleation may depend on concentration. More specifically, increased benefits seem to be observed for highly concentrated systems. In theory, if protein denaturation were mainly driven by adsorption onto the ice-water interface, the percentage of unfolded protein molecules should decrease when increasing the

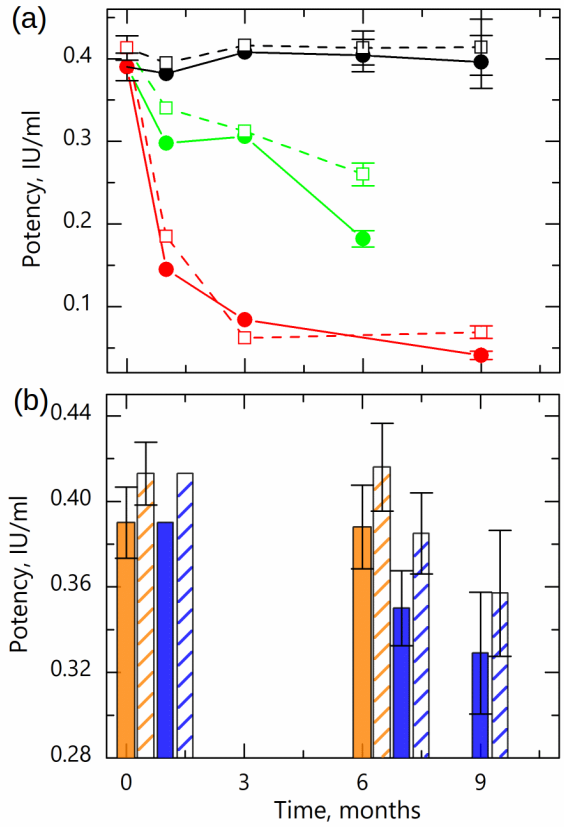

Figure 5.8: (a) Evolution of FVIII potency during storage at -20 °C (black line), 37 °C (green line) and 45 °C (red line), as measured after freeze drying with spontaneous (solid line, filled circles) or controlled (dashed line, empty squares) nucleation. (b) Evolution of FVIII potency during storage at 4 °C (orange bars), and 20 °C (blue bars), as measured after freeze drying with spontaneous (plain bars) or controlled (dashed bars) nucleation. The error bars displayed in the figure correspond to 95 % confidence intervals.

protein concentration, because of the finite extension of the ice-freeze concentrate interface [308]. This may also explain why, generally, the more concentrated the protein is, the more stable it is to freeze drying. Interestingly, the potential positive effect of controlled nucleation on protein stability is less pronounced for less stable, i.e. low-concentration, systems. This result represents an interesting observation, that warrants further investigation.

Table 5.5: *Mean predicted degradation rates expressed as % loss per year after storage for 9 months.

	*Mean predicted % loss per year [95% upper confidence limits]			
	-20 °C	+4 °C	+20 °C	+37 °C
Spontaneous nucleation	0.001 [0.001]	0.198 [0.382]	5.994 [8.598]	80.216 [80.702]
Controlled (VISF) nucleation	0.000 [0.000]	0.003 [0.009]	0.640 [1.155]	60.951 [66.728]

* These results are based on stability data obtained from 4 time points over a 9-months period.

Reconstitution Time of Plasma Samples

The reconstitution time is also a crucial parameter for many pharmaceutical lyophilizates, and the freeze drying process should be designed so as to guarantee the achievement of an injectable solution from the dried cake in a reasonable amount of time. Here, the reconstitution time of freeze dried plasma was measured after storage at different temperatures for 1, 3, 6 or 9 months, in the case of both controlled and spontaneously nucleated samples. The results of this analysis are reported in Table 5.6 and indicate a remarkable improvement in reconstitution time when the VISF technique was applied. This is evident already for the case of storage at low temperature, where VISF approximately halved the reconstitution time. For instance, it took about 4, 5.2 and 7.7 min to obtain a clear solution from the VISF samples stored for 6 months at -20 °C, 4 °C or 20 °C, respectively, while the corresponding times for conventional freezing were 11.7, 15.3 and 15.4 min.

Table 5.6: Reconstitution time (n=1) of freeze dried plasma after storage at controlled temperature.

	Reconstitution Time, min									
Time	-20 °C storage		4 °C storage		20 °C storage		37 °C storage		45 °C storage	
months	Spon.	VISF	Spon.	VISF	Spon.	VISF	Spon.	VISF	Spon.	VISF
1	13.5	7					21*	20	#	22
3	10.6	5.5					22	20	#	#
6	11.7	4	15.3	5.2	15.4	7.7	#	20.8*		
9	11.5	5			15.0	10.5			#	20*

* Some insoluble clumps still present; # Not going to reconstitute within 25 min

Therefore, when storage at high temperature is considered, application of the VISF technique made it possible to reconstitute samples that would not return to a liquid, clump-free solution if freeze dried by conventional freezing. For instance,

the VISF sample stored for 1 month at 45 °C could be reconstituted, while this was not possible in the case of conventional freezing. Similarly, it was possible to reconstitute the sample frozen by controlled nucleation and then stored for 6 months at 37 °C, even though some insoluble clumps were still present after 25 min, while the same result could not be achieved in the case of spontaneous freezing. When rehydrating a freeze dried product, the gas within the pores should be displaced by the reconstitution medium so as to allow wetting of the cake. Afterwards, hydration of the solid may take place. A large pore size may promote the displacement of gas from the cake, and this is probably the reason for the observed behaviour. This same explanation was proposed in a previous work [65], where application of the ice-fog technique shortened the reconstitution time of highly-concentrated protein formulations. A similar reduction in reconstitution time was observed when the depressurization technique was applied to highly concentrated monoclonal antibody solutions [67].

5.5 Conclusions

In previous studies, controlled nucleation by depressurization was reported to suppress glass fogging, i.e., the undesired migration of protein solutions up on the inner walls of glass vials during the freezing step of lyophilization, and to result in higher stability against shaking stress [67]. Application of the ice-fog technique [307] to lyophilized monoclonal antibody formulations stored at different temperatures reduced particle formation in highly concentrated systems. However, the addition of polysorbates resulted in an overall lower particle level, with no further advantage of controlled nucleation on protein stability. At low concentration, no difference with respect to particle formation between the controlled and spontaneously nucleated samples was detected. These results are in line with our data for cycles 1 and 2, where HPLC-SEC and a cell-based potency assay seemed to give evidence for no dramatic difference in the behaviour of hGH at low concentration when either VISF or spontaneous nucleation were used, even though improvements may be possible, for instance by adjusting the moisture content within the samples.

However, the results obtained for FVIII in cycles 3 and 4 seem to indicate a difference between the two freezing protocols, with the VISF technique resulting not only in improved process efficiency, but also in enhanced protein stability. This increase in protein activity may be related to the smaller ice-water interface resulting from application of controlled nucleation. As evident from the SEM images, the VISF technique promoted the formation of structures having larger ice crystals, that expose a reduced surface area compared to spontaneously nucleated samples. As a result, the risk that the protein adsorbs and denature at the ice interface is reduced, likely promoting the observed preservation of the native structure. However, it must also be borne in mind that FVIII is a much larger, multi-domain

protein (see Figure 5.2) with complex intermolecular interactions which can impact on its activity, and so direct comparison to a small protein like hGH may not be straightforward.

Overall, VISF was found to have no detrimental effects for hGH bioactivity, and for its aggregation behavior. In the case of freeze-dried plasma, increased stability and quicker reconstitution were observed. These observations, coupled with the positive effect of controlled nucleation on both process performance and product homogeneity, suggest that implementation of a controlled nucleation approach would be beneficial for the freeze-drying process.

5.6 A Quick Note on the Use of Vacuum Induced Nucleation in Different Containers

All the experiments prevously described were performed in non-coated vials or ampoules, but for some applications surface treatments are required to modify the chemical properties of the container. For instance, hydrophobic coatings are sometimes used to reduce residual volumes and improve dosing accuracy after reconstitution, or prevent disruption of the cake and sidewall fogging. Therefore, as a last point, the applicability of vacuum induced surface freezing for different types of containers has also been investigated.

Ultrapure water, and 0.025 % Tween 20 were used as model formulations, and filled in the following containers (glass type I, Schott supplied by Adelphi Tubes, Haywards Heath, UK):

- 3 ml siliconized ampoules;
- 3 ml non-siliconized ampoules;
- 2R siliconized vials;
- 2R TopLyo vials (Si-O-C-H hydrophobic coating).

The different containers were partially stoppered with 13 mm diameter halobutyl rubber lyo-closures (West Pharma, supplied by Adelphi). In all cases, a 1 ml filling volume was used, and controlled nucleation was performed on a LyoBeta 15 freeze dryer (Telstar Azbil SpA, Terrassa, Spain). The product was first equilibrated at -5 °C for about 1 h, and the pressure subsequently reduced to a low value (about 1 mbar), to trigger nucleation.

The test was repeated 3 times, and in all cases it was observed that VISF could not be successfully applied to containers with a hydrophobic coating (siliconized/TopLyo ampoules/vials) in the case of ultrapure water, because in this case it was impossible to achieve proper wetting of the container surface. Both

bubbling and boiling were observed in the solution while pulling the vacuum, well above the pressure required for inducing nucleation (see Figure 5.9).

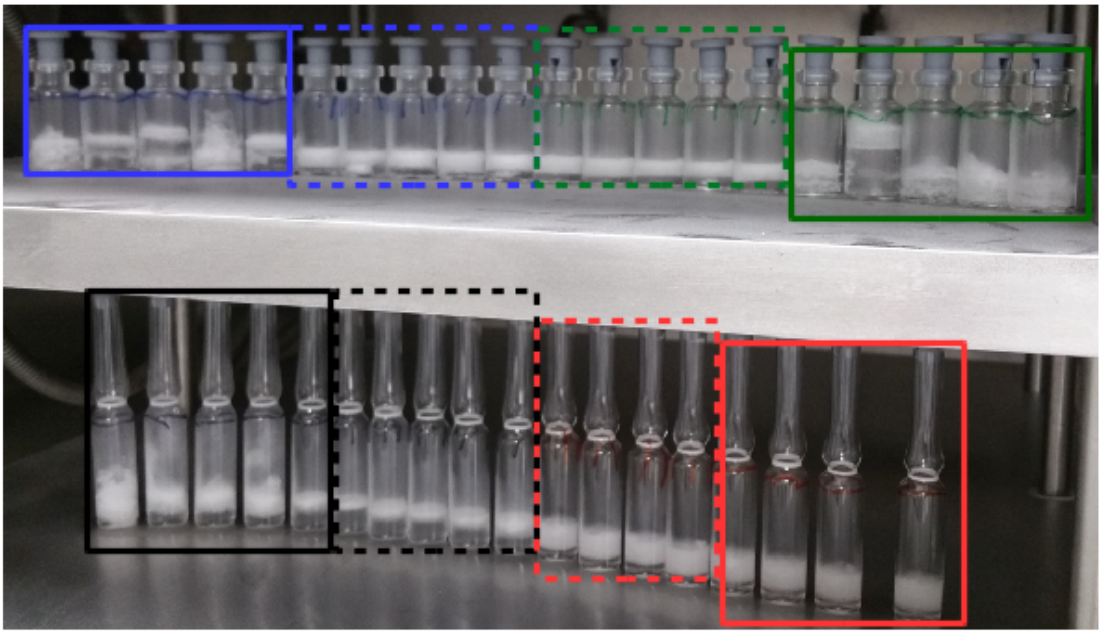

Figure 5.9: Frozen product obtained after VISF in siliconized (black) or non-siliconized (red) ampoules and siliconized (blue) or TopLyo (green) vials. Solid line: ultrapure water, Dashed line: 0.025 % Tween 20.

It was also found that the problem could be solved by adding a surfactant (Tween 20). The addition of surfactant resulted in improved wetting of the container surface, and in a diminished risk of bubbling/boiling phenomena (Figure 5.9). To the best of our knowledge, this phenomenon is not reported in the literature, and highlights the importance of the surface properties of the container for a successful control of nucleation.

List of Symbols

D_p	m	pore size
P	Pa	pressure
t	s	time
T	K	temperature
T_c	K	collapse temperature
T_g'	K	glass-transition temperature

T_m	K	holding temperature for ice crystals growth
T_n	K	nucleation temperature

Chapter 6

Optimizing the Freezing Step of Freeze Drying: A Trade-Off between Process Efficiency and Product Quality

In this Chapter, attention will be focused on the freezing step of freeze drying. As mentioned in the Introduction and in Chapter 4, the freezing step is crucial as it can influence the ice crystal size, and hence the degree of protein denaturation as a result of ice formation. Based on the results discussed in the previous Chapter, it is evident that the possibility to control the extension of the ice-water surface area would have a significant impact on protein stability.

The product morphology is determined by cooling rate and nucleation temperature, with a low cooling rate and/or high nucleation temperature promoting the formation of large ice crystals.

The morphology of the frozen product also has a strong impact on the rate of the subsequent drying phases [57–60, 148, 309]. Provided that neither shrinkage nor collapse of the cake occurs, the size of the ice crystals formed during freezing corresponds to the size of the pores within the dried product. This pore size, in turn, influences both the sublimation and desorption rates [59,83,84]. A large pore size speeds up the removal of water by sublimation, but is detrimental to the rate of desorption. Consequently, primary drying benefits from a large pore size, while the secondary drying rate is penalized. Moreover, the pore size influences the maximum temperature reached by the product as well, with higher temperatures reached in the case of smaller pores. Control of the maximum temperature reached by the product is of utmost importance for the freeze-drying process. If the maximum temperature overcomes a limit value, which is specific for each formulation, collapse of the cake is observed. A collapsed cake loses its esthetic properties and is generally characterized by a higher residual moisture content, which could damage the active

ingredients. Finally, the pore size of a dried product also affects the reconstitution time, as discussed in Chapter 5, which is a not negligible factor to be considered, especially for highly concentrated products. While low-concentrated formulations generally reconstitute in less than 1 min [310], it may take up to 1 h for higher concentrated systems to yield an injectable solution [311].

Hence, the picture of interrelated effects on protein stability during freezing as outlined in previous Chapters becomes even more complex when considering freeze-drying process duration and efficiency. The selection of optimal freezing conditions is therefore not straightforward, and may substantially benefit from a mathematical model capable of describing the effects of critical process parameters on both product quality and cycle efficiency.

6.1 Overview of the Chapter

With this objective in mind, a simulation approach will be presented in this Chapter, that makes it possible to understand the effect of critical input parameters during freezing.

In this framework, attention will be first focused on the possibility to predict the extension of the ice-water interface as function of process conditions used during freezing. Considering the crucial role of the ice surface previously discussed, much attention has been given to this problem in the literature, and empirical laws have been developed to respond to this request. Here, a mechanistic model for predicting the average size of ice crystals formed during freezing of pharmaceutical solutions will be presented. This model will be shown to provide physical insight on the relationship between ice crystal size, velocity of the freezing front, and temperature gradients within the product being frozen, which was postulated by the empirical laws previously proposed. Model simulations will be validated upon experimental observations obtained by scanning electron microscopy. The proposed model will be demonstrated to be valid over a wide range of freezing conditions, and for solutions containing both amorphous and crystalline solutes.

Afterwards, a Quality by Design (QbD) approach [85] will be proposed for the selection of optimal freezing conditions. QbD prioritizes process understanding and control in order to guarantee the desired characteristics of the final product [86,87]. The advantages of QbD are manifold and include increased manufacturing efficiency, reduction of costs, and higher level of assurance of product quality, thanks to a better understanding of phenomena involved. In the QbD framework, design space is a tool that may be used to provide useful information about the effect of input variables on output critical parameters. Here, the effect of cooling rate and nucleation temperature will be investigated using this tool. A comprehensive procedure, based on mathematical modeling, will be described, that allows calculation of the design space for the freezing process. Hence, in the framework of the freeze

drying process of biopharmaceuticals, some considerations on process performance will be advanced, and the impact of freezing conditions on process efficiency will be analyzed.

Finally, the design space approach will be applied to investigate protein stability during freezing, and two opposite scenarios will be hypothesized, depending on the relative stability of proteins in bulk or at interfaces. In the first one, the ice-water interface is identified as the dominant cause of denaturation for proteins with high bulk stability, while in the second case the duration of the freezing process itself is suggested to be the key parameter for proteins that are highly sensitive to cold denaturation. The results of experimental tests on lactate dehydrogenase and myoglobin will be reported, that support the model results. While the complexity of the phenomena involved in the freezing of a protein formulation can hardly be fully described by a mechanistic model, the proposed approach still aims to provide some useful indications. These indications may help to guide lyo-professionals into the selection of appropriate freezing conditions, which are beneficial to the preservation of protein biological activity, and ensure process economic efficiency at the same time [1].

6.2 A Mathematical Model for the Prediction of the Ice Crystal Size Distribution in Freeze Dried Formulations

As evident from the previous discussion, the ice-water interface represents a major cause of protein denaturation during the freezing and freeze-drying processes. The larger this interface is, the greater the risk of losing therapeutic activity is. A strict control should therefore be carried out on the extension of this surface, but a preliminary requirement to achieve this result is the possibility to predict the effect of process variables on the ice crystal size.

Prediction of the average size of solvent crystals formed during freezing of a solution would also prove beneficial in several other fields. For instance, the ice crystal size has a strong impact on the organolectic properties of frozen food [316], while in the case of metal alloys the crystal size greatly affects the mechanical properties of the material [317]. As a consequence of the importance of the subject, many researchers tried to find a quantitative relationship that correlates the thermal history of the product being frozen with the average size of solvent crystals. However, all these models are essentially empirical formulas.

For example, early in 1991 Bald [318] proposed an empirical formula that correlates, in the case of dendritic growth, the average size of ice crystals to cooling

[1] Part of this Chapter is based on already published works [312–315]

rate ($\partial T/\partial t$),

$$D_p = \alpha\left(\frac{\partial T}{\partial t}\right)^{-\beta} \tag{6.1}$$

where α and β are experimentally determined parameters.

In the literature a number of empirical laws sharing the same structure has also been proposed. These equations correlate the crystal size with the freezing front rate (ν) and the temperature gradient within the frozen product (θ),

$$D_p = C\nu^{-\lambda_1}\theta^{-\lambda_2} \tag{6.2}$$

where C, λ_1 and λ_2 depend on type of product, application and processing conditions. However, these parameters were found to change even within the same application, as summarized in Table 6.1.

Table 6.1: Empirical formulas for crystal sizing

Application	Correlation	Reference
Metal solidification at low rates-1	$D_p \propto \nu^{-1}\theta^{-1}$	[317]
Freezing of apples	$D_p \propto \nu^{-0.5}\theta^{-0.5}$	[319]
Freezing of pharmaceutical solutions	$D_p \propto \nu^{-0.5}\theta^{-0.5}$	[320]
Alloy solidification at high rates	$D_p \propto \nu^{-0.25}\theta^{-0.5}$	[321]
Metal solidification at low rates-2	$D_p \propto \nu^{-0.5}$	[317]

Starting from these premises, it is clear that the development of a relationship, based on balance equations rather than empirical observations, and applicable over a wide range of conditions, would be desirable. In the following, a mathematical model will be proposed to predict the average size of ice crystals, and its distribution within the product being frozen, with the aim to cover this gap. The model will use, as input variables, the freezing front rate and temperature gradients within the product. The final equation will still include a physical parameter, related to the superficial morphology and crystal habit of the crystals, that can hardly be determined from direct experimental measurements. However, this parameter can be obtained by regression of experimental data without losing any theoretical background of the system, and is independent of the freezing protocol used, thus making the model more generally applicable. Scanning Electron Microscopy (SEM) of lyophilized solutions will also be used to validate the model predictions, exploiting the coincidence between the size of pores formed after lyophilization, and the dimension of ice crystals formed after freezing.

6.2.1 Mathematical Formulation

The model here presented applies to the crystal growth phase, once nucleation has occurred. Temperature is assumed to be constant within the liquid layer, as well as

at the freezing front. More specifically, temperature at the freezing front is assumed to be equal to the equilibrium freezing temperature of the solvent. This assumption is realistic if we take into account that temperature gradients within the frozen zone are orders of magnitude greater than those within the liquid phase.

For the mathematical derivation, the product is divided into three parts, namely, solid, mushy and liquid, where the mushy zone refers to a suspension of solid crystals in liquid. A given volume of the sample is then considered, having cylindrical shape with base diameter D and length Δz, and in which crystallization is taking place. In other words, Δz can be considered as the fraction of the mushy zone where crystal growth is occurring. Hence, the domain under investigation modifies its position during the process, as it follows the mushy zone. A schematic of this domain is shown in Figure 6.1, where a typical glass vial for the lyophilization of pharmaceuticals has been considered.

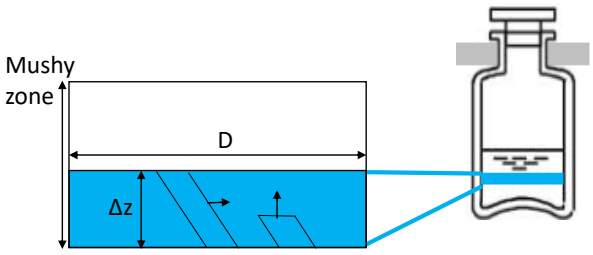

Figure 6.1: Scheme of the domain investigated by the model here presented.

As illustrated in Figure 6.2, it is assumed that all crystals are cylinders with diameter D_p, and having an inclination described by the angle δ with respect to the axial direction z. It should be further taken into account that δ is also linked to the tortuosity τ of the porous structure resulting by the freeze-drying of the frozen product, $\tau = \cos\delta^{-1}$.

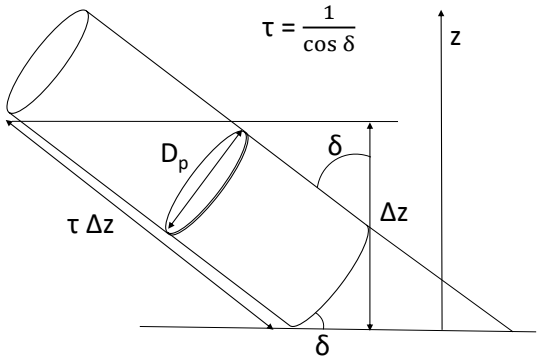

Figure 6.2: Crystals shape as assumed by the model.

During the crystal growth process, the freezing front moves along the vial height; hence, the axial interval Δz under investigation changes its position during the process, as it follows the freezing front. The vial height is therefore discretized into a finite number i = 1,2...n of axial intervals, and the number and diameter of ice crystals within the i^{th} of these intervals will be denoted with subscript i.

The enthalpy balance for the given volume can be written as,

$$\begin{pmatrix}\text{heat removed}\\ \text{by the}\\ \text{frozen product}\end{pmatrix}+\begin{pmatrix}\text{heat transferred}\\ \text{from the}\\ \text{lateral surface}\end{pmatrix}+\begin{pmatrix}\text{heat generated}\\ \text{by}\\ \text{crystallization}\end{pmatrix}+\begin{pmatrix}\text{enthalpy change}\\ \text{due to new}\\ \text{interface generation}\end{pmatrix}=0$$

$$-k_{fr}\theta_i\pi\frac{D^2}{4}-\pi D\Delta z h\Delta T+\frac{dV_{ice,i}}{dt}\varrho_{ice}\Delta H_f-\frac{dS_{ice,i}}{dt}\gamma=0 \tag{6.3}$$

where k_{fr} and θ are the thermal conductivity and temperature gradient within the frozen zone respectively, $\Delta T = T - T_{air}$ the temperature difference between air and the product being frozen, V_{ice} and S_{ice} the ice crystals volume and surface, respectively.

Integrating Equation 6.3 in time from 0 to Δt_i, that is the time interval required for the mushy zone to move ahead of Δz, it is possible to write,

$$-k_{fr}\bar{\theta}_i\pi\frac{D^2}{4}\Delta t_i-\pi D\Delta z h\Delta T\Delta t_i+\Delta V_{ice,i}\varrho_{ice}\Delta H_f-\Delta S_{ice,i}\gamma=0 \tag{6.4}$$

where $\bar{\theta}_i$ is the mean value of the temperature gradients within the frozen zone during Δt_i. Then, assuming that the initial volume of crystal nuclei is negligible compared to the final dimension of ice crystals,

$$\Delta V_{ice,i}=N_i\pi\frac{D_{p,i}^2}{4}\tau\Delta z \tag{6.5}$$

$$\Delta S_{ice,i}=N_i\pi D_{p,i}\tau\Delta z a_{S,i} \tag{6.6}$$

where N_i and $D_{p,i}$ are the final number and dimension of ice crystals, respectively, and a_S accounts for the real ice crystals habit and surface morphology, and is equal to the ratio between the real surface area and the fictitious surface area calculated considering the ice crystals as perfect cylinders of height Δz,

$$a_{S,i}=\frac{\text{real surface area}}{N_i\pi D_{p,i}\Delta z\tau} \tag{6.7}$$

The coefficient a_S is actually the product of two terms. The first, b_1, corrects for the real ice crystal habit, and can be assumed to be constant. The second one, $b_{2,i}$, takes into account surface irregularities, and can be hypothesized to be inversely proportional to the total surface area of the crystals. The greater the total surface area is, the smaller the contribution given by surface irregularities is,

$$a_{S,i}=b_1 b_{2,i} \tag{6.8}$$

Substituting Equations 6.5 and 6.6 into Equation 6.4, and dividing all terms by Δt_i,

$$-k_{fr}\bar{\theta}_i\pi\frac{D^2}{4} - \pi D\Delta z h\Delta T + N_i\pi\frac{D_{p,i}^2}{4}\varrho_{ice}\tau\nu_i\Delta H_f - N_i\pi D_{p,i}\tau\nu_i\gamma a_{S,i} = 0 \quad (6.9)$$

where ν is the freezing front velocity.

It is further assumed that the mass of ice that crystallizes in $\frac{1}{4}\pi D^2\Delta z$ is directly proportional to $\bar{\theta}$. This assumption is much more physically grounded than simply assuming D_p to be proportional to $\theta^{-\lambda_2}$, and is supported by the observation that the heat released by ice crystallization, and that is removed, increases if $\bar{\theta}$ increases,

$$N_i\pi\frac{D_{p,i}^2}{4}\varrho_{ice}\tau\Delta z = a\bar{\theta}_i \quad (6.10)$$

where a is a proportionality coefficient.

Substitution of Equation 6.10 into Equation 6.9 leads to the following expression

$$N_i\pi D_{p,i}\tau\nu_i\gamma a_{S,i} = -k_{fr}\bar{\theta}_i\pi\frac{D^2}{4} - \pi D\Delta z h\Delta T + a\bar{\theta}_i\Delta H_f\frac{\nu_i}{\Delta z} \quad (6.11)$$

This last Equation implies that the rate of generation of crystals surface area increases with the temperature gradients within the frozen zone, and thus with the total mass of water that crystallizes, and with the freezing front velocity. This is again in accordance with experimental observations, as it is widely known that a slower freezing process results in bigger ice crystals, and therefore smaller rate of generation of new surface area.

Next, the mass balance for the given volume of liquid being frozen is considered,

$$\sum_{i=1}^{n} N_i\pi\frac{D_{p,i}^2}{4}\varrho_{ice}\tau\Delta z = \phi m_w \quad (6.12)$$

where m_w is the total mass of water in the solution and ϕ is a coefficient close or equal to one, at least for dilute solutions, that accounts for the mass of water which remains in the amorphous phase and does not crystallize. For crystalline materials, nearly all water gets frozen and it is possible to assume $\phi = 1$ [322]. In the case of amorphous solutes, the final product usually contains some unfrozen water, but for the dilute solutions typically used in pharmaceutical freeze drying the value of ϕ is anyway close to one [148]. In the following, it will be therefore assumed for the sake of simplicity that ϕ equals 1.

Combining Equations 6.10 and 6.12 it is possible to determine an expression for a,

$$a = \frac{m_w}{\sum_{i=1}^{n}\bar{\theta}_i} \quad (6.13)$$

and thus Equation 6.10 can be rewritten as,

$$N_i \pi \frac{D_{p,i}^2}{4} \varrho_{ice} \tau \Delta z = m_w \frac{\bar{\theta}_i}{\sum_{i=1}^{n} \bar{\theta}_i} \tag{6.14}$$

Substituting Equation 6.14 into Equation 6.9 and rearranging gives,

$$D_{p,i} = \frac{16 m_w \bar{\theta}_i \gamma a_{S,i} \nu_i}{\varrho_{ice}(4 m_w \bar{\theta}_i \nu_i \Delta H_f - k_{fr} \bar{\theta}_i D^2 \pi \Delta z \sum_{i=1}^{n} \bar{\theta}_i - 4\pi \Delta^2 z \sum_{i=1}^{n} \bar{\theta}_i D h \Delta T)} \tag{6.15}$$

Equation 6.15 describes the effects of freezing front velocity ν and temperature gradients θ on D_p. The freezing process is extremely complex and several other factors should rigorously be taken into account. However, previous works [317, 319–321] have shown that the temperature gradients within the frozen zone and the freezing front velocity are the most relevant and allow a satisfactory description of the process.

As previously mentioned, the a_S coefficient, which takes into account the real ice crystals habit and surface irregularities, is the product of a constant term and of a term which is assumed to be inversely proportional to the surface area of the crystals. Equation 6.10 states that the volume of ice crystallized in $\frac{1}{4}\pi D^2 \Delta z$ is proportional to $\bar{\theta}$. Hence, if the volume is proportional to $\bar{\theta}$, the surface area of the crystals can be assumed to be proportional to $\bar{\theta}^{2/3}$. As a result:

$$a_{S,i} = b_1 b_{2,i} = \frac{b}{\bar{\theta}_i^{2/3}} \tag{6.16}$$

where b is determined by regression of experimental observations for pore size vs product thickness.

Unlike previous approaches, Equation 6.15 has been obtained from balance equations. Therefore, it might be used to provide insight into the freezing process behavior, and will be demonstrated in the following to be applicable over wide ranges of conditions. Equation 6.15 also has the advantage of automatically respecting the mass balance.

However, it must be pointed out that the model is not entirely theoretical. Even if balance equations have been used, the numerical value of one parameter, namely, γb, needs to be calculated from experimental data. A simulation approach, that involves molecular dynamics, has also been proposed to compute the value of this parameter, as discussed in [314], but further investigation is needed in this direction. In any case, all the terms used in the final equation have a physical meaning and a clear explanation.

Simplified model for dilute solutions

The model can be further simplified for dilute solutions. More specifically, it can be assumed that the term ND_p is large and hence heat transferred by natural convection is negligible. For example, if water constitutes 95% of the whole system mass and the air inside the freezing chamber is stationary, heat removed by natural convection is only about 0.5% of the heat generated by water crystallization.

Moreover, it is possible to write that,

$$\varepsilon\pi\frac{D^2}{4} \approx N_i\pi\frac{D_{p,i}^2}{4}\tau \tag{6.17}$$

where ε is the ratio between the volume of ice and the total volume of the system. Thus, combining Equation 6.9 with Equation 6.17,

$$D_{p,i} \approx \frac{4\varepsilon\gamma a_{S,i}\nu_i}{\varepsilon\varrho_{ice}\Delta H_f\nu_i - k_{fr}\bar{\theta}_i} \tag{6.18}$$

where a_S is given in Equation 6.16.

6.2.2 Calculation of ν and θ

The values of ν and θ, which are then required for the evaluation of D_p using Equations 6.15 or 6.18, take into account all those parameters that influence the crystal size, such as the cooling rate and the nucleation temperature, which are therefore necessary inputs for the modelling approach. The nucleation temperature, in particular, is a stochastic variable, randomly distributed within the batch, unless a controlled nucleation technique is used.

Two different approaches may be used for the calculation of ν and θ. On the one hand, the entire temperature profile may be calculated using the mathematical model described in Nakagawa *et al.* and Pisano and Capozzi [320, 323]. Details of this model can be found in the original papers, but a short description of its main features is also reported in Appendix C.

An example of the temperature profile predicted by the model, and compared with the experimental one, is shown in Figure 6.3. The liquid solution is first supercooled to a value below the equilibrium freezing point (segment A-*). Supercooling represents a metastable state, during which water molecules tend to form clusters with long-living hydrogen bonds [324]. However, these clusters are still unstable, and break up quickly. When the temperature is low enough to allow crossing of the energy barrier for the nucleation process, ice crystallization occurs rapidly in the whole product (point * in Figure 6.3). In all pharmaceutical solutions heterogeneous nucleation is observed, and the degree of supercooling often lies in the range 10-15 °C or more [58]. A sharp increase in product temperature to a value close to

the equilibrium freezing point is observed at the onset of nucleation. The ice crystals growth (segment B-C) then proceeds through the addition of molecules to the interface. Here, the latent heat of crystallization is almost compensated for by the heat removed through the already frozen product, and the temperature remains nearly constant. Finally, to ensure complete solidification, the frozen product is typically cooled down to -40 / -50 °C (segment C-D).

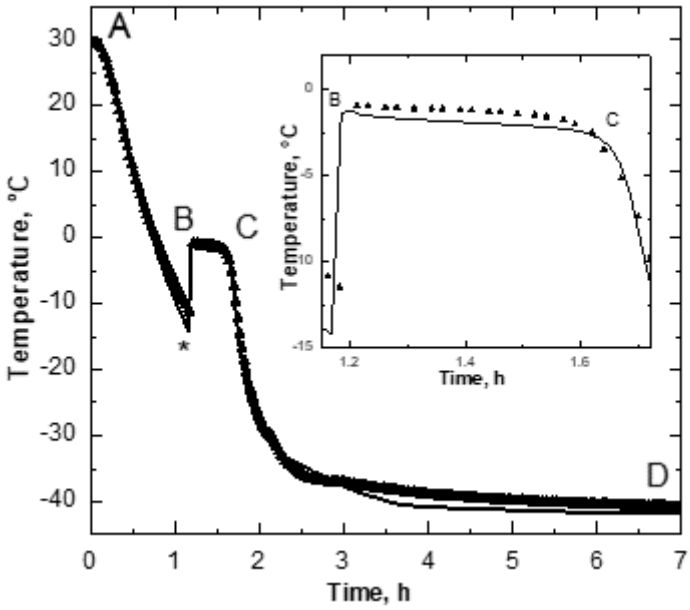

Figure 6.3: Evolution of temperature during freezing as measured experimentally (▲) and as predicted by the Nakagawa *et al.* [320] and Pisano and Capozzi [323] model (—) for test A in Table 6.2 (mannitol 5% w/w). The inset shows an enlargament of the temperature profile during the ice crystals growth phase (segment B-C).

The terms ν and θ can afterwards be determined from the temperature profiles as calculated by model simulations. More specifically, if the evolution of the temperature profile is known, e.g. see Figure 6.4, the freezing front position can be easily tracked, as its temperature corresponds to the equilibrium freezing value.

The freezing front rate, ν, can hence be calculated as the ratio between the variation of the freezing front position Δz_f and the time interval Δt, as illustrated in Figure 6.4,

$$\nu = \frac{\Delta z_f}{\Delta t} \tag{6.19}$$

Once the temperature profile along the whole product height is known, e.g. see Figure 6.5, the temperature gradient θ corresponds to the slope of the curve,

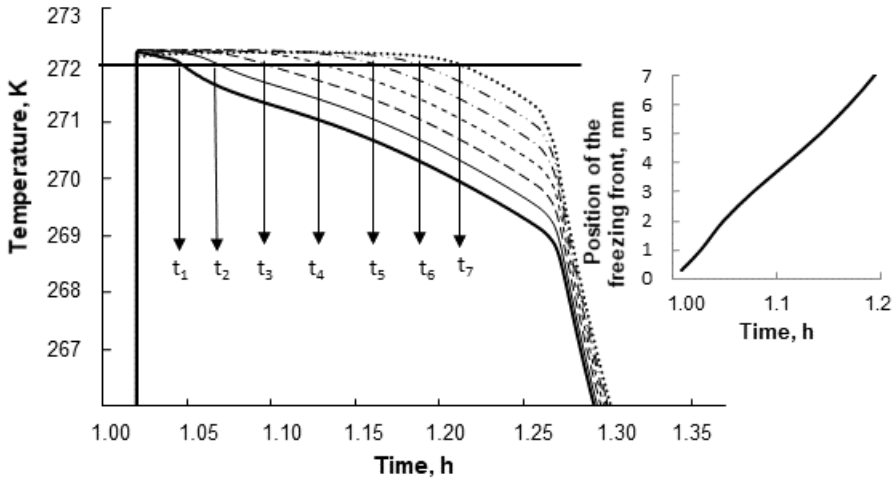

Figure 6.4: Evolution of the temperature profile within the solution being frozen at (▬) $z = 1$ mm, (—) $z = 2$ mm, (– – –) $z = 3$ mm, (- - -) $z = 4$ mm, (– · –) $z = 5$ mm, (– · · –) $z = 6$ mm, (· · · · ·) $z = 7$ mm for test A in Table 6.2.

calculated at a given time, within the frozen phase.

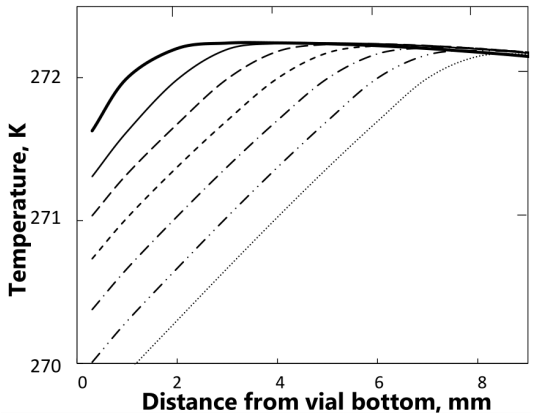

Figure 6.5: Temperature profile within the solution being frozen as observed at (▬) t_1, (—) t_2, (– – –) t_3, (- - -) t_4, (– · –) t_5, (– · · –) t_6, (· · · · ·) t_7 for test A in Table 6.2.

Alternatively, an iterative procedure, based on the following steps, may be used [270]:

1. Definition of the shelf temperature profile over time ($T_{shelf}(t)$).

2. Discretization of the sample height into n horizontal layers, having height Δz, identified by the index i = 1,2,3....n. Here, $\Delta z = 1$ mm was used.

3. A value for the average ice crystal size, $D_{p,i}$, for the layer i under investigation is hypothesized.

4. The average number of ice crystals N_i within the layer is calculated from,

$$\varepsilon \frac{\pi D^2}{4} = N_i \pi \frac{D_{p,i}^2}{4} \tau \tag{6.20}$$

5. Once N_i and $D_{p,i}$ are known, it is possible to calculate the heat generated by crystallization Q_v,

$$Q_v = N_i \pi \frac{D_{p,i}^2}{4} \varrho_{ice} \tau \Delta z \Delta H_f \tag{6.21}$$

It is also possible to evaluate the enthalpy change due to generation of new ice-freeze concentrate interface Q_s,

$$Q_s = N_i \pi D_{p,i} \tau \Delta z \gamma a_{S,i} \tag{6.22}$$

6. The time Δt_i, required for the freezing front to move ahead of Δz, can be calculated from the following energy balance equation

$$Q_v - Q_s = K_{v,e}(T_{eq} - \bar{T}_{shelf}) A_b \Delta t_i \tag{6.23}$$

where $K_{v,e}$ is the overall heat transfer coefficient between freezing front and shelf, A_b the surface area of the vial base in contact with the shelf, and $\bar{T}_{shelf}$ the average value of shelf temperature during Δt_i. $K_{v,e}$ can be calculated from the heat transfer coefficient $K_{v,f}$ between the shelf and the product base, as follows

$$K_{v,e} = \frac{1}{1/K_{v,f} + i\Delta z / k_{fr}} \tag{6.24}$$

7. Once Δt_i is known, it is possible to compute the freezing front velocity ν_i,

$$\nu_i = \frac{\Delta z}{\Delta t_i} \tag{6.25}$$

The temperature gradients θ_i within the already frozen zone can also be calculated, by imposing the following equivalence of heat fluxes,

$$k_{fr} \theta_i = K_{v,e}(T_{eq} - \bar{T}_{shelf}) \tag{6.26}$$

8. Once the freezing front velocity ν_i and temperature gradients θ_i within the frozen zone are known, it is possible to compute $D_{p,i}$, using Equations 6.15 or 6.18.

9. The value of $D_{p,i}$ obtained at point 8 is used as a new guess estimate, and the procedure is iterated starting from step 3, until the value hypothesized at step 3, and that calculated at point 8, do not differ significantly.

10. Points 3-9 are repeated for all values of i. The crystal size distribution along the vertical axis of the sample is thus obtained, and an average value can be calculated, as well

$$D_p = \frac{\sum_{i=1}^{n} D_{p,i}}{n} \tag{6.27}$$

The iterative procedure just outlined allows a faster estimation of ice crystal size compared to the 2D model by Nakagawa *et al.* [320] and Pisano and Capozzi [323]. The mean value of the ice crystal size is well predicted by this iterative approach (the difference with the 2D simulations is less than 5 %), even though the amount of information that can be extracted from this 1D model is limited compared to the case of a 2D simulation.

6.2.3 Experimental Approach

Mannitol solutions at different concentrations (5% and 10% w/w) and sucrose or dextran (molecular weight ≈ 40,000) solutions at 5% w/w were prepared from distilled water and mannitol, sucrose or dextran powder; all reagents were purchased from Sigma Aldrich and used as received. The vials employed (type 1, 10R, 45x24 mm, Schott AG, Germany), having 24 mm diameter and 45 mm height, where filled with 3 ml of sample solution, while igloo stoppers (type 1319 4432/50/ Westar, West Pharmaceutical Services, Terrassa, Spain) were used as closures. A freeze-dryer LyoBeta 25 (Telstar, Terrassa, Spain) was used for the freeze-drying cycles. A system of T-type copper/constantan miniature thermocouples was used to monitor the temperature of the shelves and of the product. More specifically, the thermocouples were inserted close to the center of the vial bottom. This is the last portion of the product to be dried, and, thus, the measure performed by thermocouples should be accurate during almost the whole drying process, because the risk of thermocouples losing contact with ice is minimized. Moreover, for a better evaluation of the nucleation temperature distribution, the vials were filmed during freezing, and the distribution of nucleation times evaluated from analysis of the video. More specifically, the nucleation time was recorded for the front row of the vials. Then, knowledge of cooling rate and specific heat capacity of the solution being frozen, allowed conversion of the nucleation time distribution into a nucleation temperature distribution.

The experiments performed for the validation of the model were carried out using two different freezing modalities, namely conventional freezing and the controlled nucleation technique (vacuum induced surface freezing) introduced in Chapter 5. For the conventional freezing tests, the tubing vials were placed onto the freeze-dryer shelves at room temperature and then the refrigerating fluid temperature was decreased with a precise cooling rate ($\partial T/\partial t$) until a value of 228 K was reached. As regards the controlled nucleation tests, the product was first equilibrated at T_n for about 45 min and, then, the pressure was reduced to 130 Pa for 1 min. The pressure within the chamber was subsequently released to the atmospheric value, and the temperature of the heat transfer fluid held at a value (T_m) below the onset of ice melting for about 1 h. To complete the freezing of the product, the fluid temperature was finally decreased to 228 K and held for at least 1 h. As regards the controlled freezing technique, different combinations of T_n and T_m values were considered, as detailed in Table 6.2. The subsequent primary drying phase was performed at 10 Pa and 263 K, for mannitol and dextran, or 253 K, for sucrose.

Table 6.2: Operating conditions of experimental tests

	solution	nucleation type	$\frac{\partial T}{\partial t}$, K min^{-1}	T_n	T_m
A	mannitol 5% w/w	conventional	0.8	-	-
B	mannitol 10% w/w	conventional	0.8	-	-
C	mannitol 5% w/w	conventional	0.3	-	-
D	mannitol 5% w/w	conventional	0.1	-	-
E	sucrose 5% w/w	conventional	0.8	-	-
F	dextran 5% w/w	conventional	0.8	-	-
G	mannitol 5% w/w	controlled	-	263	263
H	mannitol 5% w/w	controlled	-	268	268
I	mannitol 5% w/w	controlled	-	268	263

The pore dimension of the products obtained after freeze-drying, which corresponds to the dimension of ice crystals formed during freezing, was analyzed using a Scanning Electron Microscope (SEM, FEI type, Quanta Inspect 200, Eindhoven, the Netherlands). SEM images were taken at various positions along the previously metallized sample in order to obtain a quantitative estimation of within-vial heterogeneity. This quantitative analysis was done in terms of average size (D_p) and its distribution. An example of SEM images of the samples analyzed is shown in Figure 6.6.

6.2.4 Validation of the Model

The number of crystal nuclei strongly depends on both nucleation temperature and cooling rate. Variables such as cooling rate and nucleation temperature are taken into account within the model by ν and θ. A decrease in the cooling rate or an increase in the nucleation temperature translate into a decrease in ν and θ; according to the model, this means that the specific surface area exposed by the ice crystals decreases, and therefore larger ice crystals are formed.

In addition, if the time left for temperature equilibration is too short, large temperature gradients are generally observed within the liquid being frozen, and crystal nuclei are not uniformly distributed [58]. As general guideline, short equilibration times result in signifcant crystal size heterogeneity.

Moreover, the final product structure is linked to the crystals aspect ratio which, in turn, depends on the crystals growth rate. If the growth rate is controlled by diffusion, crystals mainly grow in the direction where the mass transfer rate is larger, assuming a hoppered shape. On the contrary, if the rate determining step is the surface kinetics process, that is the process in which atoms are incorporated into the crystal, the aspect ratio is mainly determined by the differences in the intrinsic growth rate of the different crystal surfaces [325, 326].

The differences in number and dimension of crystals along the product height can also be explained assuming that, in some situations, crystals preferentially grow in the radial direction with respect to the axial one, and vice versa. In the case of preferential growth along the radial direction, crystals are mostly confined within the zone where they are formed and thus the product can be divided into layers where crystal growth is independent of one another and where the number of crystals is determined by the number of nuclei formed during nucleation. On the contrary, in the case of preferential growth in the axial direction, these layers are no more independent of one another and the number of ice crystals formed in a layer is also influenced by the number of crystals coming from the surrounding layers. The parameter b defined in Equation 6.16 is linked to all these microscopic aspects of cristal growth, even though the model derived is macroscopic.

A list of numerical values for the parameters used in the model is reported in Tables 6.3 and 6.4. As can be seen from Table 6.4, the value of the parameter γb is the same for the two amorphous solutes tested, namely sucrose and dextran, while it is significantly different for mannitol. Moreover it is not affected by cooling rate, nucleation temperature or solute concentration. In addition, γb increases for the controlled nucleation technique with respect to conventional freezing.
Attention will be first focused on mannitol, a crystalline excipient, and for the case of conventional freezing. The parameter γb was determined minimizing the sum of squared residuals between experimental observations and calculated data, as shown for test A in Figure 6.7. Using this one parameter only, the entire crystal size profile could be predicted. This same parameter, calculated for 5% w/w mannitol and 0.8

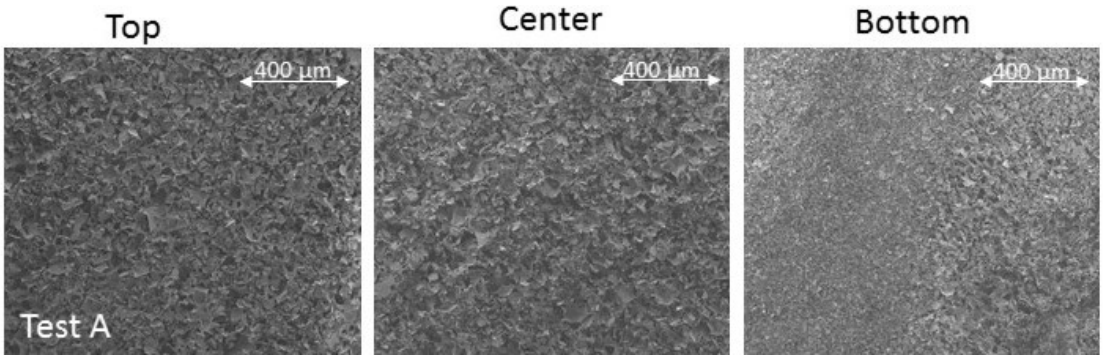

Figure 6.6: Scanning electron microscope pictures of mannitol 5% w/w lyophilized samples (test A in Table 6.2).

Table 6.3: Numerical values of the parameters employed

k_{fr}	2.5	W m^{-1} K^{-1}
h	5	W m^{-2} K^{-1}
ϱ_{ice}	918	kg m^{-3}
ΔH_f	333500	J kg^{-1}

K min^{-1} as cooling rate, was also found to be valid for a different solid content (10% w/w in the case of test B) or a different cooling rate (0.3 K min^{-1} for test C), as evident from Figure 6.7. Both the detailed model and the simplified one could fit fairly well the experimental distribution, also at the higher solid content (10% w/w), suggesting that model predictions are consistent with experimental observations in the whole range 5 - 10% w/w. This was true also for the simplified model, even though it was derived for dilute solutions. This suggests that the simplified model can be effectively applied to almost all typical pharmaceutical formulations, whose concentration is generally smaller than 10% w/w.

In Figure 6.7, the same image analysis technique described in Chapter 5 was used. More specifically, a large number (approximately 100) of pores per image was manually selected and each pore was approximated to an ellipse. Then, the dimension, D_p, attributed to the pore was the diameter of the circle having the same area to perimeter ratio of the approximating ellipse. In the case of this traditional, manual approach, the time required for the analysis of each image is ranging from 30 min to 1 h, depending on the number of pores that have to be analyzed. However, an alternative, automatized approach, based on image segmentation, may be employed, considerably reducing the amount of time needed. Details on this method are not reported here for the sake of conciseness, but can be found in [312]. In Table 6.5, the product morphology, as measured using the traditional image analysis

Table 6.4: Numerical values of parameter γb

	γb, J K$^{2/3}$ m$^{-8/3}$	
solute	conventional freezing	controlled nucleation
mannitol	$7{\cdot}10^4$	$26{\cdot}10^4$
sucrose	$23{\cdot}10^4$	-
dextran	$23{\cdot}10^4$	-

approach or the segmentation technique and as predicted by the model, was characterized using the pore dimension in the center of the product cake for different operating conditions and concentrations. As can be observed, the simplified model could estimate with good accuracy the pore size, giving results that were in line with both the traditional image analysis approach and the segmentation technique. It is also interesting to notice that the model could effectively predict the crystal size increase with decreasing cooling rate (tests A and D), which, as previously discussed, is taken into account by the model in the ν and θ terms.

Table 6.5: Crystal size in the center of the product for different mannitol concentrations and different freezing conditions, as evaluated by analysis of SEM images (traditional technique and segmentation approach) and by the simplified model.

	D_p, μm		
	Traditional	Segmentation	Simpl. Model
test A	28	27-30	28
test B	24	25	25
test D	40	45	39

Attention will now be shifted to the case of two amorphous solutes, sucrose and dextran, studied in tests E and F, respectively. For these two solutes the same γb value was estimated (see Table 6.4), and even the cooling rate and freezing protocol were the same for both tests. However, the two samples were characterized by a significantly different nucleation temperature, 267 K for sucrose and 258 K for dextran. This is reflected by the much smaller pores for test F if compared to test E, as can be seen in Figure 6.8. Therefore, both models consistently predicted that crystal size increases as nucleation temperature increases. Figures 6.7 and 6.8 suggest that the models may be applied to both crystalline and amorphous systems.

The models were also validated using a different freezing protocol, namely the vacuum induced surface freezing introduced in Chapter 5. Three different tests using mannitol (G, H and I in Table 6.2) were carried out changing the values of T_n and T_m and, thus, modifying the product nucleation temperature. As can be seen in Figure 6.9, both models gave accurate quantitative estimations of the average pore size and its variation along the axial direction of the product, demonstrating

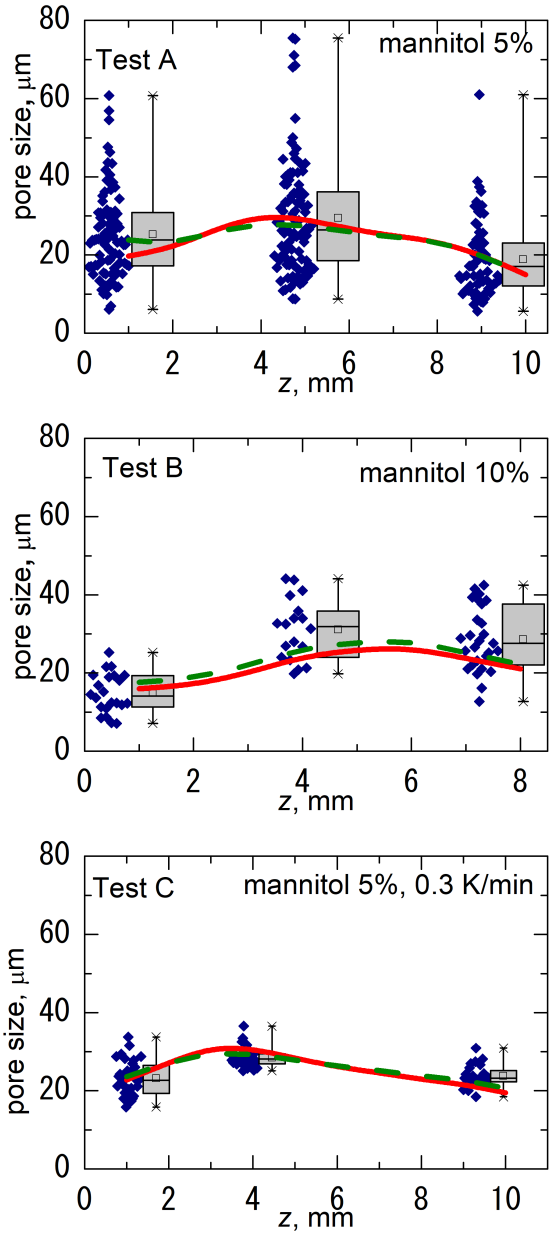

Figure 6.7: Comparison between detailed model predictions (Eq. 6.15, solid red line), simplified model predictions (Eq. 6.18, dashed green line) and SEM observations (box plots) for the average pore size in the case of tests A, B and C. For box plots, top bar is maximum observation, lower bar is minimum observation, top of box is third quartile, bottom of box is first quartile, middle bar and internal square are median and average value respectively and bullets are original data.

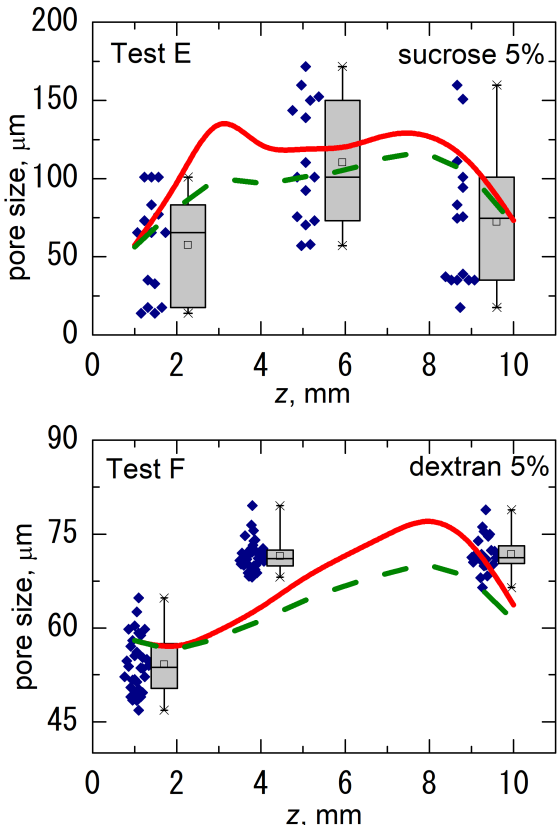

Figure 6.8: Comparison between detailed model predictions (solid red line), simplified model predictions (dashed green line) and SEM observations (box plots) for the average pore size in the case of tests E and F (sucrose or dextran 5% w/w).

that they can effectively be used independently of the freezing protocol. The pore size generally increases as T_n increases; accordingly, for test I the average pore dimension predicted by the model was greater (100 μm, data not shown) with respect to test G (85 μm), in line with experimental data.

In contrast with the general effect of T_n, the pore size of samples obtained in test H was significantly smaller than that observed for test G (see Figure 6.9). This result may be related to the different value of T_m chosen for the two tests. Test H was carried out at a higher value of T_m (268 K) and this temperature may have been too high to allow completion of freezing during the holding phase. As a consequence, part of the solution froze later on, during the ramp to 228 K, and freezing of that part occurred as in a conventional freezing cycle. This phenomenon was already observed by Oddone *et al.* [83] and the simulated temperature profiles for test H confirmed that the growing phase could not be completed during the holding time at T_m, but occurred at least in part during the temperature ramp

from T_m to 228 K (inset of Figure 6.9). The model here presented could therefore predict crystal size even for this very particular situation.

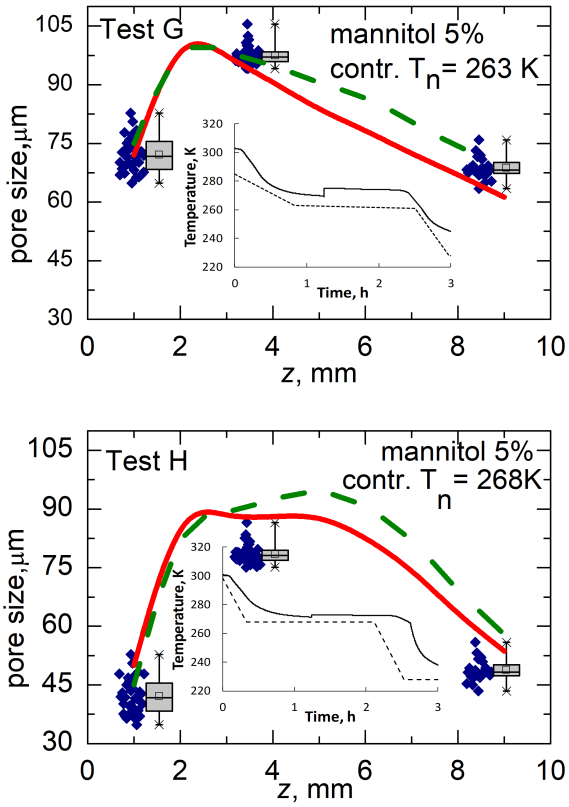

Figure 6.9: Comparison between detailed model predictions (solid red line), simplified model predictions (dashed green line) and SEM observations (box plots) for the average pore size in the case of tests G and H (controlled freezing). The insets show the temperature profiles of the product (solid line) and of the shelf (dashed line).

Finally, the mechanistic model here presented was compared to the empirical laws listed in Table 6.1. Tests C and H were selected for this comparison, as shown in Figure 6.10. In both situations, the empirical laws by Bomben and King [319] and Kochs *et al.* [321] gave similar results as the mechanistic model. On the contrary, the equations proposed by Kurz and Fischer [317] seemed to be less accurate for this particular application. The values of the C coefficient used in the empirical laws are reported in Table 6.6.

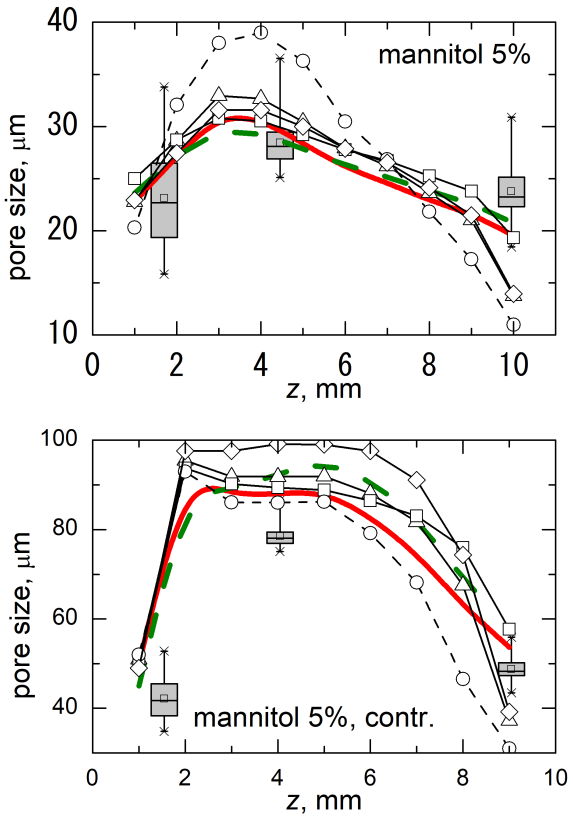

Figure 6.10: Comparison between detailed model predictions (solid red line), simplified model predictions (dashed green line) and empirical laws described in: Bomben and King [319] (—△—), Kochs *et al.* [321] (—□—), Kurz and Fischer-1 [317] (- -◯- - -) and Kurz and Fischer-2 [317] (—◇—) for tests C and H. SEM observations are reported in the box plots. Only the boxes are reported for sake of clarity

Table 6.6: Numerical values of parameter C employed

	C, $m^{1+\lambda_1-\lambda_2}\ s^{-\lambda_1}\ K^{\lambda_2}$	
	Test C	Test H
Bomben and King [319]	$1.6{\cdot}10^{-6}$	$7{\cdot}10^{-6}$
Kochs *et al.* [321]	$28{\cdot}10^{-6}$	$120{\cdot}10^{-6}$
Kurz and Fischer-1 [317]	$0.1{\cdot}10^{-6}$	$0.5{\cdot}10^{-6}$
Kurz and Fischer-2 [317]	$0.09{\cdot}10^{-6}$	$0.32{\cdot}10^{-6}$

6.3 The Freezing Step of Freeze Drying: Considerations for Process Efficiency

The crystal size D_p of the frozen product also equals the pore dimension of the dried cake, provided that no collapse occurs. In turn, the pore dimension determines the mass transfer resistance to vapor flow during primary drying, and, therefore, primary drying time t_d and maximum temperature reached during this phase T_{max}. As previously discussed in this Chapter, either the model introduced in Nakagawa *et al.* [320] and Pisano and Capozzi [323] or the iterative approach introduced in section 6.2 could be used to compute the average value of ice crystal size as function of freezing conditions, i.e., cooling rate and nucleation temperature. Afterwards, in order to assess the effect of the ice crystal size on t_d and T_{max}, a simple 1-dimensional model could be used [327], as detailed in Appendix C.

A design space approach could then be employed, as follows:

1. An appropriate range of cooling rates and nucleation temperatures is selected;
2. This range is discretized into a matrix of operating conditions to be explored;
3. The ice crystal size is computed for the selected operating conditions;
4. The mass transfer resistance to vapor flow within the dried product is computed, using as input the dimension of ice crystals (Equation C.11);
5. t_d and T_{max} are computed for the selected mass transfer resistance.

The design space for freezing could eventually be obtained, which may have several applications. For instance, it could be used to select the appropriate freezing conditions in order to obtain the desired pore dimension within the product. By adjusting the average ice crystal size, a remarkable speed up of primary drying could be achieved while avoiding any dangerous stresses for the product being dried. For example, freezing conditions that avoid product collapse can be chosen without excessively compromising process efficiency.

6.3.1 Experimental Approach

Materials and Instrumentation

Both crystalline, i.e., 5% w/w mannitol, and amorphous, i.e., 5% w/w sucrose or dextran formulations were chosen, so as to test the design space approach. Among the solutes under investigation, sucrose is particularly heat sensitive, and, using freeze-drying microscopy (microscope: BX51, Olympus Europa, Hamburg, Germany; temperature controller: PE95-T95, Linkam, Scientific Instruments, Tadworth, Surrey, UK), a collapse temperature of 241 K was observed. Batch preparation and monitoring of the nucleation temperature were performed as described

in paragraph 6.2.3. It was observed that, in our tests, the nucleation temperature varied less than 6.5 K between different vials.

Freeze-Drying Protocol

The experiments performed for testing the design space were carried out using shelf-ramped freezing, by decreasing the shelf temperature with a precise cooling rate until 228 K. The subsequent primary drying phase was performed at 10 Pa and 263 K, for mannitol and dextran, or 253 K, for sucrose, with a 1 K min^{-1} ramp from 228 K to the final primary drying temperature. The operating conditions of the experimental tests used for validation of the mathematical approach are listed in Table 6.7.

Table 6.7: Operating conditions of the experimental tests performed.*

		Freezing			Primary Drying	
	Formulation	Cooling Rate	Nucleation T		Fluid T	Pressure
		K min^{-1}	Mean, K	Range, K	K	Pa
1	Mannitol	0.1	264.5	5.0	263	10
2	5%	0.38	263.3	6.1		
3		0.71	261.9	5.8		
4		0.1	264.4	5.0		
5	Dextran	0.54	262.7	2.4	263	10
6	5%	0.71	264.6	3.6		
7	Sucrose	0.1	260	4.1	253	10
8	5%	0.7	258.2	3.5		

* Both the average nucleation temperature and its range of variation (difference between maximum and minimum value of observed nucleation temperature) are listed.

Since the thermocouples were placed at the vial bottom, the change in slope of measured temperature profiles indicated that the sublimation front had reached the bottom of the sample and that primary drying had finished. The primary drying time could also be estimated from a comparative pressure measurement, i.e., Pirani vs. capacitance manometer as discussed in section 5.3 [293]. In this way, the drying time could be quantified from experimental data.

Determination of Pore Size

The cakes obtained after drying seemed to show no defects. Neither shrinkage (visual inspection) nor micro-collapse (SEM images) were observed. Moreover, the maximum temperature measured by thermocouples was sufficiently below the maximum available product temperature to guarantee absence of micro-collapse. Thus,

the pore dimension of the products obtained after freeze-drying should reliably correspond to the dimension of the ice crystals formed during freezing, and was analysed using a Scanning Electron Microscope (SEM, FEI type, Quanta Inspect 200, Eindhoven, the Netherlands). The samples were cut along their vertical axis, and SEM images were taken at different points along this section. More specifically, three images were taken at the top, center and bottom of the product cake, and analyzed so as to estimate the pore size, D_p, of the product as previously described.

6.3.2 The Design Space Allows Prediction of Drying Time and Thermal Stress for the Product

The variation in ice crystal size as function of cooling rate and nucleation temperature was evaluated, using the mechanistic model introduced in this Chapter (Equation 6.15). The graphs obtained are shown in Figure 6.11, where the experimental points used for testing the design space are displayed as red bullets.

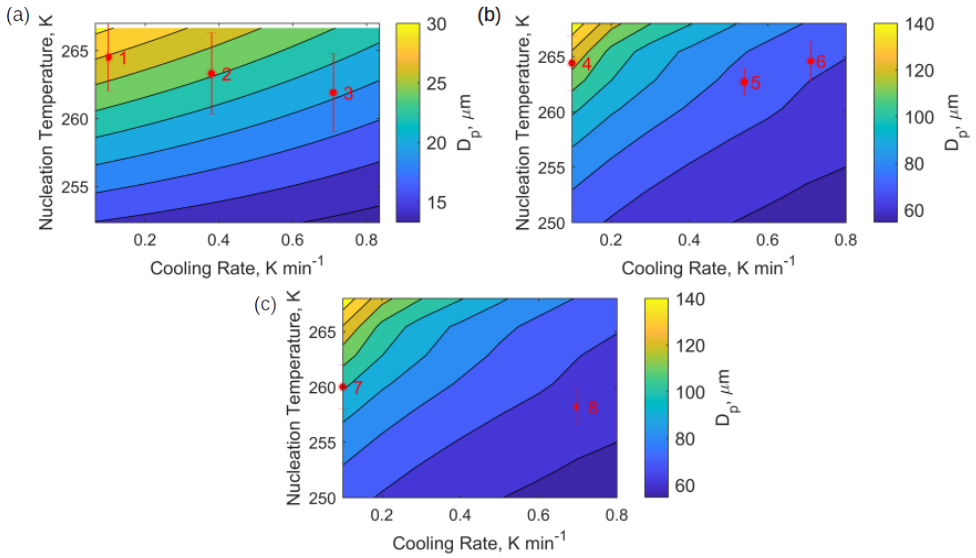

Figure 6.11: Design space for freezing showing the dimension of ice crystals as function of nucleation temperature and cooling rate, for mannitol (a), dextran (b) and sucrose (c). Points 1 to 8 refer to the experimental tests summarized in Table 6.7. The vertical error bars shown on the graphs correspond to the measured variation in nucleation temperature.

As can be easily observed, the model predicted that the mean ice crystal size increases with decreasing cooling rate and increasing nucleation temperature. This is in line with experimental results. For instance, while at 0.1 K min^{-1}, as in test 1, the mean ice crystal size was 30 μm, at 0.71 K min^{-1}, as in test 3, the mean

dimension of ice crystals was much lower, 18 μm. In Table 6.8, the comparison with experimental data is summarized.

Table 6.8: Comparison between ice crystal size as measured experimentally (Exp.) and as predicted by the design space (DS).

Test	D_p, μm Exp.	D_p, μm DS	Error, %
1	30	28	6.7
2	24	24	0
3	18	20	7.1
4	125	125	0
5	80	78	2.5
6	68	71	4.4
7	115	110	4.3
8	71	70	1.4

In the second column, the mean ice crystal size, as evaluated from SEM images for the eight experimental tests performed, is listed, while the following columns correspond to the design space predictions, and the percentage errors made when using the model. The design space seems to allow a fairly accurate estimation of the experimental value, with a relative error that was always smaller than 10%.

Let us consider some examples of application. For instance, values of D_p in the range 24-34 μm may be desired for a 5% w/w mannitol formulation, and it may be known that the nucleation temperature for that specific production will range between 262 and 267 K. This information could be obtained once a statistically significant number of tests has been performed, and the nucleation temperatures obtained have been measured and analyzed. If a cycle were performed at 0.7 K min^{-1} as cooling rate, a product with D_p in the range 18-24 μm would be obtained according to the design space, which does not correspond to the desired morphology. Experimental data confirmed this result; for test 3 in Table 6.7 the cooling rate was 0.71 K min^{-1}, the nucleation temperature was approximately 262 K and a value of 18 μm was observed for D_p. However, according to the design space, a cooling rate of 0.1 K min^{-1} would exactly lead to the desired range of D_p. In line with the design space, using 0.1 K min^{-1} as cooling rate and having 264.5 K as nucleation temperature a value of about 30 μm was experimentally measured for D_p in test 1.

The same considerations can be made for the case of an amorphous excipient, such as dextran. A product with D_p ranging between 110 and 130 μm may be desired. It may be known, as previously discussed, that the nucleation temperature will range between 262 and 267 K. If freezing were performed at 0.5 K min^{-1} as cooling rate, according to the design space, D_p would range between 78 and 99 μm, which does not satisfy the desired requirements. Experimental data confirmed this

result; for test 5 in Table 6.7 the cooling rate was 0.54 K min^{-1}, the nucleation temperature was approximately 262.7 K and a value of 80 μm was obtained for D_p. However, the design space indicates that 0.1 K min^{-1} as cooling rate would exactly lead to the desired range of D_p. Accordingly, using 0.1 K min^{-1} as cooling rate and having 264.4 K as nucleation temperature, as in test 4, $D_p = 125$ μm was experimentally observed.

The design space revealed that the influence of the nucleation temperature on the mean ice crystal size is lower at high cooling rates. For example, for mannitol as excipient and with 0.8 K min^{-1} as cooling rate, the mean crystal size varied from about 14 μm at a nucleation temperature of 252 K, to about 22 μm at a nucleation temperature of 267 K. Thus, at this cooling rate the crystal size changed by 57% in the selected range of nucleation temperatures. On the contrary, if a cooling rate of 0.1 K min^{-1} were selected, the mean ice crystal size would increase approximately by 100% in the same range of nucleation temperatures, from 17 to 34 μm.

The same observation is valid if the cooling rate is varied at equal nucleation temperature. For instance, at 266 K as nucleation temperature and changing the cooling rate from 0.8 to 0.1 K min^{-1}, the average crystal size of mannitol-based products would increase by 48%, from 23 to 34 μm. At lower nucleation temperature, i.e., 252 K, the same reduction in cooling rate would make the ice crystal size change by 31% only, approximately from 13 to 17 μm. This suggests that at high nucleation temperature the influence of cooling rate on the ice crystal size is more important.

The design space obtained for the freezing step of freeze-drying may allow a speed-up of process design, along with a better understanding of phenomena involved. At present, process operating conditions are chosen using a non-systematic approach. This means that operating conditions are adjusted without a real knowledge of the effect they will produce and the impact they will have on the process. By contrast, the design space for freezing would allow a selection of operating conditions driven by deep understanding of the process. For given cooling rate and nucleation temperature, it would be possible to know the pore size that will be obtained. By coupling this knowledge with a proper primary drying model, it would also be possible to quantify the impact of pore size on drying performances. The net result may be an improved control of the process.

From this viewpoint, two very important variables of the primary drying phase are the processing time t_d and the maximum temperature T_{max} reached within the product. As regards the drying time, it determines the profitability of the process and the objective is to reduce it as much as possible. However, there is another constraint which must be fulfilled. During freeze-drying, care must be taken to keep the product temperature below a limit value that is characteristic of the product being processed. The goal is to avoid product degradation and also the collapse, or shrinkage, of the dried cake. In order to do this it is necessary to keep the maximum temperature reached within the product, T_{max}, below the collapse temperature of

the formulation.

In this framework, in the present work two graphs were built, which describe the variation of t_d and T_{max} upon nucleation temperature and the cooling rate used during freezing, see Figure 6.12. More specifically, for modelling primary drying, the fluid temperature and chamber pressure listed in Table 6.7 were used.

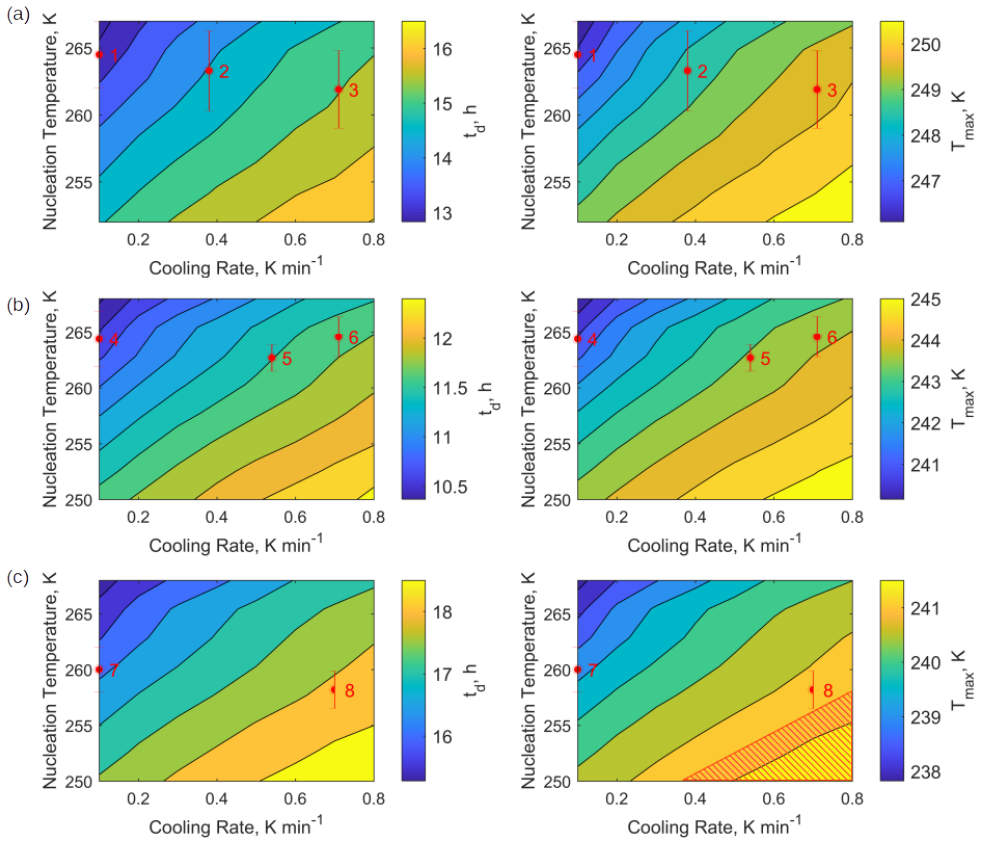

Figure 6.12: Design space for freezing showing the drying time t_d and maximum temperature T_{max} as function of nucleation temperature and cooling rate, for mannitol (a), dextran (b) and sucrose (c). Points 1 to 8 refer to the experimental tests summarized in Table 6.7. The vertical error bars shown on the graphs correspond to the measured variation in nucleation temperature. For modelling primary drying, the fluid temperature and chamber pressure listed in Table 6.7 were used.

These graphs allow fast estimation of drying performances for varying possible operating conditions during the freezing step. At high cooling rate and low nucleation temperature, the pores obtained within the dried cake are small. According to Equation C.11, this leads to a very high value of product resistance to vapor flow. The result is that the process becomes longer, and the maximum temperature

reached within the product is higher. Thus, small pores within the product are undesired for the primary drying phase of freeze drying. The design space obtained was also tested upon experimental data, and the results of this process are shown in Table 6.9.

Table 6.9: Comparison between mean value of drying time t_d and maximum product temperature T_{max} as measured experimentally (Exp.) and as predicted by the design space (DS).

Test	t_d, h Exp.	t_d, h DS	Error, %	T_{max}, K Exp.	T_{max}, K DS	Error, %
1	13.30	13.25	0.4	245.2	246.5	0.5
2	14.53	14.70	1.2	248.7	248.2	0.2
3	15.21	15.40	1.2	250.0	249.4	0.2
4	10.89	10.70	1.7	241.3	241.1	0.1
5	11.22	11.30	0.7	243.0	243.3	0.1
6	11.75	11.70	0.4	244.0	243.7	0.1
7	16.00	16.40	2.5	238.7	239.1	0.2
8	18.00	17.80	1.1	240.0	240.6	0.3

In the case of test 1 in Table 6.7, the freezing cycle was performed at 0.1 K min^{-1} cooling rate, and the nucleation temperature varied between 262 and 267 K. According to Figure 6.12a, it would be possible to predict that the drying time should range between 12.25 and 14.25 h. If a value for the nucleation temperature is then fixed, for example 264.5 K which is the centre of the observed range, the design space predicts an average drying time of 13.25 h. Experimentally, a drying time of 13.30 h (offset time) was observed. The drying time predicted by the model is thus in good agreement with experiments.

The same approach can be used for tests 2 to 8, and the results of such analysis are shown in Table 6.9, which suggests that the design space could allow accurate prediction of drying time t_d. This is a remarkable finding, as the duration of primary drying determines the costs and efficiency of the process.

On the contrary, the maximum temperature reached by the product during primary drying is related to the quality of the final product. A great advantage of the design space is that it can also be used to predict the maximum temperature. Let us consider sucrose, whose collapse temperature is 241 K, and let us also suppose that the nucleation temperature in the cycle being studied will be in the range 250 - 268 K. If a cooling rate of 0.7 K min^{-1} were chosen, according to Figure 6.12c the maximum temperature reached by the product during primary drying would range between 240 and 241.5 K. Thus, using the design space, it would be possible to conclude that the selected cooling rate is not suitable, as it may result in collapse of the dried cake. However, if nucleation occurred at higher temperature, for instance

in the range 256-268 K, then collapse would be avoided according to the design space. This result was experimentally confirmed; in the case of test 8, the cooling rate and nucleation temperature were 0.7 K min^{-1} and 258.2 K, respectively, and no collapse was observed. On the contrary, if a cooling rate of 0.1 K min^{-1} were selected, the design space would predict a maximum temperature between 238 and 240 K, i.e., below the collapse temperature. Thus, product collapse should be avoided in this case. The result obtained was again confirmed by experimental results; for test 7, the measured maximum temperatures were in the range 238.4 - 239.0 K, which is included in the interval predicted using the design space.

The design space shown in this work is thought to be coupled with knowledge of the thermal and mechanical behaviour of the system under investigation. For example, if the collapse temperature of the formulation under investigation is known, the design space could be used to fix suitable operating conditions to avoid collapse. In this perspective, the collapse temperature of each formulation could be taken into account as constraints to the calculated design space. In the case of sucrose, for instance, knowledge of the collapse temperature made it possible to distinguish the region of operating conditions which will likely lead to product collapse, shaded in red in Figure 6.12c, from the region of conditions that will not compromise the product.

The design space for freezing calculated in this work could also be coupled with the design space for primary drying, as calculated in Fissore *et al.* [328] and Pisano *et al.* [329]. For instance, in Figure 6.13, the design space for primary drying of a 5% w/w sucrose solution is shown. The graph shows the maximum allowable fluid temperature, as a function of chamber pressure, for an average pore size equal to 60 μm (dashed red line) or 140 μm (solid black line). These values were chosen because they correspond to the 2 extreme ice crystal sizes calculated in this work for sucrose (Figure 6.11).

From Figure 6.13, it is possible to notice that at chamber pressure of 10 Pa and fluid temperature of 263 K, which are the conditions employed in the present study, the sample with 60 μm pore size would be collapsed, whereas the 140 μm one would still be below its collapse temperature. This is confirmed by the design space for freezing, Figure 6.12c, where the 60 μm sample would be in the shaded area, whereas the 140 μm one would lie in the region of allowed operating conditions.

Again, Table 6.9 shows that the design space allowed accurate prediction of the maximum temperature T_{max} reached by the product. This suggests that, using the design space, the impact of selected operating conditions for freezing on drying performances may be quantified. This could allow not only an increased manufacturing efficiency, but also an improved quality of the product obtained.

Finally, it should also be mentioned that the same approach here outlined could be used to build a design space for the VISF controlled nucleation approach, as discussed in [270]. In this case the nucleation temperature T_n is not a stochastic variable any more, as its value can now be controlled. However, other parameters

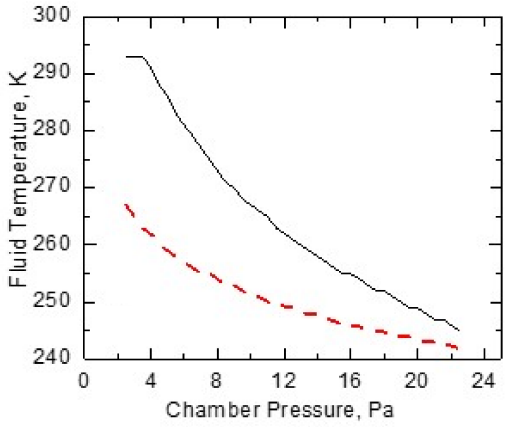

Figure 6.13: Design space for primary drying showing the maximum allowable fluid temperature in the case of sucrose as excipient (241 K collapse temperature), as function of the pressure within the drying chamber. The solid black curve and dashed red line were calculated for an average pore size of 140 and 60 μm, respectively.

come into play, such as the temperature value of the second holding step, T_m, or the holding time at this temperature. It may be interesting to know what happens if these values are modified.

During holding at T_m, not all the water freezes, because, according to the solute-water phase diagram, a maximum equilibrium value of solid content is eventually reached, which cannot be exceeded. Thus, the remaining water freezes later on, during the final ramp to, for instance, 228 K.

However, if the holding time at T_m is too short, the equilibrium value of solid content cannot be reached, and the freezing is mostly accomplished as in a conventional shelf-ramped freezing process.

Thus, it is evident that both T_m and the holding time at T_m have an impact on product morphology, and an example of design space, showing their effect on the average ice crystal size D_p, is shown in Figure 6.14.

The design space was calculated for sucrose 5% w/w, and cooling rate and nucleation temperature equal to 0.7 K min^{-1} and 268 K, respectively. Compared to the modelling approach previously outlined, a further condition must be considered in the case of VISF, i.e., the impossibility to overcome the maximum equilibrium solid content during the holding stage at T_m. Thus, when $T_{shelf} = T_m$, a constraint equation must be added to the model. In our case, this equation was obtained using data for the sucrose-water state diagram [330, 331].

As can be noticed, a reduction in T_m resulted in a reduction in ice crystal size, as well. As an example, with a 1 h holding time, the crystal size changed from

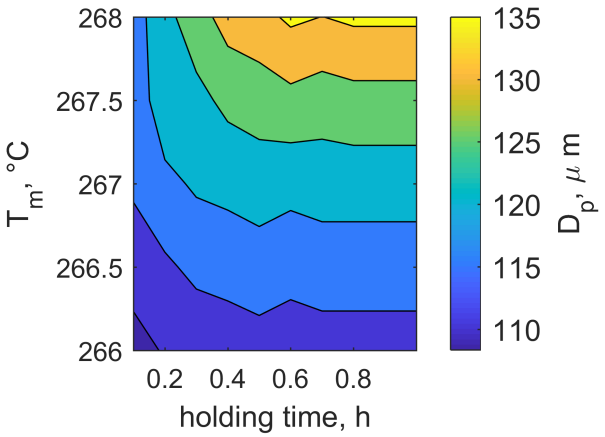

Figure 6.14: Design space showing the effects of T_m and holding time at T_m on the average ice crystal size D_p. Data refer to VISF of sucrose 5% w/w.

about 140 μm to about 110 μm, when T_m was decreased from 268 K to 266 K.

However, if the holding time was too short, and, thus, no sufficient time was left to reach equilibrium, the resulting ice crystal size was smaller than expected. For instance, if T_m was equal to 268 K, and the holding time was decreased from 50 min to 10 min only, then the average ice crystal size would decrease as well, from about 140 μm to about 115 μm. This happens because freezing cannot be completed during the holding time, and eventually occurs during the ramp to 228 K, as already discussed when tests G and H in Table 6.2 were compared.

Thus, it can be concluded that, in order to optimize the freezing cycle, the T_m should be high, but, in this case, a sufficiently long holding time should be used to ensure that equilibrium is reached before the beginning of the ramp to 228 K.

6.4 What Freezing Protocol Preserves Protein Stability the Best? Answers from the Design Space

The effect of cooling rate and nucleation temperature on protein stability will now be investigated using a Design Space approach, and considering a simple two-state unfolding process. Two opposite situations will be identified, based on the ratio between protein stability in the bulk and at the ice-water interface.

6.4.1 Simulation Approach

Freezing of a protein-based active ingredient was considered, and a two-state unfolding process of a protein, from the native (N) to denatured (U) state, was hypothesized. A two-state kinetic model is valid if the following assumptions are met: (i) a single folded state, and a large number of unfolded states can be identified; (ii) each unfolded state gives a small contribution to the total partition function of unfolded states; (iii) transition rate constants are inversely proportional to the partition functions of single conformational states, and each unfolded state can transition to several other unfolded states; (iv) the boundary factors in transition rate constants are not influenced by changes in folding conditions; (v) there are no large barriers between unfolded states which could result in the protein being trapped for a long time [332, 333].

In a two-state process, the mole fraction of unfolded protein, at thermodynamic equilibrium, is given by,

$$f_U = \frac{[\mathrm{U}]}{[\mathrm{N}] + [\mathrm{U}]} = \frac{K}{1+K} \tag{6.28}$$

where the equilibrium constant K depends on the free energy of unfolding ΔG,

$$K = e^{-\Delta G/\mathrm{R}T} \tag{6.29}$$

R is the universal gas constant, and T the absolute temperature. The free energy change depends on temperature and concentration c of an excipient according to,

$$\Delta G(T,c) = \Delta H_0 - T\Delta S_0 + \Delta c_p[T - T_0 - T\ln T/T_0] + mc \tag{6.30}$$

where ΔH_0, ΔS_0 and Δc_p are the enthalpy, entropy and specific heat change upon unfolding at a reference temperature T_0. The proportional coefficient m depends on the particular excipient considered. In most cases, thermodynamic equilibrium is not reached during the freezing process, and a kinetic approach should be used. For a reversible two-step process, it is possible to write the following system of differential equations,

$$\begin{aligned} \tfrac{d[\mathrm{U}]}{dt} &= -k_f[\mathrm{U}] + k_u[\mathrm{N}] \\ \tfrac{d[\mathrm{N}]}{dt} &= +k_f[\mathrm{U}] - k_u[\mathrm{N}] \end{aligned} \tag{6.31}$$

where the kinetic constants k_f and k_u depend on temperature, osmolyte concentration and solution viscosity μ [234, 235]. Moreover, we can assume the unfolding process as an activated process, characterized by the presence of a transition state TS,

$$\begin{aligned} k_u(T,c,\mu) &= A\tfrac{\mu_0}{\mu}e^{-\Delta G^{\text{N-TS}}/\mathrm{R}T} \\ k_f(T,c,\mu) &= A\tfrac{\mu_0}{\mu}e^{-\Delta G^{\text{U-TS}}/\mathrm{R}T} \end{aligned} \tag{6.32}$$

A is a constant for any protein, and μ_0 the viscosity at some reference conditions (293 K, 1 bar and pure water in our case, i.e., $\mu_0 = 1$ cP). In practice, we describe the conformational dynamics of proteins in viscous solutions using a version of the Kramers' theory [334–336] where the contribution given by the internal friction of the protein has been considered negligible [235]. This assumption is particularly justified in the case of a freezing process, where the solvent friction should dominate because of the large values of viscosity that are eventually reached.

The free energy change associated with the passage from native or denatured state, and the transition state (TS), is given by,

$$\Delta G^{\text{X-TS}}(T,c) = \Delta H_0^{\text{X-TS}} - T\Delta S_0^{\text{X-TS}} + \Delta c_p^{\text{X-TS}}[T - T_o - T\ln T/T_0] + m^{\text{X-TS}}c \quad (6.33)$$

with X=N, U. In a two-state process, we also have that $\Delta G = \Delta G^{\text{N-TS}} - \Delta G^{\text{U-TS}}$.

Here, realistic values were chosen for all the parameters related to protein stability, but with no reference to any specific protein. These values were also varied among different simulations, to cover an as wide as possible range of protein folding stabilities. The results presented in the following aim, therefore, to have general validity, qualitatively describing a variety of situations that may occur during the freezing of a protein formulation.

A typical pharmaceutical process, where a protein formulation is filled into vials and loaded onto the shelves of a freeze dryer, was considered during the simulations. More specifically, freezing of a 5% w/w sucrose formulation with a 150:1 sucrose to protein mole ratio was considered. During the simulations, the shelf temperature was linearly decreased from 293 to 233 K, with varying cooling rates over the range 0.1-1 K min^{-1}. The nucleation temperature was also varied in the range 248-265 K, and an 8x8 matrix of cooling rate x nucleation temperature conditions was considered to build the design space. The average ice crystal size D_p within the product was computed for each point of this matrix using the mechanistic model previously described, and the resulting ice-water surface area S_{ice} was computed from,

$$S_{ice} = \frac{4m_w}{D_p \varrho_{ice}} \quad (6.34)$$

where ϱ_{ice} is the ice density, m_w the initial mass of water in the vial, and the assumption was made that ice crystals are cylinder-shaped.

The increase in sucrose concentration as a result of the freezing process was computed assuming equilibrium conditions, using the data reported in [331]. The solution viscosity as function of temperature and excipient concentration was then calculated using the scaled Arrhenius equation,

$$\log_{10}\frac{\mu}{\mu_0} = c_0 + c_1\left(\frac{T_g}{T}\right) + c_2\left(\frac{T_g}{T}\right)^2 \quad (6.35)$$

where T_g is the glass transition temperature, and the coefficients c_0, c_1 and c_2 were taken from [134].

For the estimation of ΔG, values of 0.2 kJ kmol^{-1} K^{-1}, 3 kJ kmol^{-1} K^{-1} and 295 K were considered for ΔS_0, Δc_p and T_0, respectively. These values were chosen because are in the range often encountered for proteins. By contrast, the values of ΔH_0 and m were varied from simulation to simulation, to sample different situations of protein bulk stability. The kinetic constants k_u and k_f were then computed assuming $\Delta G^{\text{N-TS}} = 1.5\Delta G$ and $\Delta G^{\text{U-TS}} = 0.5\Delta G$, while the constant A in Equation 6.32 was again changed from simulation to simulation. Equations 6.31 were then solved in Matlab R2017a, using a 0.1 s time-step. The percentage of unfolded protein as a result of bulk denaturation could therefore be computed.

However, surface-induced denaturation may also occur, as a result of adsorption to the ice-water interface. From equation 6.34 it was possible to calculate the final extension $S_{ice}(t_{end})$ of the ice-water surface interface, while its change during the process was estimated from,

$$S_{ice}(t) = S_{ice}(t_{end})\frac{c_{w,end}}{c_w(t)} \tag{6.36}$$

where $c_{w,end}$ is the water mass fraction at the end of freezing, while c_w represents its current value. The percentage of surface-denatured molecules was finally computed assuming that the protein behavior was perturbed whenever it was closer than $d_i = 4$ nm to the ice surface. This threshold was chosen because a long-ranged effect of the ice surface, up to 4 nm from the interface, was observed in Chapter 3 for the GB1 peptide, and confirmed in Chapter 4 for protein L. To determine the number n_p of adsorbed protein molecules (having molar mass M_p and molecular volume V_p), the following system of equations was solved,

$$\begin{aligned} n_w M_w &= c_w(n_w M_w + n_p M_p + n_s M_s) \\ S_{ice} d_i &= V_w n_w + V_p n_p + V_s n_s \end{aligned} \tag{6.37}$$

where n_s and n_w are the number of adsorbed sucrose and water molecules, respectively, while M_w, M_s and V_w, V_s are their molar masses and molecular volumes. It was assumed that the ratio n_s/n_p was the same as in the bulk, i.e., 150. It was also assumed that the surface-driven denaturation of the protein occurred with no thermodynamic barrier. This means that 50% of the adsorbed protein molecules are in the unfolded state. For the simulations, both the completely reversible situation, and the completely irreversible one, were considered. In the latter case, Equations 6.31 were modified to,

$$\begin{aligned} \tfrac{d[\text{U}]}{dt} &= +k_u[\text{N}] \\ \tfrac{d[\text{N}]}{dt} &= -k_u[\text{N}] \end{aligned} \tag{6.38}$$

The irreversible situation corresponds, for instance, to the case of proteins that tend to form aggregates when they are in the unfolded state. In this case, the unfolded molecules that cluster to form an aggregate and precipitate cannot convert back to the native state anymore. The parameters used for the simulations performed in this work are listed in Table 6.10. The simulated conditions were chosen so as to explore a wide range of protein bulk stability, from the point of view of both thermodynamics (different ΔH_0, see Figure 6.15, and different m), and kinetics (different pre-exponential factor A), while keeping fixed the behavior at the ice surface. Simulations 1 and 3 correspond to proteins that are significantly more stable in bulk solution than at the surface. By contrast, the model proteins described in simulations 2 and 4 are unstable in bulk, and tend to unfold quickly at the low temperature experienced during freezing.

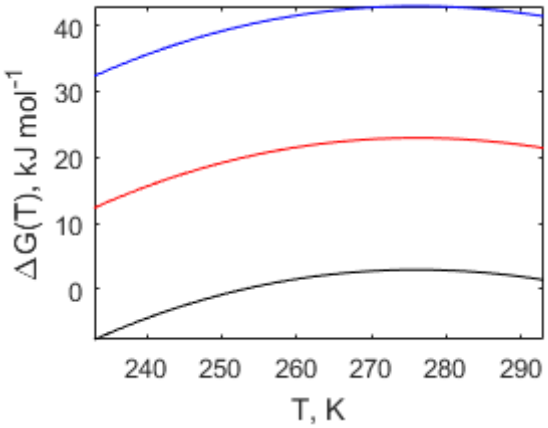

Figure 6.15: Free energy change ΔG upon unfolding, calculated for ΔH_0 equal to 100 kJ mol^{-1} (blue curve), 80 kJ mol^{-1} (red curve) or 60 kJ mol^{-1} (black curve).

Table 6.10: Details of the simulations performed.

	ΔH_0, kJ mol^{-1}	m, kJ mol^{-1} M^{-1}	A, s^{-1}	Reversible
1	80	4	$2{\cdot}10^2$	yes
2	60	1	$2{\cdot}10^2$	yes
3	100	4	$2{\cdot}10^1$	no
4	80	4	$2{\cdot}10^2$	no

6.4.2 Experimental Validation

Some experimental tests were performed to validate the simulation results. Myoglobin (Mb from equine heart, Sigma Aldrich, Milan, Italy) and lactate dehydrogenase (LDH from rabbit muscle, Sigma Aldrich, Milan, Italy) were selected as model proteins.

Freeze-Thaw Cycles on Myoglobin

Myoglobin was dissolved in 10 mM sodium citrate buffer at pH 3.7. Citrate buffer was selected because its pH remains constant during freeze-thawing [69]. Mb concentration was adjusted to either 0.1 or 0.2 mg/ml, and the presence of the surfactant Tween 80 (Sigma Aldrich, Milan, Italy) at 0.01% w/v was also considered for this study. All the solutions were prepared using water for injection (Fresenius Kabi, Verona, Italy), and filtered using 0.2 μm filters.

To assess the effect of different ice-water surface areas on protein stability, three different freezing protocols were used. In the first one, samples were frozen by immersing vials (0.5 ml Screw Cap GeNunc Storage Vials, HDPE, Sterile, Thermo Fisher Scientific, Rochester, NY, USA) into liquid nitrogen for 5 min, and then thawed in air at room temperature. In the second and third protocol, 2 ml of each sample were filled into 4R 16x45 mm vials (Nuova Ompi glass division, Stevanato Group, Piombino Dese, Italy), partially stoppered with silicon stoppers (West Pharmaceutical Services, Milan, Italy) and loaded onto the shelves of a freeze-dryer. In one case, shelf-ramped freezing was performed at 1 K/min, from 283 to 238 K using a Revo (Millrock Technology, Kingston, NY, USA) freeze-dryer. In the other case, the VISF technique was used. This cycle was performed in a LyoBeta 25 (Telstar, Terrassa, Spain) freeze-dryer, where the vials were first equilibrated at 268 K for 1 h. At that point, pressure inside the chamber was lowered to about 1 mbar to induce nucleation. Pressure was subsequently released to the atmospheric value, and the product equilibrated at 268 K for 1 h to promote the formation of big ice crystals. Finally, temperature was decreased to 238 K at 0.5 K/min. For both cycles, the cooling rate was also monitored by means of T-type miniature thermocouples placed inside some non-active (containing only the buffer) vials. In all cases, samples were thawed in air at room temperature.

In the case of quench cooling in liquid nitrogen, a large ice-water surface area should be formed, while the biggest ice crystals should be obtained using the VISF protocol, where the high nucleation temperature (268 K) and the holding time at 268 K should promote the ice crystals growth. In the case of quench and shelf-ramped freezing, three freeze-thaw cycles were performed, and samples were analyzed both after the first and the third cycle. By contrast, only one freeze-thaw cycle was performed using the VISF technique. After thawing, the protein solutions were centrifuged at 13000 rpm for 5 min (Heraeus Megafuge 8 Centrifuge Series, Thermo Fisher Scientific, Milano, Italy), and the percentage of aggregates was

thereafter calculated from the decrease in protein concentration (UV detection at 280 and 410 nm) after centrifugation. Optical density (OD) was measured against a solvent-matched reference using a 6850 UV/VIS spectrophotometer (Jenway, Stone, Staffordshire, UK).

The peak centered at 280 nm in the Mb absorbance spectrum is due to aromatic amino acids (primarily tryptophan and tyrosine) in the polypeptide and to the heme iron, and its location is characteristic of most proteins. The other absorption peak centered at approximately 410 nm is due entirely to the heme, and is commonly referred to as the Soret band [337].

The presence of stabilizing additives was also considered, and a freeze-thaw cycle was therefore performed on 0.1 mg/ml myoglobin in 5% w/w sucrose, 5% w/w trehalose, or 5% w/w mannitol, again using the Revo freeze-dryer. All the formulations were prepared in 10 mM sodium citrate buffer at pH 3.7, and a 0.1 mg/ml myoglobin solution in buffer only (without cryo- or lyoprotectants) was also considered. The formulations were prepared both with and without 0.01% w/w Tween 80, to investigate the effect of surfactants. Freezing was performed using a 1 K/min cooling rate, and the protein recovery was measured by UV/VIS spectroscopy at 410 nm after thawing in air at room temperature, using a spectrophotometric multiwell plate reader (Multiskan FC Microplate Photometer, Thermo Fisher Scientific, Milano, Italy).

Freeze Thaw Cycles on Lactate Dehydrogenase

LDH was dialyzed against 10 mM sodium citrate buffer at pH 6.5. Dialysis was performed at 277 K, and the buffer was changed 3 times (the first 2 times every 3 h, while the third dialysis step was carried out overnight). The concentration of LDH after dialysis was determined using UV/VIS spectroscopy. The peak at 280 nm was monitored, and an extinction coefficient of 1.44 mL/(mg cm) was used for calculations.

Two different freeze-thaw protocols were tested on 0.1 mg/ml LDH in 10 mM sodium citrate buffer at pH 6.5, 5% w/w sucrose, 5% w/w trehalose, or 5% w/w mannitol. All the formulations were prepared in 10 mM sodium citrate buffer at pH 6.5, both with and without 0.01% w/w Tween 80, and filtered using 0.2 μm filters. The first protocol was quenching in liquid nitrogen, while the second one consisted in shelf-ramped freezing at 1 K/min cooling rate. The same approach previously described for Mb was used for both protocols. After thawing in air at room temperature, the enzymatic activity of LDH after 1 freeze-thaw cycle was calculated from the increase in absorbance at 450 nm due to the reduction of NAD to NADH. A standard curve built with 1.25 mM NADH standard was used to calculate the amount of NADH generated in each well. For this analysis, 0.2 mg/ml LDH in 10 mM sodium citrate buffer at pH 6.5, with or without 0.01% w/v Tween 80, was also considered. The degree of aggregation in each sample was

also measured from the increase in absorbance at 500 nm, which is a measure of turbidity. However, this last analysis was performed after 9 freeze-thaw cycles, so as to increase the strength of the signal. In all cases, a spectrophotometric multiwell plate reader was used for the analyses.

Low-Temperature Storage

Mb and LDH were selected as model proteins because they are characterized by very different cold denaturation temperatures (283 K for Mb at pH 3.7 according to [338] and 245 K for LDH at neutral pH [339]). To further confirm their different behavior at low temperature, 0.1 mg/ml Mb in 10 mM sodium citrate buffer at pH 3.7 and 0.1 mg/ml LDH in 10 mM sodium citrate buffer at pH 6.5 were kept for 24 h at 268 K, where no freezing was observed. The presence of 0.01 % w/v Tween 80 was also considered. Protein stability over this time was monitored from the decrease in absorbance at 410 nm, in the case of Mb, or the ability to reduce NAD to NADH, in the case of LDH, as previously described.

6.4.3 An Insight Into Protein Behavior during Freezing

A typical output of a simulation is shown in Figure 6.16.
Before the nucleation event, the solution viscosity μ is too low to significantly hinder the conformational changes of the protein, and the kinetic constants show moderately high values. Immediately after the nucleation event, i.e., after approximately 0.7 h in the case of Figure 6.16, the excipient concentration c increases sharply because of the formation of ice crystals (Figure 6.16a), and this augments the bulk stability of the protein. This is evident from Figure 6.16b, where the free energy change ΔG is shown both including (red curve) or neglecting (black curve) the effect of the osmolyte. In this case, the effect of the excipient is stabilizing (positive m-value), as it shifts the free energy of unfolding to larger values. While the two curves are almost superimposed before nucleation, they clearly split apart as soon as the first ice crystals are formed. The solution viscosity increases as well (Figure 6.16a), kinetically hindering any protein movement. As a consequence, the kinetic constants k_u and k_f drop to very low values (Figure 6.16d). Surface-driven denaturation becomes therefore dominant after nucleation, and a not negligible amount of protein molecules adsorbs to the ice surface (Figure 6.16c), potentially undergoing conformational changes. The resulting fraction of unfolded (U) and native (N) protein molecules as function of freezing time is shown in Figure 6.17.

In the case of Figure 6.17a, simulation 1 in Table 6.10 is considered. In this case, the protein is stable in bulk, as the folding rate constant (k_f) is remarkably larger than the unfolding (k_u) one (see Figure 6.16d). Therefore, the protein molecules that unfold almost immediately convert back to the native state, and no notable unfolding occurs before nucleation. Once nucleation occurs, surface-driven

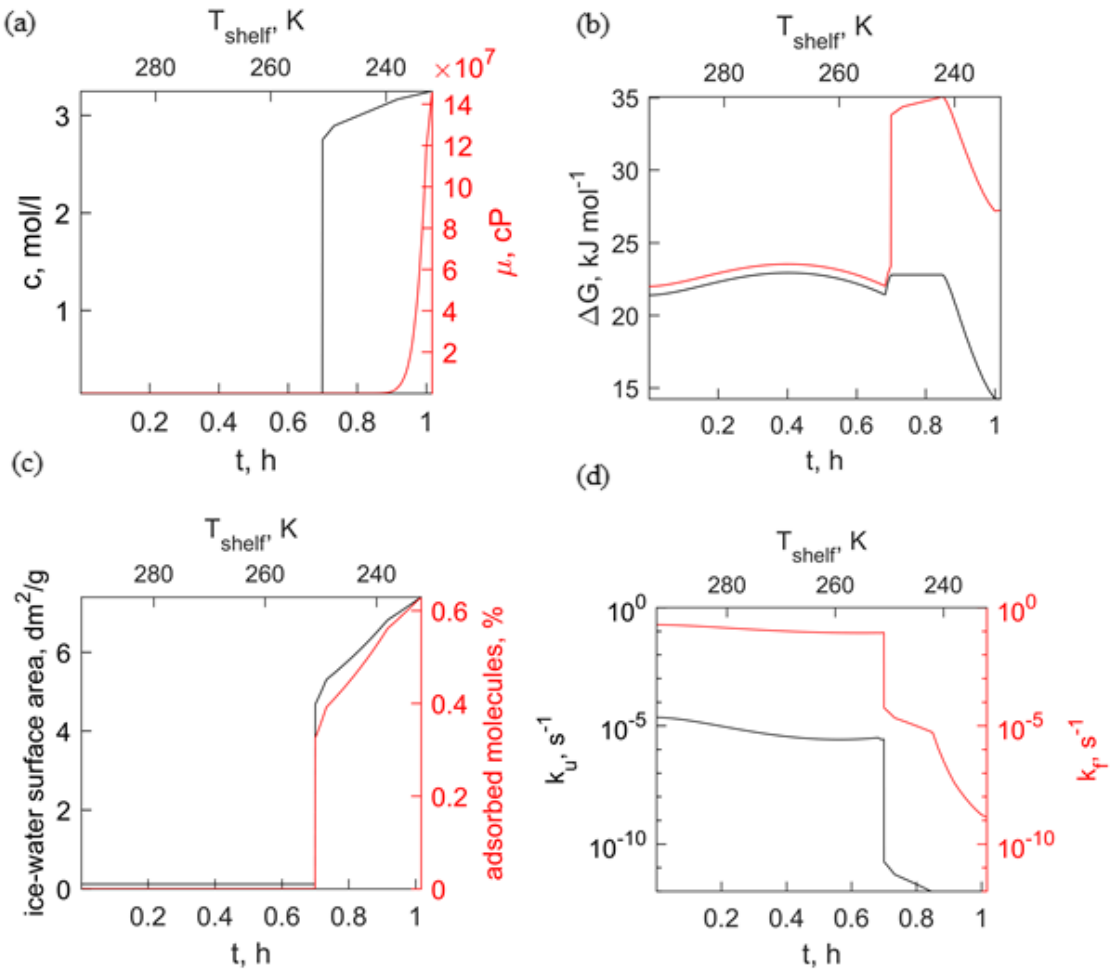

Figure 6.16: Example of model prediction of protein behavior during freezing at 1 K min^{-1} as cooling rate and 260 K as nucleation temperature. Evolution of (a) sucrose concentration c and solution viscosity μ, (b) free energy change of unfolding ΔG including (red curve) or neglecting (black curve) the effect of the osmolyte, (c) extension of the ice-water surface area (in dm^2/g, using the solvent mass as reference) and percentage of adsorbed molecules, (d) rate constants k_u and k_f as function of the freezing time, for the conditions of simulation 1 in Table 6.10. The evolution of the shelf temperature T_{shelf} is shown on the upper x-axis.

denaturation prevails and, as observed in Figure 6.17a, the amount of unfolded proteins suddenly increases. On the other hand, if the same values of ΔH_0, m and A are used, but an irreversible process (simulation 4 in Table 6.10) is considered (Figure 6.17b), a significant percentage of the protein undergoes conformational changes before nucleation, and the denaturation that is subsequently induced by

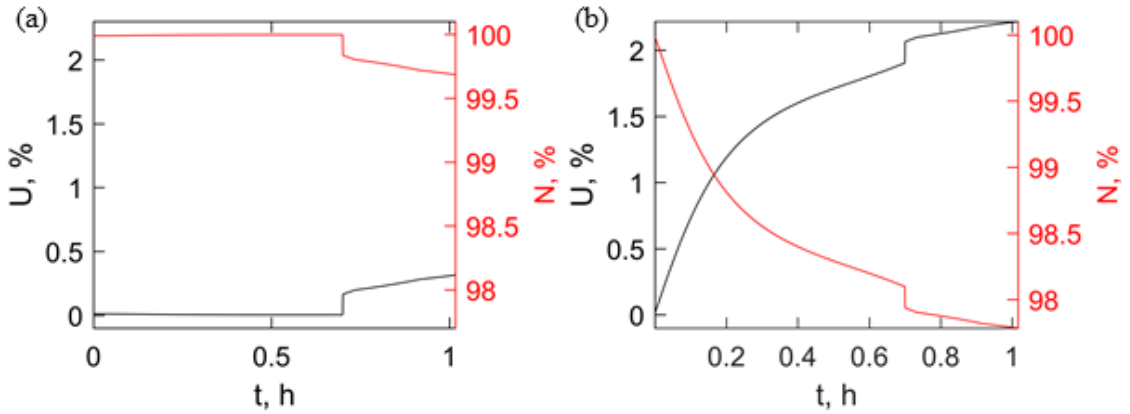

Figure 6.17: Percentage of unfolded (U) and native (N) protein molecules as function of the freezing time, in the case of (a) simulation 1 and (b) simulation 4 in Table 6.10. Cooling rate and nucleation temperature were 1 K min^{-1} and 260 K, respectively.

the ice surface plays only a secondary role. As evident from this first example, an important role is therefore played by the ratio between bulk and surface stability. A design space approach was therefore used to further investigate the implications of this observation.

6.4.4 The Design Space Suggests the Existence of Two Opposite Behaviors

The objective here is to identify the parameters that may be responsible for an increased loss of therapeutic activity during the freezing step of freeze drying, and should therefore be monitored during the process. As we are not interested in the absolute value of unfolded protein molecules in each configuration, the graphs will show the percentage of unfolded molecules (U), normalized by the maximum value observed in each design space (U_{max}). This will allow an easier comparison between different simulations.

In line with previous considerations, the design space for simulation 1 (Figure 6.18a) shows that the worst condition for protein stability ($U/U_{max} = 100\%$) corresponds to the region of high cooling rates and low nucleation temperature. As previously discussed, small ice crystals are formed in these conditions, that result in a large ice-water surface, as shown by the red isocurves in Figure 6.18. It is evident that the greater the extension of the ice interface is, the more the protein unfolds. In these conditions surface-induced denaturation prevails, and the optimal process should maximize the ice crystal size. This may be achieved using a low cooling rate, or inducing nucleation at high temperature by means of a controlled

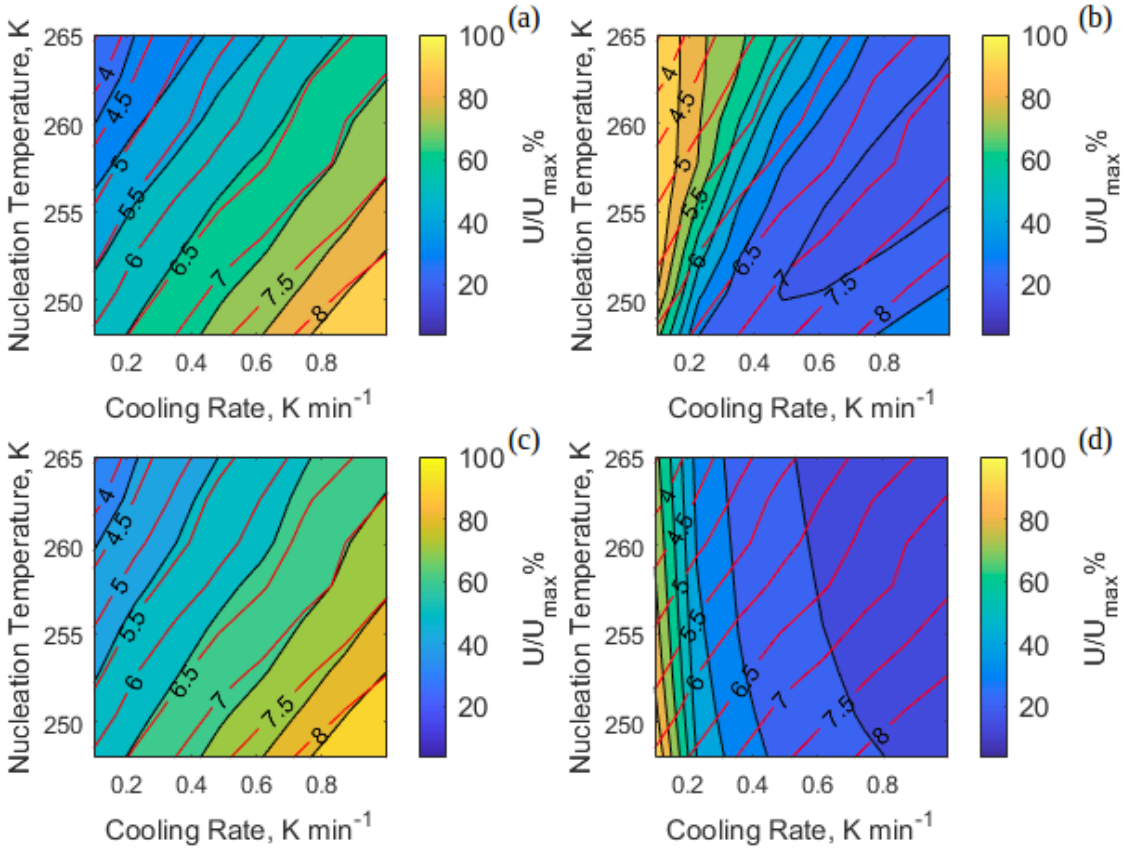

Figure 6.18: Design space showing the stress for the protein (percentage of unfolded protein, normalized by its maximum value in the design space), for simulations 1 (a), 2 (b), 3 (c) and 4 (d) in Table 6.10. The red isocurves identify the conditions resulting in the same total extension of the ice-water surface area, reported in dm^2/g (solvent mass is used as reference) on the curves.

nucleation approach, as the VISF protocol. Alternatively, the use of surfactants should be considered to minimize protein-surface interactions.

By contrast, if the protein has a reduced stability in bulk, as in simulation 2, surface-induced denaturation becomes less important. In this case (Figure 6.18b), a low cooling rate results in the highest degree of protein unfolding. This happens because the lower the cooling rate is, the longer the solution viscosity is low enough to allow fast conformational changes. Therefore, if a protein with low bulk stability is considered, a high cooling rate may be beneficial, as it would result in shorter freezing times and earlier cryo-concentration.

Similar considerations apply to the case of irreversible conformational changes.

In the case of simulation 3 (Figure 6.18c), surface-driven unfolding is again dominant and slow cooling rate/high nucleation temperature conditions should be preferred. On the other hand, if the denaturation process in the bulk solution is thermodynamically and kinetically favored, as in simulation 4 (Figure 6.18d), the extension of the ice interface is not a crucial parameter, while the duration of the freezing process should be strictly controlled, and a cryo-concentrated matrix should be formed as quickly as possible. Overall, Figure 6.18 shows that surface-induced unfolding is dominant in simulations 1 and 3 in Table 6.10, while it becomes less important in the case of simulations 2 and 4.

Finally, the effect of protein concentration has been investigated in Figure 6.19. In this case, a fixed value of 263 K was chosen for the nucleation temperature,

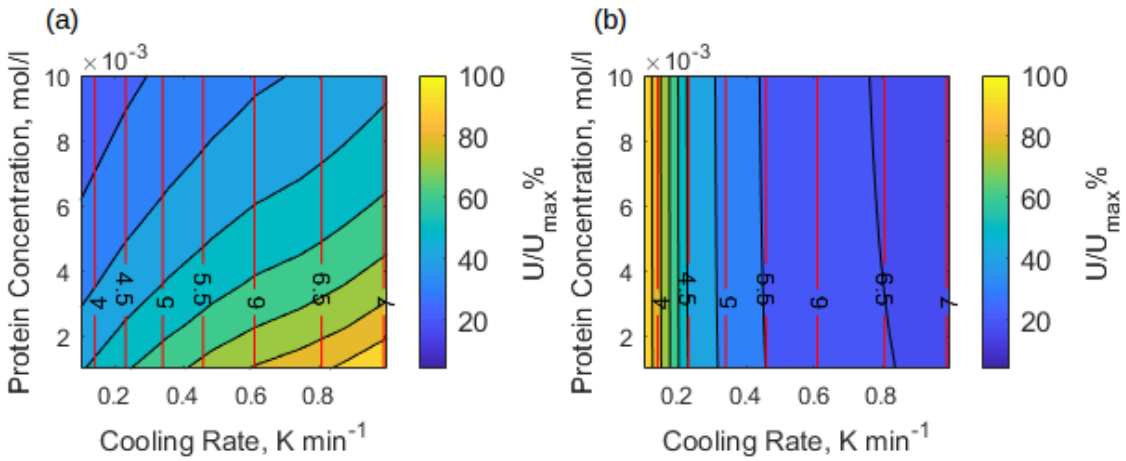

Figure 6.19: Design space showing the stress for the protein (percentage of unfolded protein, normalized by its maximum value in the design space), for simulations 3 (a), and 4 (b) in Table 6.10. The effect of protein concentration in the case of dominant surface-driven denaturation (a) or controlling bulk unfolding (b) is illustrated. The red isocurves identify the conditions resulting in the same total extension of the ice-water surface area, reported in dm^2/g (solvent mass is used as reference) on the curves.

and the cooling rate was again varied from 0.1 to 1 K/min. Conditions 3 and 4 in Table 6.10 were considered, and the protein concentration was varied between 0.001 and 0.01 mol/l. In the case of proteins having high bulk stability (conditions 3 in Table 6.10, Figure 6.19a), surface-driven denaturation is dominant, and increasing the protein concentration has a significant positive effect. This occurs because the extension of the ice-water interface is finite, and the number of protein molecules adsorbed at the surface cannot exceed a given value. When this value is reached, increasing the bulk concentration reduces the percentage of unfolded molecules. By contrast, the effect of protein concentration is not equally important when

cold denaturation in the bulk is the dominant mechanism (conditions 4 in Table 6.10, Figure 6.19b). In this case, only a very slight increase in protein recovery is observed when moving from 0.001 to 0.01 mol/l, and is again related to the smaller percentage of protein molecules adsorbed at the ice surface at higher concentration.

It should be pointed out that the contour plots shown in Figures 6.18 and 6.19 were computed for reference values of protein thermodynamics and kinetics parameters (ΔH_0, ΔS_0, Δc_p, m, A) and the relative trends only were here discussed. Considering that this is the first attempt to include protein stability in a design space approach, this was done with the aim to make the discussion as general as possible. It would anyway be possible to use this same modeling approach to obtain specific information for the protein being considered, by specifically adjusting the folding parameters included in the model.

6.4.5 Comparison with Experimental Data

For experiments, different freezing protocols were selected, that correspond to either the far left (slow process) or far right (faster process) of Figure 6.18. A very fast freezing protocol, i.e., quench cooling in liquid nitrogen, was also used as an extreme representative of a fast freezing process, with the objective to enhance the observed effects on protein stability. The behavior of proteins during freezing is extremely complicated, and several factors should in principle be taken into account. However, the simple model presented in this work made it possible to identify two possible scenarios. The first one, that primarily ascribes protein denaturation to adsorption at the ice-water interface, was confirmed in the case of LDH as model protein (see Figure 6.20). In panel (a), the recovery of protein enzymatic activity after 1 freeze-thaw cycle is shown, as measured from the increase in OD at 450 nm due to the reduction of NAD to NADH. In panel (b), the increase in OD at 500 nm due to aggregation after 9 freeze-thaw cycles has been considered to estimate protein recovery.

Both quench freezing and shelf-ramped freezing at 1 K/min were considered. Plastic vials were used for the quench freezing protocol for their better compatibility with the available equipment. However, preliminary tests for 0.1 mg/ml LDH in 10 mM sodium citrate buffer at pH 6.5 or 0.1 mg/ml myoglobin in 10 mM sodium citrate buffer at pH 3.7, with or without 0.01% w/v Tween 80, were used to verify that no significant difference existed with the results obtained in 4R 16x45 mm glass vials. It is evident that the quench freezing in liquid nitrogen (blue bars) resulted in reduced protein recovery compared to the shelf-ramped freezing at 1 K min^{-1} (grey bars). The formation of extremely small ice crystals, and therefore of a huge ice-water surface area, during quench freezing may be at the basis of this observation. In all cases, the addition of 0.01 % w/v Tween 80 remarkably reduced protein aggregation, as evident from Figure 6.20b, and often resulted also

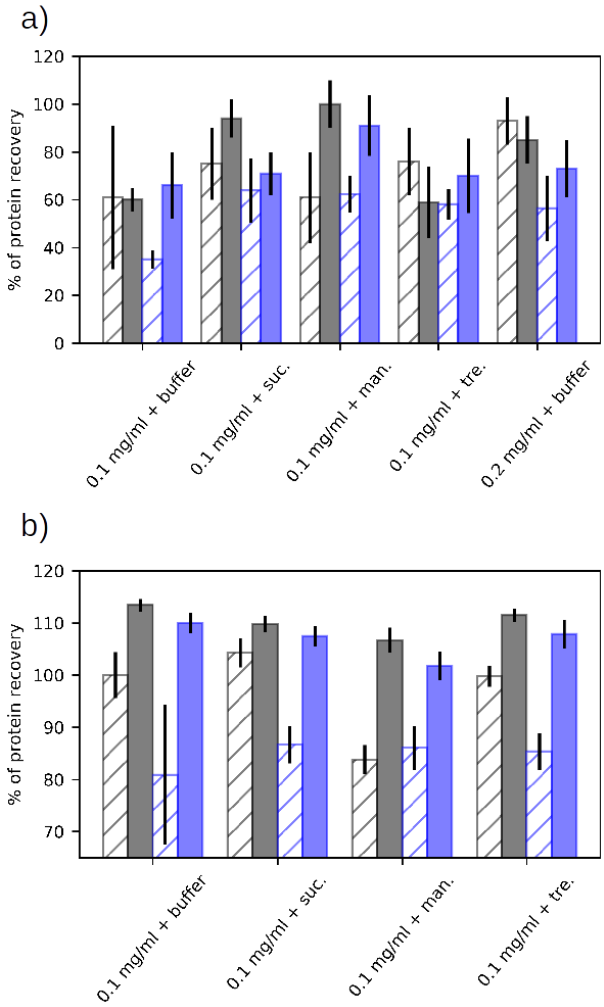

Figure 6.20: Percentage of LDH recovery after shelf-ramped freezing at 1 K min^{-1} (grey bars) or quench freezing (blue bars), both in presence (plain bars) and absence (dashed bars) of Tween 80, as measured from (a) the increase in absorbance at 450 nm due to the enzymatic reduction of NAD to NADH or (b) the increase in OD at 500 nm. For panel (a) data have been normalized relative to the formulation before freeze-thaw. For panel (b) data have been normalized relative to the buffer only formulation after shelf-ramped freezing (in absence of surfactant).

in increased protein activity, as shown in Figure 6.20a. More specifically, the addition of surfactant always led to improved recovery of LDH activity when quench freezing was performed (blue bars in Figure 6.20a). This further suggests that surface-driven denaturation was dominant, especially for the case of the quench freezing protocol. Sucrose and trehalose at 5% w/w resulted in improved protein

activity (Figure 6.20a), even though they did not remarkably decrease protein aggregation at 1 K/min cooling rate (Figure 6.20b). By contrast, the addition of the crystalline excipient mannitol, in absence of surfactants, was deleterious for protein aggregation at this cooling rate, as it resulted in increased turbidity of the solution after freeze-thawing. This increase in aggregation may be related to the formation of mannitol crystals at 1 K/min, where the protein may adsorb and denature. This explanation is also supported by the fact that when Tween 80 was added to the mannitol formulation, the protein recovery significantly increased, as regards both enzymatic activity (Figure 6.20a) and aggregation (Figure 6.20b). However, the quench frozen samples containing cryoprotectants, and in absence of surfactants, all showed an increase in protein recovery compared to the buffer only formulation, and behaved very similarly, suggesting that mannitol did not crystallize under these conditions, in line with previous observations [340, 341]. This effect of cryoprotectants and surfactants was not explicitly discussed in the simulation part of this work, but could be included in the model by modifying the value of m in Equations 6.30 and 6.33, or the adsorption behavior at the ice interface (in the case of surfactants, Equations 6.37).

Finally, the recovery of LDH enzymatic activity was also tested for a different concentration, 0.2 mg/ml in Figure 6.20a. For both shelf-ramped freezing and quench freezing, protein recovery improved when moving from 0.1 to 0.2 mg/ml. As shown in Figure 6.19a, this beneficial effect observed upon increasing the protein concentration is not surprising when surface-driven denaturation is dominant.

The behavior discussed in this work for LDH is also well-documented in the literature. For instance, using LDH as model protein, a remarkable loss of activity was observed in frozen systems, while no degradation was detected in concentrated solutions at the same temperature and composition, but without ice [61]. Similarly, solutions of the azurin protein exhibited a dramatic decrease in the average phosphorescence lifetime of the Trp-48 residue at the onset of ice formation, which is indicative of protein unfolding [20]. Moreover, phosphofructokinase (PFK), lactate dehydrogenase (LDH), glutamate dehydrogenase (GDH), interleukin-1-receptor antagonist (IL-1ra), tumor necrosis factor binding protein (TNFbp), and ciliary neurotropic factor (CNTF) were observed to form a large amount of insoluble precipitates when quench cooled in liquid nitrogen, while a smaller cooling rate caused significant less precipitation [21]. Also, the addition of a small amount of surfactant could effectively prevent the observed precipitation. These results suggest that surface-driven denaturation at the ice-water interface is dominant for many proteins, and indicate that the behavior described in Figures 6.18a and 6.18c is commonly observed in experiments.

On the contrary, it is more difficult to find in the literature reports of proteins that behave according to Figures 6.18b and 6.18d. This occurs because many of the proteins commonly used for this type of experiments show a quite high bulk stability, and are unlikely to exhibit any significant denaturation before the onset

of ice formation. In contrast, the model protein selected in this work, namely, myoglobin, is extremely sensitive to cold denaturation, especially at the low pH (3.7) selected for this study [338, 342]. In Figure 6.21, the percentage of myoglobin recovery after freeze-thawing is shown, as measured from the OD at 280 nm (Figure 6.21a) or 410 nm (Soret band, Figure 6.21b).

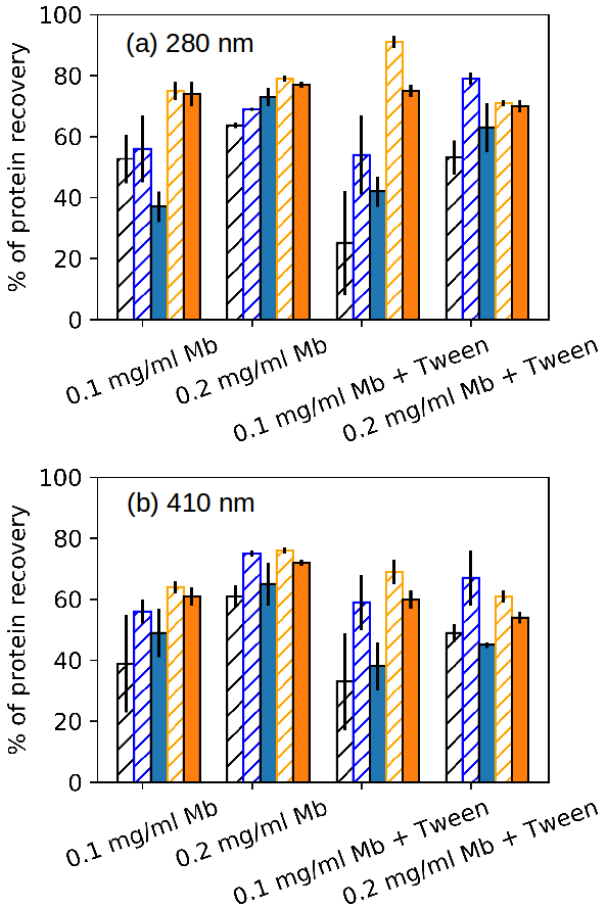

Figure 6.21: Percentage of myoglobin recovery after 1 (dashed bars) and 3 (colored bars) freeze-thaw cycles performed using the VISF controlled nucleation technique (black bars), shelf-ramped freezing at 1 K min^{-1} (blue bars) or quench freezing in liquid nitrogen (orange bars), as measured from the decrease in OD at 280 nm (a) or 410 nm (b). Data have been normalized relative to the formulation before freeze-thaw.

The peak at 280 nm is related to protein concentration. Therefore, a decrease in the peak at 280 nm after freeze-thawing and centrifugation is indicative of aggregation. By contrast, while the peak at 410 nm shows a similar dependence on concentration, it may also be affected by changes in the conformation of Mb. For instance, it

was observed that the disruption of the secondary and tertiary structure of Mb by addition of 8 M urea caused a decrease in the extinction coefficient at 410 nm, while no changes were observed for the peak at 280 nm [337]. This happens because the heme pocket modifies its conformation, and is exposed to a different environment, when Mb unfolds. The Soret band is altered whenever the physical environment of the heme changes, and therefore represents a useful indicator of the protein conformation. The different significance of the OD values at 280 nm (aggregation) and 410 nm (aggregation + denaturation) also accounts for the lower values of protein recovery generally reported in Figure 6.21b than in Figure 6.21a. Some denatured protein molecules may still be in the monomeric form, contributing to an increase in the peak at 280 nm, while not being detected at 410 nm.

As can be observed from Figures 6.21a and 6.21b, the VISF protocol (black bars) resulted in the worst protein recovery, while quench cooling in liquid nitrogen (orange bars) generally preserved activity the most, and shelf-ramped freezing at 1 K/min (blue bars) showed an intermediate behaviour. Overall, this indicates that surface-driven denaturation is not the main source of stress for the protein. In fact, the freezing protocol resulting in the largest ice-water surface area, i.e., quench freezing, led to an increase in protein recovery. By contrast, the data are compatible with bulk denaturation being the controlling mechanism. In this case, the time spent by the protein in a liquid solution at low temperature, where cold unfolding may ensue, should be minimized. The VISF protocol includes two holding steps at low temperature (268 K) and low solution viscosity (liquid or still not completely frozen matrix), that may promote the conformational changes responsible for aggregation and denaturation. The formation of a cryoconcentrated matrix occurs more quickly in the case of shelf-ramped freezing, and is extremely fast when vials are directly immersed into liquid nitrogen. These results indicate that also the second scenario identified by the simulations, and illustrated in Figures 6.18b and 6.18d, can be experimentally observed, and this suggests that the simulation approach can capture at least the main features of protein stability during freezing.

The dominant role of bulk denaturation for myoglobin at pH 3.7 is also evident from the observed effect of surfactants. When surface-driven unfolding is the main mechanism, addition of polysorbates results in significantly increased protein stability [21, 51, 176, 179]. However, addition of Tween 80 led to a reduced recovery of protein activity in our freeze-thaw experiments, as particularly evident from the absorbance values at 410 nm (Figure 6.21b). One possible explanation would be that surfactants have an effect on the protein conformation, in line with the simulation results obtained in Chapter 3. Specifically, this effect is denaturing for myoglobin at low pH. This is an interesting observation, especially considering that the literature is controversial on this point. For instance, while some authors suggested that the polysorbates may have an effect on the secondary structure of a protein in the bulk [52], others theorized that this effect was negligible [49, 53]. The disagreement on this point also arises from the close interrelation of surface-driven

phenomena and bulk effects that generally occurs when dealing with a protein formulation. Here, a situation was identified where these two effects can be more easily separated, making it possible to isolate the effect of surfactants in the bulk solution.

Finally, it is interesting that doubling the Mb concentration, from 0.1 to 0.2 mg/ml, generally increased protein recovery. According to the design space (Figure 6.19), an increase in concentration should reduce the percentage of protein molecules adsorbed onto the ice-water surface area and remarkably affect stability when surface-driven denaturation is prevailing (Figure 6.19a). However, a smaller effect should be observed when bulk unfolding is the controlling mechanism (Figure 6.19b). We may therefore hypothesize that, even though the controlling mechanism of Mb aggregation and unfolding at pH 3.7 is related to cold denaturation, adsorption onto the ice surface still plays a role, that translates into the observed improvement in protein recovery at higher concentration. This hypothesis is also supported by the smaller difference observed for the two values of protein concentration upon addition of Tween 80 (Figure 6.21). Even though Tween 80, as previously discussed, has a denaturing effect in the bulk, it still protects from ice-induced unfolding. Therefore, if minimization of surface effects is the mechanism, the improvement related to an increase in protein concentration should become negligible in presence of surfactants. The observed stabilization may, however, also be related to the formation of more stable multimers, or by volume exclusion effects that arise at higher concentration. This last explanation was also proposed in a previous work [343], where the cold denaturation temperature of β-lactoglobulin was found to decrease significantly when increasing the protein concentration.

The effect of the addition of protectants such as sucrose, mannitol and trehalose, was also investigated (see Figure 6.22), both in presence (grey bars) and absence (dashed bars) of Tween 80. Data have been normalized relative to the formulation before freeze-thaw. It is apparent that the addition of excipients improved the recovery of protein activity after freeze-thawing, and this effect was particularly remarkable in presence of Tween 80. This is in line with the stabilizing effect of sugars on cold denaturation observed in previous works [343]. It is particularly interesting that when either sucrose or trehalose was present, addition of the surfactant was not deleterious. We may hypothesize that these two amorphous excipients can effectively counteract the denaturing effect of Tween 80. As for the crystalline excipient considered in this work, i.e. mannitol, it behaved similarly to the disaccharides sucrose and trehalose in absence of surfactant. This once more suggests that surface-driven denaturation is not the main mechanism of protein denaturation for myoglobin at pH 3.7, as in this case crystallization of mannitol would create another surface onto which adsorption and unfolding may occur. However, the fraction of mannitol that remains in the amorphous phase together with the protein is probably not as effective as sucrose and trehalose in counteracting the

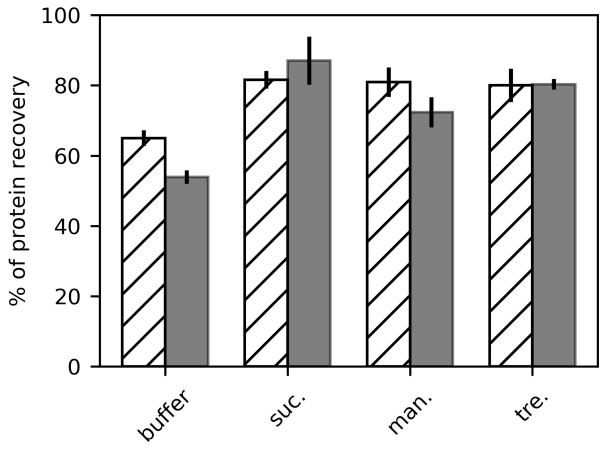

Figure 6.22: Percentage of myoglobin recovery after one freeze-thaw cycle, as measured from the decrease in OD at 410 nm. Different formulations were considered, including 5% mannitol, 5% sucrose, 5% trehalose, or citrate buffer only, both in absence (dashed bars) and presence (grey bars) of Tween 80. Data have been normalized relative to the formulation before freeze-thaw.

denaturing effect of Tween 80. As a result, a decrease in protein recovery is observed upon addition of the surfactant to the mannitol-based formulation, similarly to what occurs in citrate buffer only.

To further confirm the results obtained for Mb and LDH, their stability during a 24 h storage at 268 K was monitored, as shown in Figure 6.23.
No freezing occurred in this case, and cold denaturation could therefore be the only possible explanation for a potential decrease in protein recovery. Such decrease was not observed in the case of LDH (dashed lines in Figure 6.23) over the time considered, and the protein activity was almost completely recovered after 24 h, both in presence and absence of Tween 80. This suggests that the decrease in LDH activity observed during the freeze-thaw cycles was mostly related to the formation of the ice-water interface, in line with previous considerations [61]. Tween 80 did not remarkably affect LDH stability during storage at low temperature, and its positive effect during the freezing experiments presented in Figure 6.20 should therefore be mostly related to inhibition of protein interaction with the ice surface. The situation was reversed in the case of Mb as model protein (solid lines). In this case, a significant loss in protein recovery was observed already after 1 h storage at 268 K, and only 64 % or 56 % protein activity were eventually recovered after 24 h, in absence or presence of Tween 80, respectively. These results confirm that Mb is unstable in a bulk solution at low temperature, and that its denaturation occurs quickly, thus explaining why fast immobilization of Mb in a frozen matrix during quench freezing is beneficial. The effect of Tween 80 is again observed to be

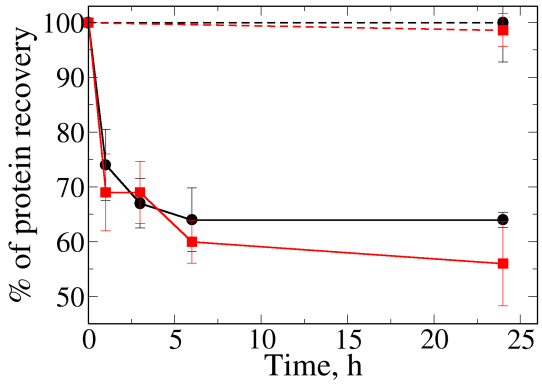

Figure 6.23: Protein recovery as function of time, in non-frozen samples, during storage at 268 K. Both 0.1 mg/ml Mb in sodium citrate buffer at pH 3.7 (solid lines) and 0.1 mg/ml LDH in sodium citrate buffer at pH 6.5 (dashed lines) were considered, either in presence (red line, square symbol) or absence (black line, circle symbol) of 0.01 % w/v Tween 80.

detrimental for Mb stability, also during storage at low temperature.

However, in addition to protein stability, process efficiency should also be considered. As previously discussed, a large pore size reduces the resistance to mass transfer, and promotes the sublimation process, thus resulting in shorter drying times. At the same time, a smaller resistance to water vapor transfer through the dried layer is associated with a lower product temperature, and therefore reduced risk of product collapse. Because of this, a low cooling rate, and a high nucleation temperature, are beneficial for process efficiency. This choice of freezing conditions also results in the best preservation of protein activity, if proteins with high bulk stability are considered. However, this is not true anymore when the protein being dried is extremely prone to cold denaturation. In this case, a high cooling rate would preserve biological activity the most, and a trade-off should therefore be achieved between efficiency and product quality.

6.5 Conclusions

Proteins are often stored in the frozen or freeze-dried state, which is generally preferred over liquid storage because the solid state results in increased stability and shelf life. To distinguish between the two scenarios here identified, i.e., controlling surface-induced or bulk degradation, the denaturation of a protein solution at low temperature, but in absence of ice, could be monitored, as shown in Figure 6.23.

If no significant aggregation/loss of activity occurs during the typical timescale of a freezing process, as for LDH at pH 6.5 in this work, then surface-induced denaturation may be dominant during freezing, and a slow freezing rate may be beneficial. Addition of surfactants in this case may help to further improve protein recovery, provided that the surfactant-protein interaction is not detrimental. On the other hand, if the protein being considered undergoes fast denaturation (i.e., faster than a typical freezing process) during the abovementioned experiment, as for Mb at pH 3.7 in this work, a fast cooling rate should be used, to entrap the protein in a solid matrix as quickly as possible. Addition of surfactants in this case may not be necessary, and may even have detrimental consequences, as observed in this work for myoglobin. If freezing is performed in larger vessels, fast freezing may also improve the homogeneity of solutes distribution throughout the geometry of the cake [344]. In all cases, addition of cryoprotectants, such as the disaccharides sucrose and trehalose, would be beneficial [343]. If the protein formulation has to be stored in the frozen state, storage should be performed below the T'_g of the matrix to avoid processes triggered by mobility above T'_g (e.g., protein unfolding and aggregation or solutes crystallization) [344]. The thawing rate may also be crucial, as slow thawing promotes the recrystallization process, where small ice crystals grow into larger ones. Recrystallization results in undesired interfacial or shear stress for proteins at the ice-water interface, which may affect their activity [345]. Considering this, a fast thawing rate may be beneficial for both bulk-denaturing and surface-denaturing proteins, as it minimizes both process duration and the risks of recrystallization. However, the effect of thawing was not addressed in the present work, and additional investigation is needed on this point. If the objective is to store the protein in a freeze-dried state, proteins with high bulk stability should be frozen at a slow cooling rate, which also results in the best process efficiency, as discussed in section 6.3. However, if the protein being dried is extremely prone to cold denaturation, a high cooling rate would preserve biological activity the most. If the increase in drying time resulting from the fast cooling rate is not economically viable, a proper choice of the formulation (i.e., addition of a sufficient amount of effective cryoprotectants) may still make it possible to obtain an acceptable protein recovery with a slower cooling rate.

List of Symbols

a	kg m K^{-1}	parameter of Equation 6.10
A		pre-exponential coefficient (Equation 6.32)
A_b	m^2	surface area of the vial base
a_S	-	empirical coefficient that accounts for the real ice crystals surface

b	$K^{2/3}\ m^{-2/3}$	coefficient of Equation 6.16
b_1	-	corrective coefficient which takes into account the real crystal habit
b_2	-	corrective coefficient which takes into account the surface irregularities
c	$mol\ l^{-1}$	excipient concentration
C	$m^{1+\lambda_1-\lambda_2}\ s^{-\lambda_1}\ K^{\lambda_2}$	empirical parameter of Equation 6.2
c_0	-	coefficient of Equation 6.35
c_1	-	coefficient of Equation 6.35
c_2	-	coefficient of Equation 6.35
Δc_p	$J\ mol^{-1}\ K^{-1}$	specific heat change upon unfolding at temperature T_0
c_w	-	water mass fraction at time t
$c_{w,end}$	-	water mass fraction at the end of freezing
d_i	m	threshold distance from the ice surface to observe an effect on protein stability
D	m	vial base diameter
D_p	m	crystals diameter
f_U	-	mole fraction of unfolded protein
ΔG	$J\ mol^{-1}$	free energy of unfolding
h	$W\ m^{-2}\ K^{-1}$	heat transfer coefficient
ΔH_0	$J\ mol^{-1}$	enthalpy of unfolding at temperature T_0
ΔH_f	$J\ kg^{-1}$	latent heat of crystallization
K	-	equilibrium constant for the unfolding process
k_f	s^{-1}	folding rate constant
k_{fr}	$W\ m^{-1}\ K^{-1}$	thermal conductivity of the frozen zone
k_u	s^{-1}	unfolding rate constant
$K_{v,e}$	$W\ m^{-2}\ K^{-1}$	heat transfer coefficient between freezing front and shelf
$K_{v,f}$	$W\ m^{-2}\ K^{-1}$	heat transfer coefficient during freezing between product base and shelf
m	$J\ mol^{-1}\ M^{-1}$	coefficient describing protein stabilization by excipients
m_w	kg	mass of water
M_p	$kg\ mol^{-1}$	molar mass of the protein
M_s	$kg\ mol^{-1}$	molar mass of sucrose
M_w	$kg\ mol^{-1}$	molar mass of water
n	-	number of discretization intervals
N	-	number of ice crystals
n_p	-	number of adsorbed protein molecules
n_s	-	number of adsorbed sucrose molecules

n_w	-	number of adsorbed water molecules
Q_s	J	enthalpy change resulting from generation of a new ice-freeze concentrate interface
Q_v	J	heat generated by crystallization
R	J mol^{-1} K^{-1}	universal gas constant
ΔS_0	J mol^{-1} K^{-1}	entropy of unfolding at temperature T_0
S_{ice}	m^2	surface of ice crystals
t	s	time
t_d	s	primary drying time
t_{end}	s	ending time of freezing
Δt	s	time interval
T	K	temperature
T_0	K	reference temperature
T_{air}	K	air temperature
T_{eq}	K	equilibrium temperature of freezing
T_g	K	glass transition temperature of amorphous solids
T_g'	K	glass transition temperature of the maximally freeze-concentrated solution
T_m	K	holding temperature after nucleation
T_{max}	K	maximum temperature during primary drying
T_n	K	nucleation temperature
T_{shelf}	K	shelf temperature
$\bar{T}_{shelf}$	K	average shelf temperature
ΔT	K	temperature difference between air and product
V_{ice}	m^3	volume of ice
V_p	m^3	molecular volume of the protein
V_s	m^3	molecular volume of sucrose
V_w	m^3	molecular volume of water
z	m	axial coordinate
z_f	m	axial coordinate of the freezing front
Δz	m	axial interval
Δz_f	m	variation of the freezing front position

Greek Letters

α	-	empirical parameter of Equation 6.1
β	-	exponent of Equation 6.1
γ	J m^{-2}	solid-solid interfacial tension
δ	-	crystals inclination angle

ε	-	ratio between the volume of ice and the total volume of the system
θ	K m^{-1}	temperature gradient within the frozen zone
$\bar{\theta}$	K m^{-1}	time average of the temperature gradient within the frozen zone
λ_1	-	exponent of Equation 6.2
λ_2	-	exponent of Equation 6.2
μ	Pa s	viscosity
μ_0	Pa s	viscosity at reference conditions
ν	m s^{-1}	freezing front velocity
ϱ_{ice}	kg m^{-3}	density of ice
τ	-	tortuosity
ϕ	-	fraction of water which crystallizes

Chapter 7

Conclusions

In this last Chapter, a summary of the main results obtained in the thesis will be provided, with the objective of highlighting the thread that runs through the entire work. This result will be achieved by answering to the questions listed at the end of Chapter 1, which summarize the core findings presented in this work.

1. What are the molecular properties related to cryo- and lyoprotection? Is the protection mechanism sequence-specific?

   In Chapter 2, the role of cryo- and lyoprotectants was addressed, using hGH and LDH as model proteins. This first analysis suggested that differences exist among excipients, with the disaccharides being better than polyols, monosaccharides and amino acids both during freezing and in the dried state. Cryoprotection has been found to be a thermodynamic mechanism, mostly related to preferential exclusion of the excipient from the protein surface, while the formation of a glassy, viscous matrix, that kinetically hinders protein motions, is mainly responsible for lyoprotection. Some molecular properties were also identified, that seem to correlate with the protective effect of stabilizers. The higher the molecular volume was, the more the osmolyte was found to be excluded from the protein surface and to act as a good cryoprotectant, while a high hydrogen bonding propensity was observed to be linked to the efficiency as a lyoprotectant.

   In Chapter 4 a “bad solvent theory” was also proposed to explain protein stabilization by excipients. According to this theory, a protein is stabilized in a poor solvent, which favors the collapsed, native fold against the expanded, unfolded states. However, proteins result from the combination of 20 different amino acids, showing different charge and hydrophobicity/philicity properties. Because of this, the surface of proteins exposes patches having extremely heterogeneous characteristics. As a result of this heterogeneity, it is not easy to define a "bad solvent" for a protein as a whole. In contrast, as shown in

Chapter 4, it is easier to observe preferential exclusion from specific regions of the protein. That is, a good solvent for specific patches on the protein surface could be, at the same time, a poor solvent for other side-chain sequences of the same protein.

On the one hand, the observation that certain amino acid sequences are more important than others seems to indicate a sequence-dependent mechanism of protein stabilization. On the other hand however, it should be considered that the osmolytes generally affect proteins in similar ways. For instance, urea is a destabilizing osmolyte for any protein, while sugars, many polyols, amino acids and trimethylamine N-oxide (TMAO) always show a protective effect. The osmolyte effect is, therefore, universal, and operates on proteins in general. As already suggested [269], this universality implies a backbone-based mechanism of protein folding.

The following scenario was therefore proposed in this thesis, where the folding transition is envisioned as a multistep process. At the beginning, the unfolded state of the protein exposes its backbone to the solvent. If the solvent is a bad solvent for the backbone, the unfavorable backbone-solvent interactions promote the refolding process. The folding transition is then guided by the side chains, that drive the protein conformation toward the specific and unique native fold [269]. According to this theory, exclusion of the osmolytes from the protein backbone is crucial for starting the folding transition. A protective osmolyte could even favorably interact with specific side chains, provided that it is a bad solvent for the backbone. For instance, it was shown in Chapter 3 that the sugars can reduce the risk of unfolding, even though they are attracted by the polar side chains on the surface of hGH. However, also the preferential exclusion from specific side-chain sequences on the protein surface may be important to prevent undesired conformational changes. Exclusion from the aggregation-prone regions, for instance, can minimize the risk of aggregation.

2. What is the role of buffers? Do they have any other property, other than pH control, which may affect protein stability?

The effect of common buffer species, namely, phosphate and citrate buffers, was also investigated in Chapter 2. The data presented in this work suggest that pH control is not the only effect that buffers may have on protein stability. For instance, they may create a sugar-buffer hydrogen-bonding network, that disrupts, and substitutes, the original matrix formed between sugar molecules. Because of this, buffers were found to increase preferential exclusion of both sucrose and trehalose, but negatively affected the sugars ability to substitute water during drying.

3. How do surfactants prevent protein aggregation and surface adsorption? Do they also have an effect on the secondary structure of proteins? If so, what is the mechanism behind their action?

   In Chapter 3, attention was shifted to surfactants, and specifically to the commonly used polysorbates Tween 20 and Tween 80. Their effect was investigated both in the bulk and at the air-water, ice-water and silica-water interfaces, using hGH and the GB1 peptide as model systems. Simulation results for hGH and Tween 20 suggest that polysorbates should be able to effectively prevent aggregation, and that they may have an effect on protein conformation in the bulk, as well. An orientation-dependent mechanism was identified, in which the protein structure is stabilized if the surfactant hydrophilic heads are oriented towards the protein. The ability of surfactants to preferentially locate at interfaces, preventing protein adsorption, was further confirmed for the case of GB1 at the air-water interface. Here, Tween 80 could prevent the unfolding process promoted by exposure of the hydrophobic residues of the protein to the vacuum. Addition of Tween 80 also limited GB1 adsorption at the silica-water interface, but no improvement in peptide stability was observed in this case. On the contrary, partial unfolding was even observed upon addition of the polysorbate. The mechanism was completely different at the ice surface, where the surfactant molecules preferentially surrounded the protein, rather than the ice-water interface. This translated into improved protein stability, again because the Tween 80 molecules oriented with their hydrophilic heads towards the peptide.

4. Which of the interfaces typically encountered during a freeze drying process is more dangerous for protein denaturation, and how do they induce conformational changes?

   In Chapter 3 it was observed that GB1 was destabilized at the air-water and ice-water interfaces, but stabilized at the silica surface. In the case of the air-water interface, destabilization was caused by direct interaction of the hairpin turn and of the hydrophobic pair formed by residues 43W-54V with the vacuum. This strong interaction promoted a large exposure of nonpolar surface area and a resulting loss of structure. In the case of the ice-water surface, the interaction of residues 43W-54V with the surface promoted a partial loss of structure; however, the turn structure was retained, and complete unfolding was not observed. By contrast, the GB1 hairpin was stabilized at the silica surface (representative of the glass-water interface) because of confinement effects and absence of strong protein-surface interactions.

5. What is the role of the ice-water interface? How does it affect protein stability?

   The results presented in Chapter 3 suggest the existence of a peculiar behavior at the ice surface, which was further analyzed in Chapter 4. Here, the protective effect of different excipients on hGH during freezing was found to be related to their preferential exclusion from specific patches of the protein surface. Interestingly, these patches corresponded to the aggregation prone regions (APR) of the protein. A similar observation could be drawn for protein L, where the stabilizing effect of glucose and denaturing effect of the ice interface were addressed using metadynamics. The ice surface was found to slow down the surrounding liquid water molecules, promoting solvent penetration of the protein, and the hydration of regions which were poorly surface-exposed in the native state. An explanation of the ice effect as an enhancement of cold denaturation phenomena was therefore advanced.

6. May a controlled nucleation approach, resulting in the formation of a small ice-water surface area, be beneficial for protein stability?

   In Chapter 5, the effect of the controlled nucleation approach known as vacuum induced surface freezing, that results in the formation of large ice crystals and therefore small ice-water surface area, has been addressed. While controlled nucleation did not result in significant improvement in protein recovery for hGH, it increased the long-term stability of FVIII in freeze dried plasma samples. This observation, coupled with the positive effect of controlled nucleation on process efficiency and reconstitution time of highly-concentrated products, makes vacuum induced surface freezing a promising technique for application in the freeze-drying industry.

7. What freezing conditions result in the best preservation of protein residual activity? Are these conditions beneficial for process performance, as well?

   The results presented in the first Chapters made it clear that the freezing step of freeze-drying plays a crucial role, especially when therapeutic proteins are considered. Chapter 6 was therefore devoted to the description of a simulation approach for the freezing of pharmaceuticals. A mechanistic model for the prediction of the ice crystal size distribution in frozen solutions was first presented, and validated upon experimental data. This model was subsequently coupled with a set of equations that can describe the heat and mass fluxes during the primary drying phase of lyophilization. As a further step, the effect of freezing on protein behavior was also taken into account, and two different scenarios were identified, based on the relative stability of the protein

in the bulk compared to the interface. In one case, surface-driven unfolding prevails for proteins having high bulk stability, and the ice crystals size should be minimized to reduce loss of activity. On the other hand, the duration of the freezing process itself is the key parameter if proteins with reduced bulk stability are considered. Both these scenarios were experimentally observed, using LDH and Mb as model proteins. However, in addition to protein stability, process efficiency should also be considered, and in this case a low cooling rate and a high nucleation temperature are beneficial. This choice of freezing conditions also results in the best preservation of protein activity, if proteins with high bulk stability are considered. This is not true anymore when the protein being dried is extremely prone to cold denaturation. In this case, a high cooling rate would preserve biological activity the most, and a trade-off should therefore be achieved between efficiency and product quality.

7.1 Closing Remarks

Overall, different modelling techniques have been used in the present work to provide insight into the freeze drying of protein-based pharmaceuticals. When the distribution of excipients around a protein is the main focus, common molecular dynamics simulations may be enough to provide the desired information, as shown in Chapter 2. However, when the interaction between protein molecules, or between proteins and surfaces, is the objective, enhanced sampling techniques may help. If the system is simple enough, umbrella sampling may be a good choice, as discussed in Chapter 3. However, when several slow degrees of freedom need to be biased, metadynamics, and in particular the variant known as parallel bias metadynamics, may help, as observed in Chapters 3 and 4. If the whole process has to be simulated, modelling should be carried out on a different scale, using for instance macroscopic balance equations as discussed in Chapter 6. It is our conviction that these simulation approaches may prove beneficial to make the whole process more time- and cost-effective, improving the recovery of therapeutic activity at the same time. These models are not thought to substitute experiments, but, rather, they should complement them. In the case of the molecular dynamics simulations used in the first part of this work, they can provide useful information about the molecular mechanisms at the basis of proteins stabilization by excipients, which are often not accessible by current experimental techniques. This knowledge of molecular-level phenomena provided by molecular dynamics could be used to design a formulation to be lyophilized using a rational, knowledge-driven approach, in opposition to the present experience-based one. However, due to the limited time and length scales accessible by molecular dynamics, and the difficulty of reproducing in a simulation box the complexity of reality, experimental tests remain an essential tool, which should be used to validate the simulation results.

At the same time, the modelling approach presented in Chapter 6 could be used to speed up the exploration of operating conditions, making it easier (and faster) to identify those parameters that maximize both process efficiency and product quality. As an example, the design space shown in Chapter 6 could be obtained in a few hours using simulations, while it would have required months if experiments had to be performed. However, the mass and energy balance equations at the basis of this *in-silico* technique needs to be validated, and some of the parameters included in the equations can only be obtained by fitting of experimental data. Hence, a theme that permeates this whole work is the complementary nature of simulations and experiments. They cannot be conceived as standalone approaches, but should strengthen each other in a joint effort to increase our knowledge of natural phenomena.

Appendix A

Water Models Used for the Simulations

A water model is a list of parameters, typically Lennard-Jones coefficients σ and ε (see Equation 1.7) and atomic charges q, used to describe in a simulation the behavior of water molecules. While water has three atoms, the complexity of its behavior often requires more than 3 sites for a complete description of all the properties of water.

In this work, the SPC/E [124], TIP3P [242], TIP4P and TIP4P/Ice [243] models have been used. The first two are three-site models, i.e., are characterized by three interaction points corresponding to the three atoms of water. The last two are four-site models, because a dummy atom M has been added near the oxygen, along the bisector of the HOH angle, to better describe the electrostatic distribution of the molecule (see Figure A.1).

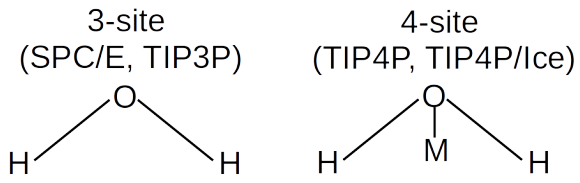

Figure A.1: Scheme of 3-site and 4-site water models.

Table A.1 reports a list of parameters for the water models used in the present work. The OH and, when applicable, OM distances, r(OH) and r(OM), are reported, together with the HOH angle. As regards the nonbonded interaction parameters, the σ and ε values for the oxygen atom are also shown. For the hydrogen and dummy atoms these values are always zero, and are therefore not listed. Finally, the charges on O, H and M atoms, used for the calculation of Coulomb interactions,

are also shown in the Table.

Table A.1: List of parameters for the water models used in the present work

	SPC/E	TIP3P	TIP4P	TIP4P/Ice
r(OH), nm	0.1	0.09572	0.09572	0.09572
r(OM), nm	-	-	0.015	0.015
HOH, °	109.47	104.52	104.52	104.52
σ, nm	0.317	0.315	0.315	0.317
ε, kJ/mol	0.650	0.636	0.649	0.882
q(O), e	-0.8476	-0.834	0	0
q(H), e	0.4238	0.417	0.52	0.5897
q(M), e	-	-	-1.04	-1.1794

Three-site models are widely used because are extremely computationally efficient, but four-site models are sometimes better for specific applications. For instance, as discussed in Chapter 4, the TIP4P/Ice model improves the description of solid ice water compared to other models. In this work, the SPC/E model was used in Chapters 2 and 3 because it is the water model suggested for use with the Gromos force field. The TIP3P, TIP4P and TIP4P/Ice water models were used in Chapter 4 in combination with the OPLS force field. More specifically, when an accurate description of water freezing was desired, the TIP4P/Ice model was used.

List of Symbols

q	C	atomic charge
r	m	distance

Greek Letters

ε	J mol^{-1}	Lennard-Jones coefficient
σ	m	Lennard-Jones coefficient

Appendix B

Convergence of the PBMetaD Simulations

In Chapter 3 the conformational stability of the GB1 peptide at the air-water, ice-water and silica-water interfaces was studied using parallel bias metadynamics (PBMetaD). The presence of surfactant molecules (Tween 80) was also considered for this study. The same computational technique was used in Chapter 4 to address the folding behavior of protein L, both in presence and absence of the ice-water interface or of a 1 M glucose concentration. In PBMetaD, several one-dimensional bias potentials are deposited on different collective variables (CVs), and for each CV a one-dimensional free energy profile is generated at the end of the simulation. Up to 5 different CVs were baised for GB1, including the GB1 antiparallel β-sheet content (β_s), radius of gyration (R_g) and AlphaBeta similarity ($\alpha\beta$), the distance between the protein COM and the surface (d), and the coordination number of the Tween 80 hydrophobic tails (CN_t) or heads (CN_h) around the protein surface. For protein L, up to 4 different CVs were selected, namely the protein L α-helix (α_h) and antiparallel β-sheet (β_s) content, its distance root mean square deviation for the backbone atoms with respect to the crystal structure (dRMSD), and the distance between the protein COM and the ice surface (d). In order to estimate the convergence of the PBMetaD simulations performed, the fluctuations of the free energy profiles were monitored. The results of this analysis are shown in Figures B.1-B.10 for GB1, and Figures B.11-B.14 for protein L. The fluctuations over the last 10% of the simulation time have been shaded in grey in the Figures.

List of Symbols

CN_h	-	coordination number of the surfactant hydrophilic heads

CN_t	-	coordination number of the surfactant hydrophobic tails
d	m	distance
dRMSD	m	distance root mean square deviation
R_g	m	radius of gyration

Greek Letters

α_h	-	α-helix content
$\alpha\beta$	-	AlphaBeta similarity collective variable
β_s	-	antiparallel β-sheet content

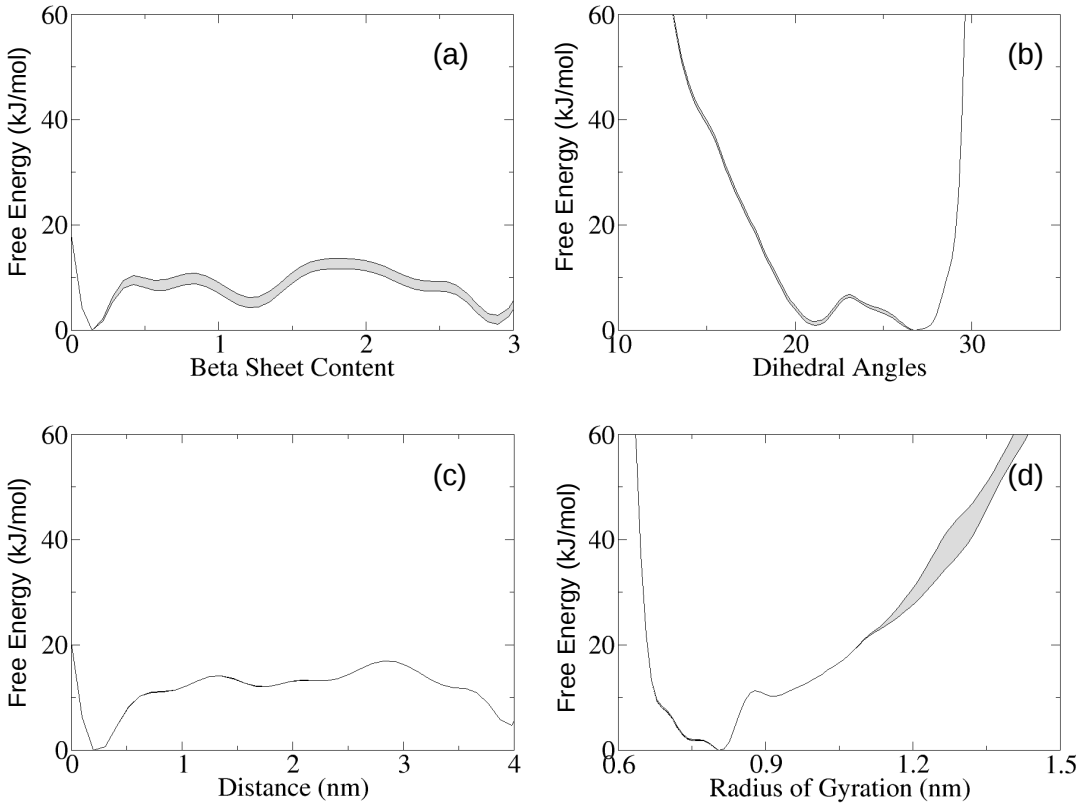

Figure B.1: Free energy profiles for the GB1 hairpin as function of (a) β-sheet content (β_s), (b) AlphaBeta similarity ($\alpha\beta$), (c) distance between the protein COM and the surface (d), and (d) radius of gyration (R_g). The graphs refer to the air-water interface, in absence of surfactants (simulation 1 in Table 3.4). The fluctuations in the measured profiles over the last 10% of the simulation time are shaded in grey.

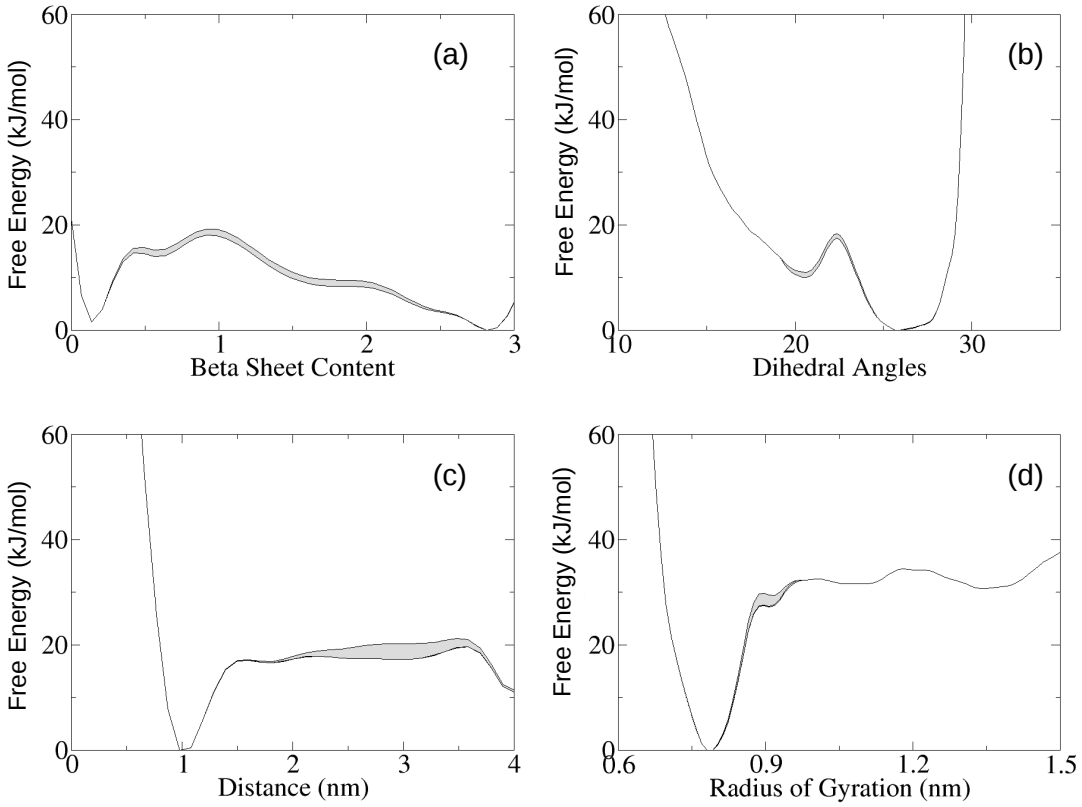

Figure B.2: Free energy profiles for the GB1 hairpin as function of (a) β-sheet content (β_s), (b) AlphaBeta similarity ($\alpha\beta$), (c) distance between the protein COM and the surface (d), and (d) radius of gyration (R_g). The graphs refer to the ice-water interface, in absence of surfactants (simulation 2 in Table 3.4). The fluctuations in the measured profiles over the last 10% of the simulation time are shaded in grey.

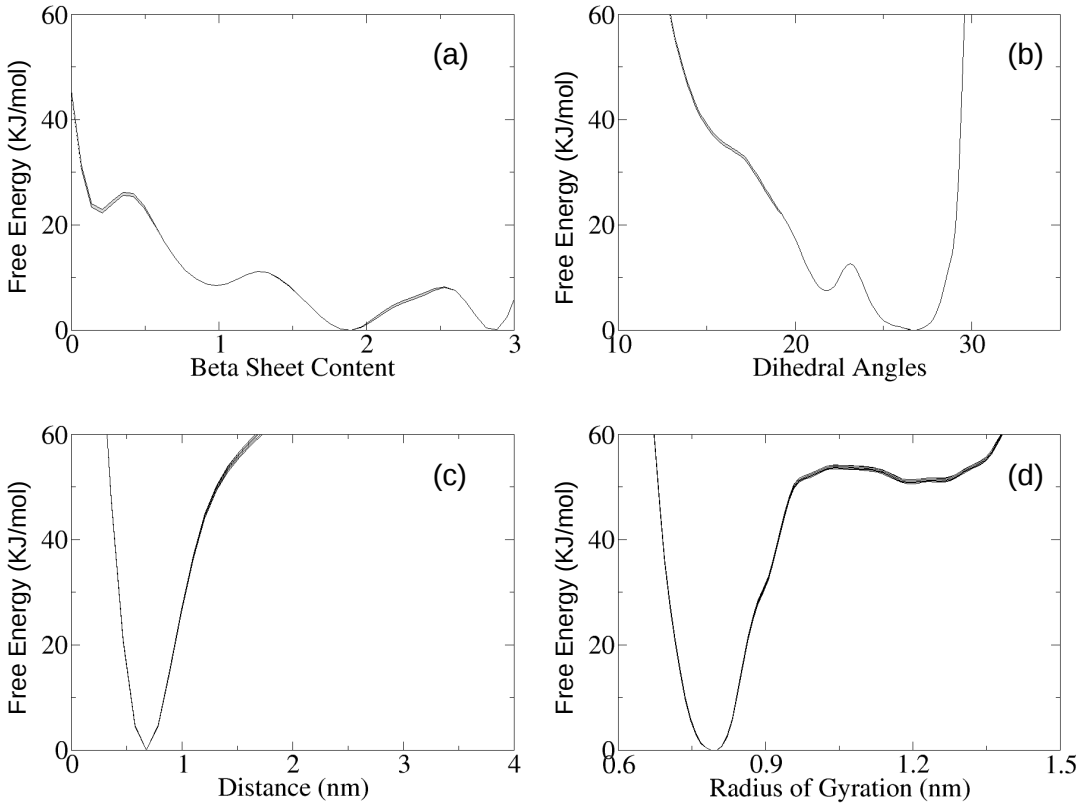

Figure B.3: Free energy profiles for the GB1 hairpin as function of (a) β-sheet content (β_s), (b) AlphaBeta similarity ($\alpha\beta$), (c) distance between the protein COM and the surface (d), and (d) radius of gyration (R_g). The graphs refer to the silica-water interface, in absence of surfactants (simulation 3 in Table 3.4). The fluctuations in the measured profiles over the last 10% of the simulation time are shaded in grey.

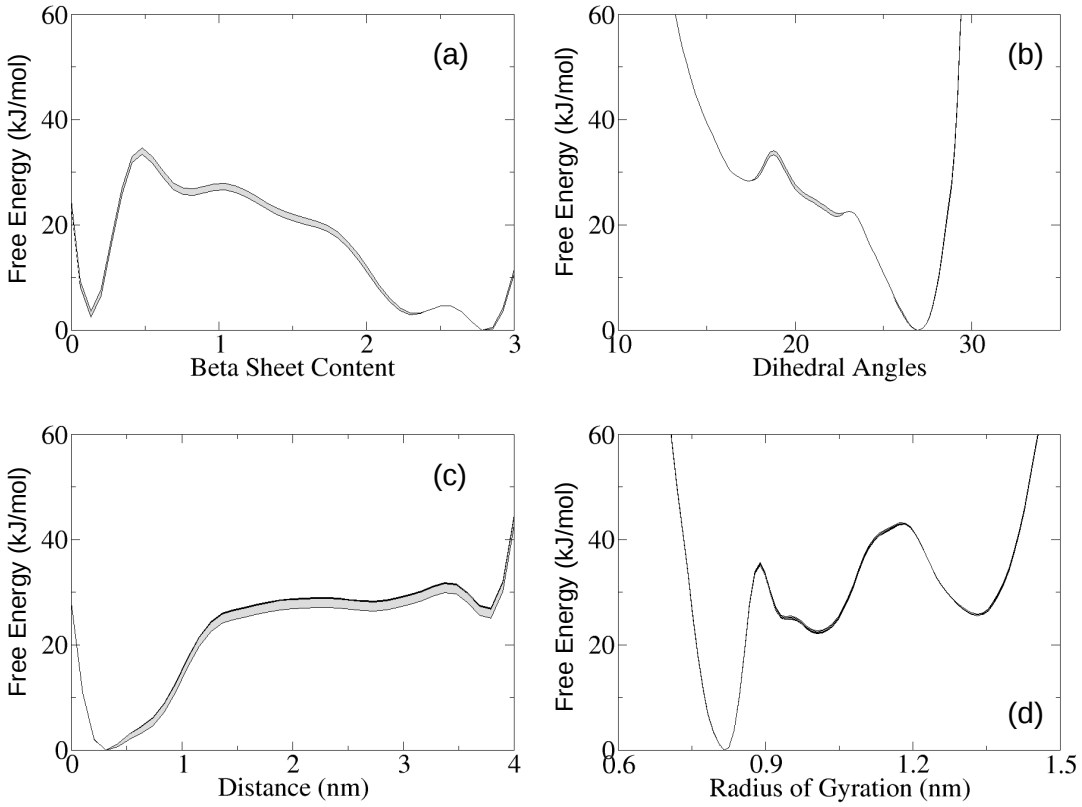

Figure B.4: Free energy profiles for the GB1 hairpin as function of (a) β-sheet content (β_s), (b) AlphaBeta similarity ($\alpha\beta$), (c) distance between the protein COM and the surface (d), and (d) radius of gyration (R_g). The graphs refer to the air-water interface, in presence of surfactants (simulation 4 in Table 3.4). The fluctuations in the measured profiles over the last 10% of the simulation time are shaded in grey.

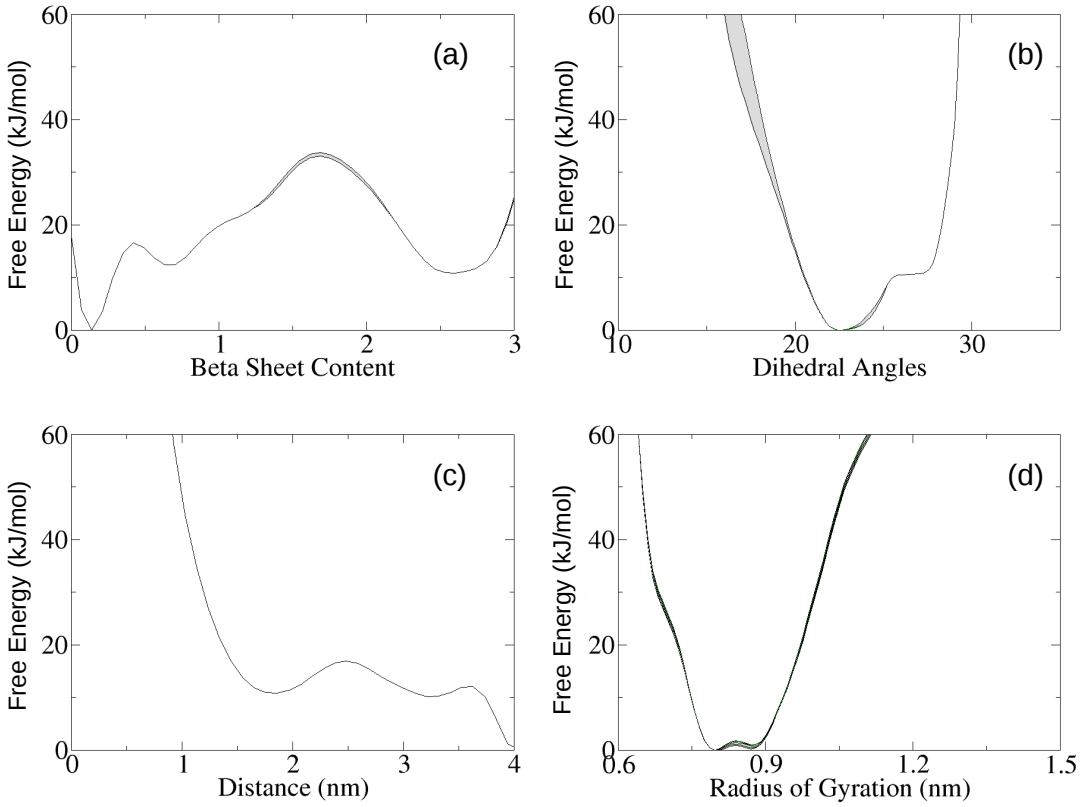

Figure B.5: Free energy profiles for the GB1 hairpin as function of (a) β-sheet content (β_s), (b) AlphaBeta similarity ($\alpha\beta$), (c) distance between the protein COM and the surface (d), and (d) radius of gyration (R_g). The graphs refer to the ice-water interface, in presence of surfactants (simulation 5 in Table 3.4). The fluctuations in the measured profiles over the last 10% of the simulation time are shaded in grey.

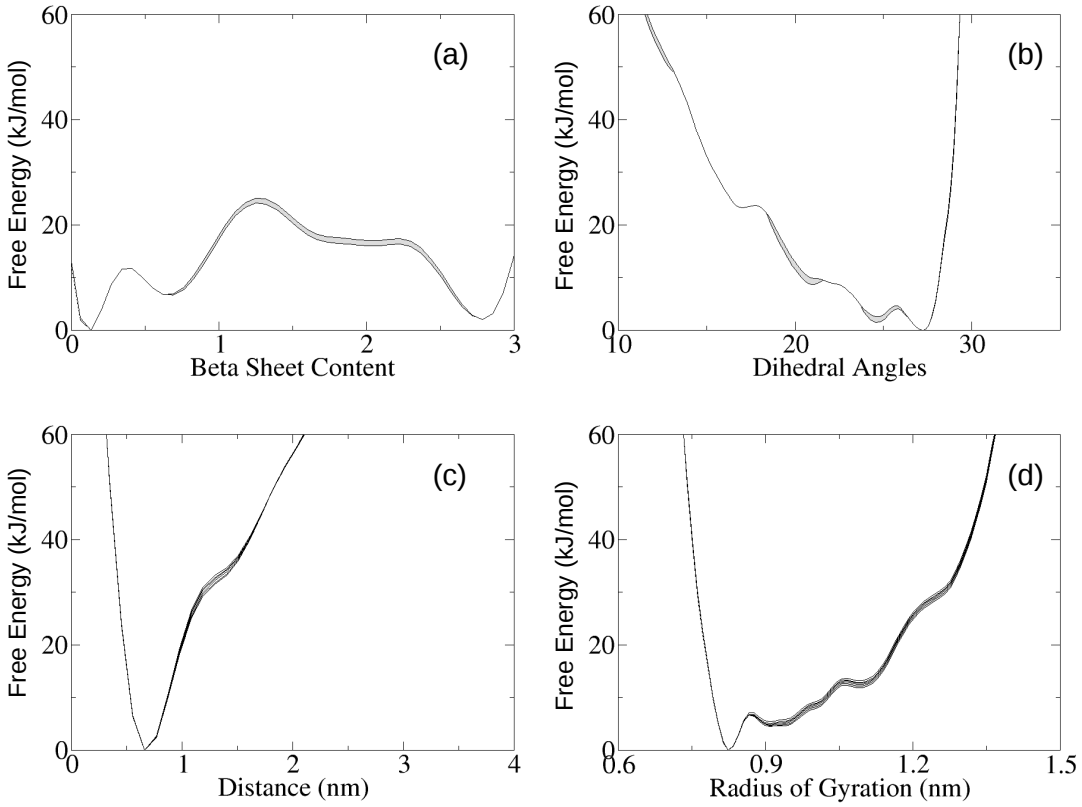

Figure B.6: Free energy profiles for the GB1 hairpin as function of (a) β-sheet content (β_s), (b) AlphaBeta similarity ($\alpha\beta$), (c) distance between the protein COM and the surface (d), and (d) radius of gyration (R_g). The graphs refer to the silica-water interface, in presence of surfactants (simulation 6 in Table 3.4). The fluctuations in the measured profiles over the last 10% of the simulation time are shaded in grey.

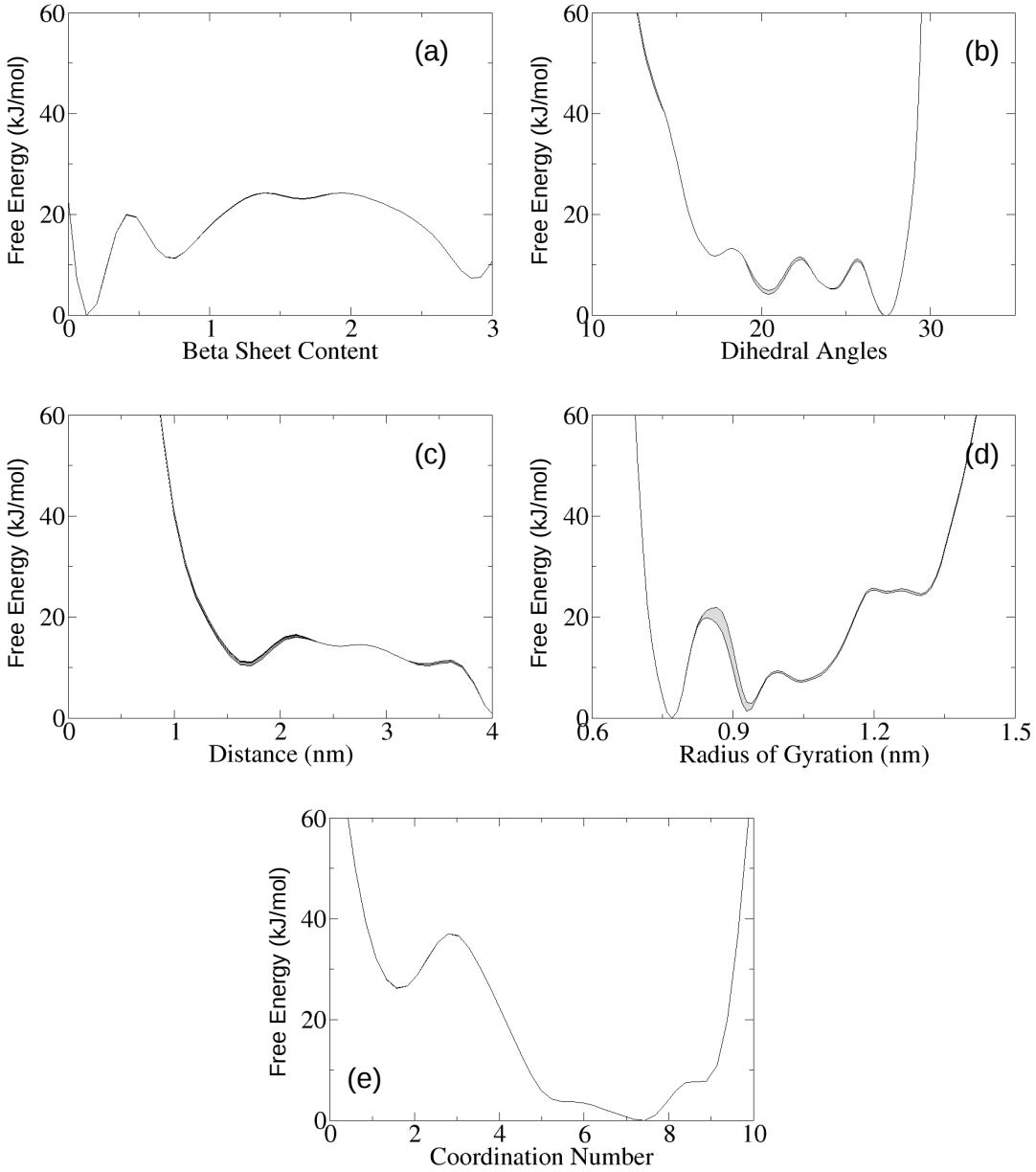

Figure B.7: Free energy profiles for the GB1 hairpin as function of (a) β-sheet content (β_s), (b) AlphaBeta similarity ($\alpha\beta$), (c) distance between the protein COM and the surface (d), (d) radius of gyration (R_g), and (e) coordination number of the Tween 80 hydrophobic tails (CN_t) around the protein surface. The graphs refer to the ice-water interface, in presence of surfactants (simulation 7 in Table 3.4). The fluctuations in the measured profiles over the last 10% of the simulation time are shaded in grey.

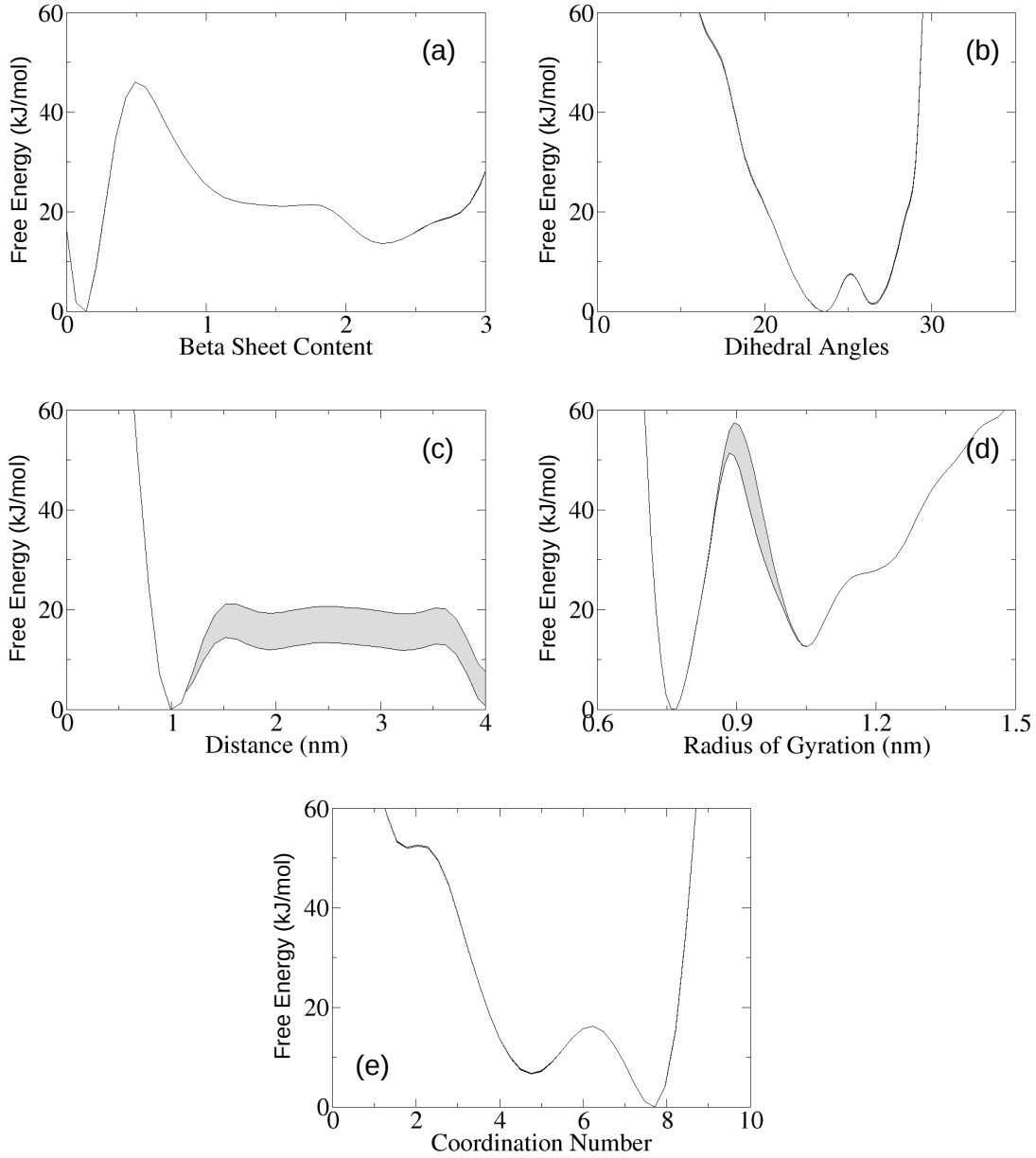

Figure B.8: Free energy profiles for the GB1 hairpin as function of (a) β-sheet content (β_s), (b) AlphaBeta similarity ($\alpha\beta$), (c) distance between the protein COM and the surface (d), (d) radius of gyration (R_g), and (e) coordination number of the Tween 80 hydrophilic heads (CN_h) around the protein surface. The graphs refer to the ice-water interface, in presence of surfactants (simulation 8 in Table 3.4). The fluctuations in the measured profiles over the last 10% of the simulation time are shaded in grey.

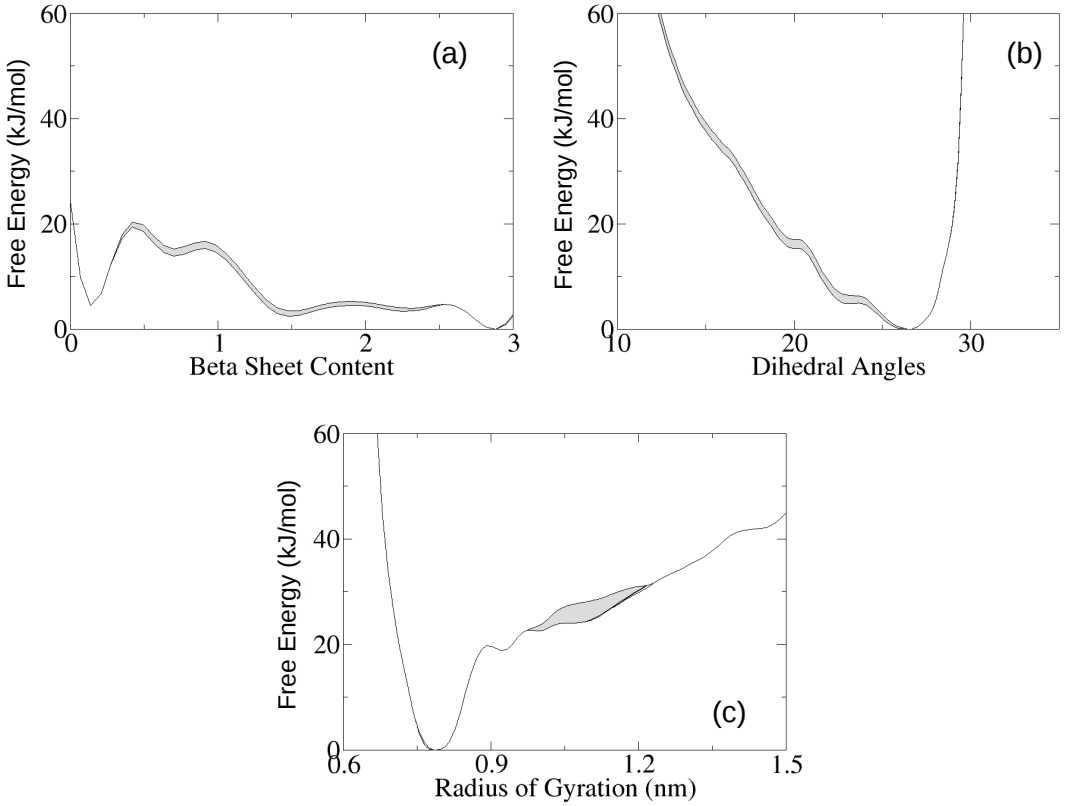

Figure B.9: Free energy profiles for the GB1 hairpin as function of (a) β-sheet content (β_s), (b) AlphaBeta similarity ($\alpha\beta$), and (c) radius of gyration (R_g). The graphs refer to the bulk solution, in absence of surfactants (simulation 9 in Table 3.4). The fluctuations in the measured profiles over the last 10% of the simulation time are shaded in grey.

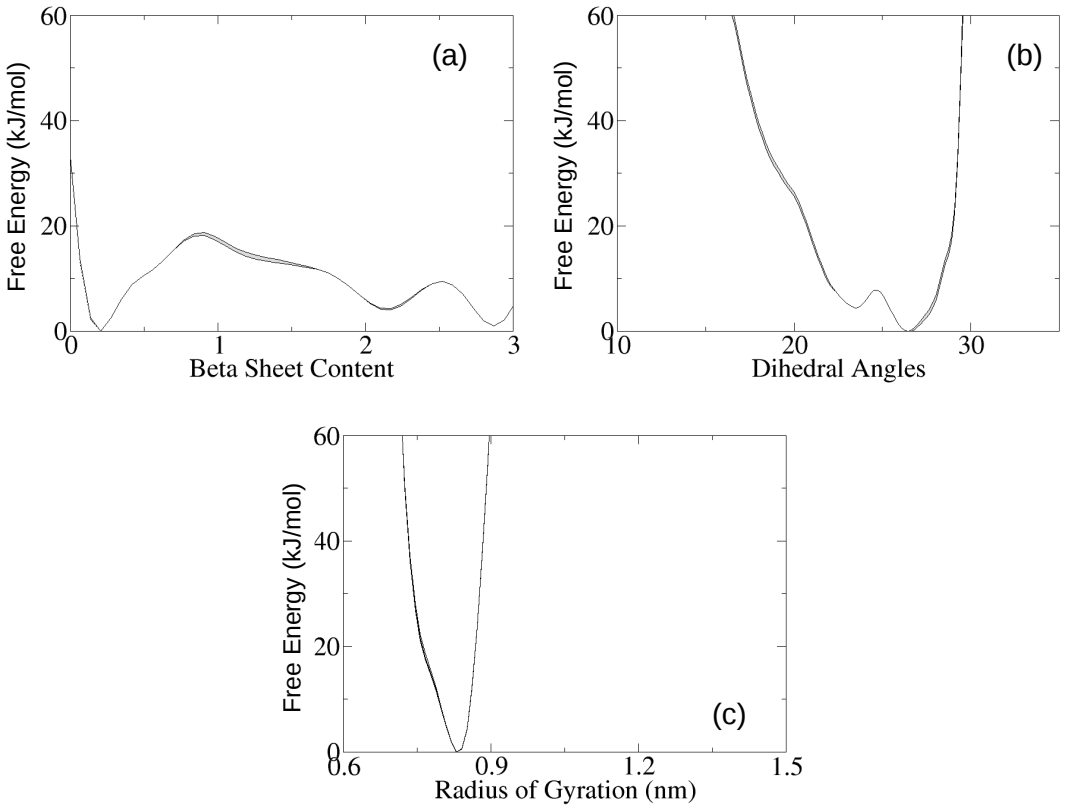

Figure B.10: Free energy profiles for the GB1 hairpin as function of (a) β-sheet content (β_s), (b) AlphaBeta similarity ($\alpha\beta$), and (c) radius of gyration (R_g). The graphs refer to the bulk solution, in presence of surfactants (simulation 10 in Table 3.4). The fluctuations in the measured profiles over the last 10% of the simulation time are shaded in grey.

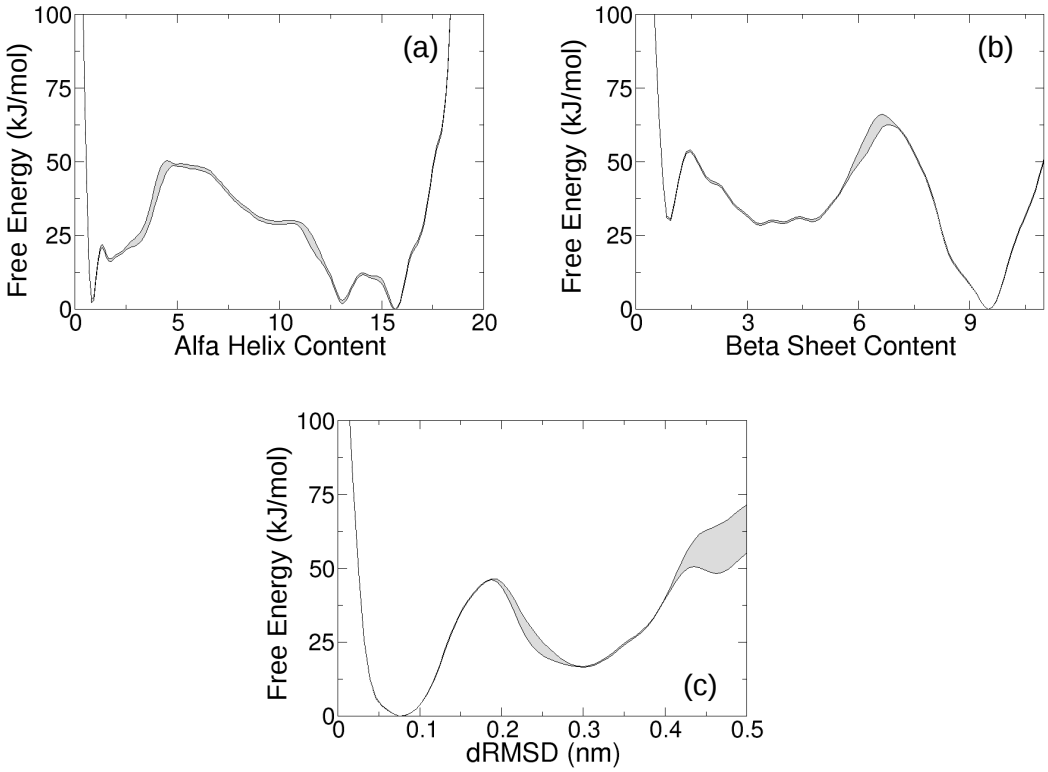

Figure B.11: Free energy profiles for protein L as function of (a) α-helix content (α_h), (b) β-sheet content (β_s) and (c) distance root mean square deviation (dRMSD). The graphs refer to the simulation in bulk water (simulation 1 in Table 4.4). The fluctuations in the measured profiles over the last 10% of the simulation time are shaded in grey.

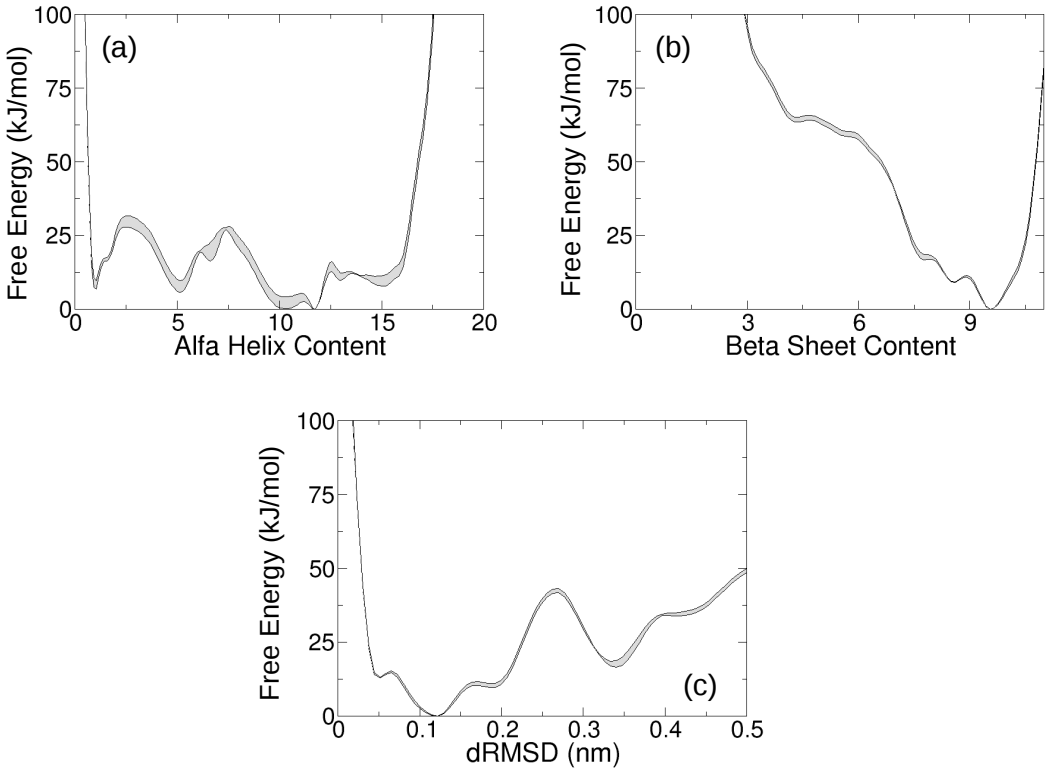

Figure B.12: Free energy profiles for protein L as function of (a) α-helix content (α_h), (b) β-sheet content (β_s) and (c) distance root mean square deviation (dRMSD). The graphs refer to the simulation in 1M glucose (simulation 2 in Table 4.4). The fluctuations in the measured profiles over the last 10% of the simulation time are shaded in grey.

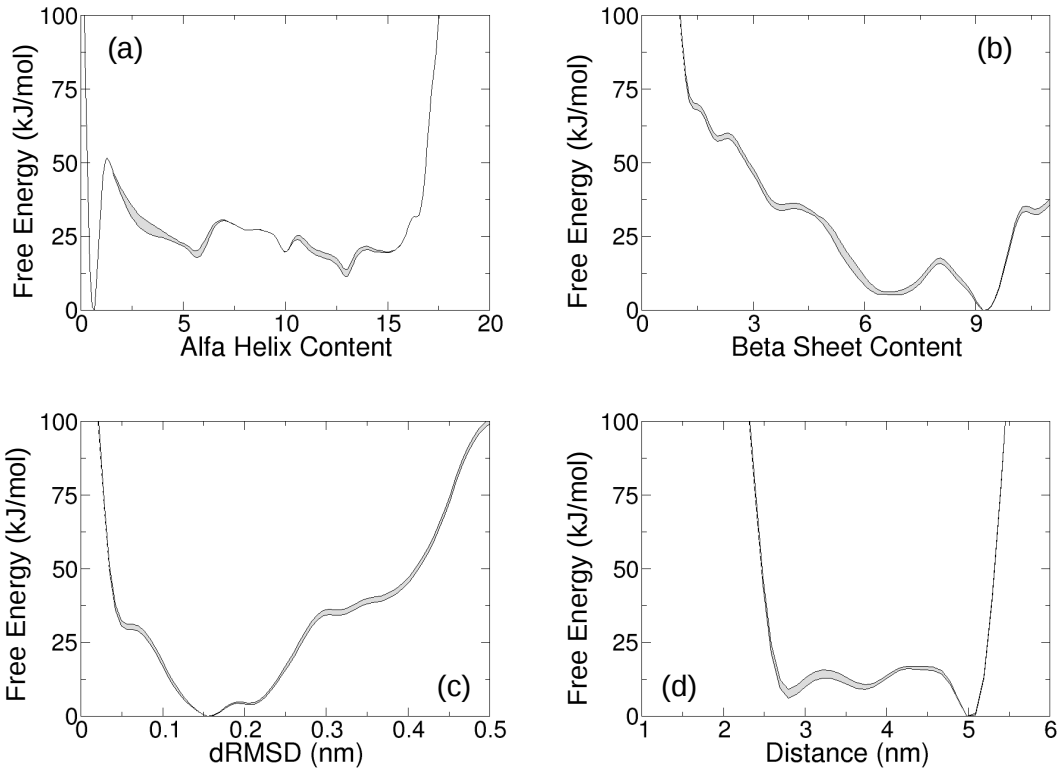

Figure B.13: Free energy profiles for protein L as function of (a) α-helix content (α_h), (b) β-sheet content (β_s), (c) distance root mean square deviation (dRMSD) and (d) distance between the protein COM and the ice surface (d). The graphs refer to the simulation at the ice-water interface (simulation 3 in Table 4.4). The fluctuations in the measured profiles over the last 10% of the simulation time are shaded in grey.

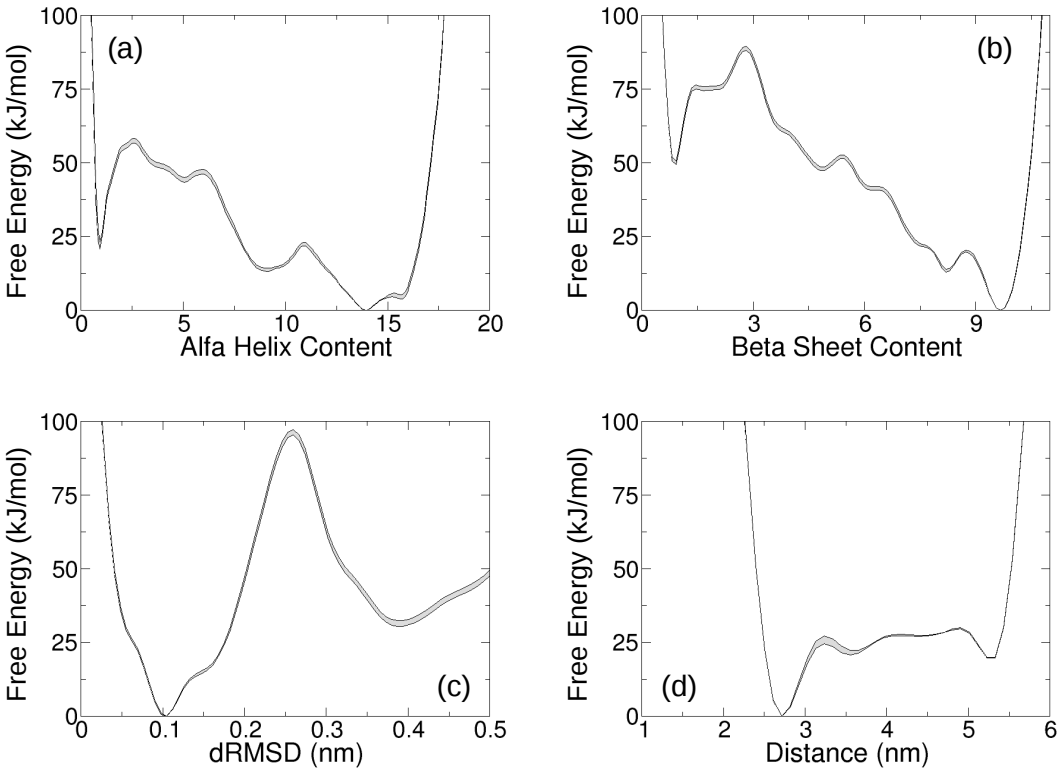

Figure B.14: Free energy profiles for protein L as function of (a) α-helix content (α_h), (b) β-sheet content (β_s), (c) distance root mean square deviation (dRMSD) and (d) distance between the protein COM and the ice surface (d). The graphs refer to the simulation at the ice-water interface, in presence of 1M glucose (simulation 4 in Table 4.4). The fluctuations in the measured profiles over the last 10% of the simulation time are shaded in grey.

Appendix C

Models of the Freezing and Drying Steps of Freeze Drying

C.1 Freezing

According to the model by Nakagawa *et al.* and Pisano and Capozzi [320,323], the cooling phase of freezing is described by the second Fourier's law,

$$\varrho_l c_{p,l} \frac{\partial T}{\partial t} = \nabla(k_l \nabla T) \tag{C.1}$$

where the density ϱ_l, specific heat $c_{p,l}$ and thermal conductivity k_l of the solution are assumed to be constant.

During the ice crystal growth period, the sample can be divided into zones corresponding to different physical states, namely a solid zone, a mushy zone (having ice fraction X_{ice}) and, possibly, a clear liquid zone. Mathematical modelling of this step is based on the following energy balance, which becomes valid only after nucleation (as previously discussed, the nucleation temperature, or nucleation time, is a parameter in the model),

$$\varrho c_p \frac{\partial T}{\partial t} = \nabla(k \nabla T) + Q_n + Q_c \tag{C.2}$$

where the terms Q_n and Q_c are source terms that represent heat generation due to nucleation and ice crystal growth respectively. It is assumed that the rate of nucleation is proportional to the degree of supercooling $T_{eq} - T^*$, where T_{eq} represents the equilibrium temperature of freezing, while T^* is the temperature of the supercooled liquid. Therefore, the source term Q_n, always positive, can be expressed as:

$$Q_n = \Delta H_f k_n (T_{eq} - T^*) \tag{C.3}$$

where ΔH_f is the latent heat of crystallization and k_n is a kinetic constant.

As regards the source term Q_c, it corresponds to the latent heat generated by ice crystallization and can be expressed as:

$$Q_c = \Delta H_f \frac{\partial}{\partial t}(\varrho_{ice} X_{ice}) \tag{C.4}$$

Moreover, assuming that the crystal growth rate is entirely controlled by heat transfer kinetics and that the ice fraction X_{ice} changes linearly with temperature, the source term Q_c can be further expressed as:

$$\Delta H_f \frac{\partial}{\partial t}(\varrho_{ice} X_{ice}) = -\Delta H_f \varrho_{ice} \frac{\partial X_{ice}}{\partial T} \frac{\partial T}{\partial t} = -\Delta H_f \varrho_{ice} \frac{\Delta X_{ice}}{\Delta T} \frac{\partial T}{\partial t} \tag{C.5}$$

where ΔT represents the temperature difference through the mushy zone. Thus, this source term can be included in the accumulation term using an apparent specific heat capacity c_p^* defined as:

$$c_p^* = \begin{cases} c_{p,l} & T > T_{eq} + \Delta T \\ \frac{c_{p,l}+c_{p,ice}}{2} + \frac{\Delta H_f}{\Delta T} \Delta X_{ice} & T_{eq} \leq T \leq T_{eq} + \Delta T \\ c_{p,ice} & T < T_{eq} \end{cases} \tag{C.6}$$

where $c_{p,l}$ and $c_{p,ice}$ are the specific heat of liquid and ice respectively. Therefore, Equation C.2 becomes:

$$\varrho c_p^* \frac{\partial T}{\partial t} = \nabla(k \nabla T) + Q_n \tag{C.7}$$

This 2D axisymmetric model for the prediction of the temperature profiles in the product and vial wall was here solved numerically in COMSOL Multiphysics (vers. 5.1).

C.2 Primary Drying

In order to assess the effect of the ice crystal size on drying time t_d and maximum temperature reached by the product T_{max}, the 1-dimensional model described in [327] was used, as will be summarized in the following. The heat flux to the product during primary drying, J_q, is proportional to the temperature difference between the heating fluid T_{fluid} and the product at the bottom of the container T_b,

$$J_q = K_v (T_{fluid} - T_b) \tag{C.8}$$

where K_v is an overall heat transfer coefficient.

At the interface of sublimation, the heat flux due to water vapor removal is given by,

$$J_q = \Delta H_s J_w \tag{C.9}$$

where ΔH_s is the heat of sublimation. Moreover, the vapor flux removed by sublimation J_w is proportional to the difference between the vapor pressure at the interface of sublimation $p_{w,i}$ and the partial pressure of water vapor in the drying chamber $p_{w,c}$,

$$J_w = \frac{1}{R_p}(p_{w,i} - p_{w,c}) \tag{C.10}$$

In this last equation, $p_{w,i}$ is a function of the product temperature at the interface of sublimation T_i, while $p_{w,c}$ can be considered to be constant and equal to the chamber pressure P_c. This last approximation is justified by the fact that during primary drying the gas phase inside the chamber is mainly composed of water [346]. R_p represents the mass transfer resistance of the dried product. The value of R_p is linked to the mean crystal size D_p of the product by the following expression [284],

$$R_p = \frac{3}{2}\frac{\tau^2}{\varepsilon}\frac{L_d}{D_p}\sqrt{\frac{\pi \mathrm{R} T}{2M_w}} \tag{C.11}$$

Thus, R_p describes the effect of the ice crystal size obtained during freezing, computed using the model described in Chapter 6, on drying performances. In other words, it allows connection of the freezing and primary drying models. The evolution of the dried cake thickness L_d can be finally quantified using the following equation,

$$\frac{dL_d}{dt} = \frac{1}{\varrho_f - \varrho_d} J_w \tag{C.12}$$

where ϱ_f and ϱ_d are the density of the frozen product and the dried cake, respectively. The time t_d at which L_d becomes equal to the sample height is called drying time. By solving Equations C.8 to C.12, it is possible to estimate the maximum temperature reached within the product T_{max} and the drying time t_d and, thus, to evaluate the effects of the ice crystal size on drying performances. This primary drying model was implemented in Matlab and solved using the finite differences approach, with a 60 s timestep.

List of Symbols

c_p	J kg^{-1} K^{-1}	specific heat capacity
c_p^*	J kg^{-1} K^{-1}	apparent specific heat capacity
$c_{p,ice}$	J kg^{-1} K^{-1}	specific heat capacity of ice
$c_{p,l}$	J kg^{-1} K^{-1}	specific heat capacity of the liquid
D_p	m	pores diameter
ΔH_f	J kg^{-1}	latent heat of crystallization
ΔH_s	J kg^{-1}	latent heat of sublimation

J_q	J s^{-1} m^{-2}	heat flux during primary drying
J_w	kg s^{-1} m^{-2}	mass flux during primary drying
k	W m^{-1} K^{-1}	thermal conductivity
k_l	W m^{-1} K^{-1}	thermal conductivity of the liquid
k_n	kg s^{-1} m^{-3} K^{-1}	kinetic constant for nucleation
K_v	W m^{-2} K^{-1}	heat transfer coefficient during drying
L_d	m	dried cake thickness
M_w	kg mol^{-1}	molar mass of water
P_c	Pa	chamber pressure
$p_{w,c}$	Pa	vapor pressure in the drying chamber
$p_{w,i}$	Pa	vapor pressure at the sublimation interface
Q_n	J s^{-1} m^{-3}	source term representing the generation due to nucleation
Q_c	J s^{-1} m^{-3}	source term representing the generation due to crystal growth
R	J mol^{-1} K^{-1}	universal gas constant
R_p	m s^{-1}	mass transfer resistance
t	s	time
t_d	s	drying time
T	K	temperature
T_b	K	temperature at the bottom of the product
T_{eq}	K	equilibrium temperature of freezing
T_{fluid}	K	fluid temperature
T_i	K	temperature at the sublimation interface
T^*	K	temperature of the supercooled liquid
T_{max}	K	maximum temperature during drying
ΔT	K	temperature difference
X_{ice}	-	ice fraction
ΔX_{ice}	-	change in ice fraction

Greek Letters

ε	-	porosity
ϱ	kg m^{-3}	density
ϱ_d	kg m^{-3}	density of the dried cake
ϱ_f	kg m^{-3}	density of the frozen product
ϱ_{ice}	kg m^{-3}	density of ice
ϱ_l	kg m^{-3}	density of the liquid
τ	-	tortuosity

Bibliography

[1] Evaluate Pharma, World preview 2016, outlook to 2022, London: Evaluate Ltd. http://info.evaluategroup.com/ rs/607-YGS-364/images/wp16.pdf (2016).

[2] EU Commission, The 2015 EU industrial R&D investment scoreboard, http://iri.jrc.ec.europa.eu/scoreboard15.html (2015).

[3] International Federation of Pharmaceutical Manufacturers & Associations, The pharmaceutical industry and global health, facts and figures 2017, https://www.ifpma.org/resource-centre/ifpma-launches-its-2017-facts-figures-report (2017).

[4] PhRMA, Chart pack biopharmaceuticals in perspective, Washington DC: Pharmaceutical Research and Manufacturers of America. http://phrma.org/sites/default/files/pdf/chart-pack-biopharmaceuticals-in-perspective.pdf (2016).

[5] J. A. DiMasi, H. G. Grabowski, R. W. Hansen, Innovation in the pharmaceutical industry: New estimates of R&D costs, J. Health Econ. 47 (2016) 20–33.

[6] PhRMA, Biopharmaceutical research industry profile, Washington DC: Pharmaceutical Research and Manufacturers of America. http://phrma.org/sites/default/files/pdf/biopharmaceutical-industry-profile.pdf (2016).

[7] P. Troein, Market development in Europe and globally-A perspective on biologicals and biosimilars, 15th Biosimilar Medicines Conference (Medicines for Europe), London: QuintilesIMS.

[8] E. Moorkens, N. Meuwissen, I. Huys, P. Declerck, A. G. Vulto, S. Simoens, The market of biopharmaceutical medicines: A snapshot of a diverse industrial landscape, Front. Pharmacol. 8 (2017) 314.

[9] MordorIntelligence, Biopharmaceuticals market - growth, trends, and forecast (2019 - 2024), https://www.mordorintelligence.com/industry-reports/global-biopharmaceuticals-market-industry (2018).

[10] R. Pisano, A. Arsiccio, Formulation design and optimization using molecular dynamics, in: F. Davide, R. Pisano, A. Barresi (Eds.), Freeze Drying of Pharmaceutical Products, CRC Press (Taylor & Francis), Boca Raton, Florida,

U.S.A., In press, pp. 11–31.
[11] W. Wang, C. J. Roberts, Aggregation of Therapeutic Proteins, John Wiley & Sons, Hoboken, New Jersey, 2010, 486 pp.
[12] C. J. Roberts, Therapeutic protein aggregation: Mechanisms, design, and control, Trends Biotechnol. 32 (2014) 372–380.
[13] W. Wang, Lyophilization and development of solid protein pharmaceuticals, Int. J. Pharm. 203 (2000) 1–60.
[14] W. Jiskoot, T. W. Randolph, D. B. Volkin, C. R. Middaugh, C. Schöneich, G. Winter, W. Friess, D. J. Crommelin, C. J. F, Protein instability and immunogenicity: Roadblocks to clinical application of injectable protein delivery systems for sustained release, J. Pharm. Sci. 101 (2012) 946–954.
[15] P. L. Privalov, Cold denaturation of proteins, Crit. Rev. Biochem. Mol. Biol. 25 (1990) 281–305.
[16] F. Franks, Protein destailization at low temperatures, Adv. Protein Chem. 46 (1995) 105–139.
[17] G. Graziano, F. Catanzano, A. Riccio, G. Barone, A reassessment of the molecular origin of cold denaturation, J. Biochem. 122 (1997) 395–401.
[18] C. F. Lopez, R. K. Darst, P. J. Rossky, Mechanistic elements of protein cold denaturation, J. Phys. Chem. B 112 (2008) 5961–5967.
[19] S. Matysiak, P. G. Debenedetti, P. J. Rossky, Role of hydrophobic hydration in protein stability: A 3D water-explicit protein model exhibiting cold and heat denaturation, J. Phys. Chem. B 116 (2012) 8095–8104.
[20] G. B. Strambini, E. Gabellieri, Proteins in frozen solutions: Evidence of ice-induced partial unfolding, Biophys. J. 70 (1996) 971–976.
[21] B. S. Chang, B. S. Kendrick, J. F. Carpenter, Surface-induced denaturation of proteins during freezing and its inhibition by surfactants, J. Pharm. Sci. 85 (1996) 1325–1330.
[22] W. Liu, D. Q. Wang, S. L. Nail, Freeze-drying of proteins from a sucrose-glycine excipient system: Effect of formulation composition on the initial recovery of protein activity, AAPS PharmSciTech 6 (2005) 150–157.
[23] A. Al-Hussein, H. Gieseler, The effect of mannitol crystallization in mannitol-sucrose systems on LDH stability during freeze-drying, J. Pharm. Sci. 101 (2012) 2534–2544.
[24] M. J. Pikal, Mechanisms of protein stabilization during freeze-drying storage: The relative importance of thermodynamic stabilization and glassy state relaxation dynamics, In: "Freeze-Drying/Lyophilization of Pharmaceutical and Biological Products", Informa Healthcare, London, 2010, pp. 198–232.
[25] L. van den Berg, D. Rose, The effect of freezing on the pH and composition of sodium and potassium solutions: The reciprocal system KH_2PO_4-Na_2HPO_4-H_2O, Arch. Biochem. Biophys. 81 (1959) 319–329.
[26] T. J. Anchordoquy, J. F. Carpenter, Polymers protect lactate dehydrogenase during freeze-drying by inhibiting dissociation in the frozen state, Arch.

Biochem. Biophys. 332 (1996) 231–238.
[27] G. Gomez, M. J. Pikal, N. Rodríguez-Hornedo, Effect of initial buffer composition on pH changes during far-from-equilibrium freezing of sodium phosphate buffer solutions, Pharm. Res. 18 (2001) 90–97.
[28] N. Murase, F. Franks, Salt precipitation during the freeze-concentration of phosphate buffer solutions, Biophys. Chem. 34 (1989) 293–300.
[29] M. C. Heller, J. F. Carpenter, T. W. Randolph, Manipulation of lyophilization induced phase separation: Implications for pharmaceutical proteins, Biotechnol. Prog. 13 (1997) 590–596.
[30] J. A. Rupley, G. Careri, Protein hydration and function, Adv. Protein Chem. 41 (1991) 37–172.
[31] S. Ohtake, Y. Kita, T. Arakawa, Interactions of formulation excipients with proteins in solution and in the dried state, Adv. Drug Deliv. Rev. 63 (2011) 1053–1073.
[32] S. Timasheff, The control of protein stability and association by weak interactions with water: How do solvents affect these processes?, Ann. Rev. Biophys. Biomol. Struct. 22 (1993) 67–97.
[33] T. Arakawa, S. N. Timasheff, Stabilization of protein structure by sugars, Biochemistry 21 (1982) 6536–6544.
[34] J. G. Kirkwood, F. P. Buff, The statistical mechanical theory of solutions. I, J. Chem. Phys. 19 (1951) 774.
[35] A. Ben-Naim, Molecular Theory of Solutions, Oxford University Press, New York, 2006.
[36] Y. Mi, G. Wood, L. Thoma, Cryoprotection mechanisms of polyethylene glycols on lactate dehydrogenase during freeze-thawing, AAPS J. 6 (2004) e22.
[37] J. F. Carpenter, S. J. Prestrelski, T. Arakawa, Separation of freezing and drying-induced denaturation of lyophilized proteins using stress-specific stabilization. I. Enzymatic activity and calorimetric studies., Arch. Biochem. Biophys. 303 (1993) 456–464.
[38] J. H. Crowe, L. M. Crowe, D. Chapman, Preservation of membranes in anhydrobiotic organisms: The role of trehalose, Science 223 (1984) 701–703.
[39] J. F. Carpenter, J. H. Crowe, An infrared spectroscopic study of the interactions of carbohydrates with dried proteins, Biochemistry 28 (1989) 3916–3922.
[40] J. F. Carpenter, J. H. Crowe, T. Arakawa, Comparison of solute-induced protein stabilization in aqueous solution and in frozen and dried state, J. Dairy Sci. 73 (1990) 3627–3636.
[41] J. L. Green, C. A. Angell, Phase relations and vitrification in saccharide-water solutions and the trehalose anomaly, J. Phys. Chem. 93 (1989) 2880–2882.
[42] F. Franks, Long-term stabilization of biologicals, Biotechnology 12 (1994) 253–256.
[43] S. J. Hagen, J. Hofrichter, W. A. Eaton, Protein reaction kinetics in a room-temperature glass, Science 269 (1995) 959–962.

[44] P. S. Belton, A. M. Gil, IR and Raman spectroscopic studies of the interaction of trehalose with hen egg white lysozyme, Biopolymers 34 (1994) 957–961.

[45] D. Corradini, E. G. Strekalova, H. E. Stanley, P. Gallo, Microscopic mechanism of protein cryopreservation in an aqueous solution with trehalose, Sci. Rep. 3 (2013) 1218.

[46] S. Corezzi, M. Paolantoni, P. Sassi, A. Morresi, D. Fioretto, C. L, Trehalose-induced slowdown of lysozyme hydration dynamics probed by EDLS spectroscopy, J. Chem. Phys. 151 (2019) 015101.

[47] H. J. Lee, A. McAuley, K. F. Schilke, J. McGuire, Molecular origins of surfactant-mediated stabilization of protein drugs, Adv. Drug. Deliv. Rev. 63 (2011) 1160–1171.

[48] T. W. Randolph, L. S. Jones, Surfactant-Protein Interactions, In: "Rational Design of Stable Protein Formulations: Theory and Practice", Kluwer Academic, New York, 2002, pp. 159–175.

[49] N. B. Bam, J. L. Cleland, T. W. Randolph, Molten globule intermediate of recombinant human growth hormone: Stabilization with surfactants, Biotechnol. Prog. 12 (1996) 801–809.

[50] N. B. Bam, T. W. Randolph, J. L. Cleland, Stability of protein formulations: Investigation of surfactant effects by a novel EPR spectroscopic technique, Pharm. Res. 12 (1995) 2–11.

[51] N. B. Bam, J. L. Cleland, J. Yang, M. C. Manning, J. F. Carpenter, R. F. Kelley, T. W. Randolph, Tween protects recombinant human growth hormone against agitation-induced damage via hydrophobic interactions, J. Pharm. Sci. 87 (1998) 1554–1559.

[52] S. Deechongkit, J. Wen, L. O. Narhi, Y. Jiang, S. S. Park, J. Kim, B. A. Kerwin, Physical and biophysical effects of polysorbate 20 and 80 on darbepoetin alfa, J. Pharm. Sci. 98 (2009) 3200–3217.

[53] L. S. Jones, T. W. Randolph, U. Kohnert, A. Papadimitriou, G. Winter, M. L. Hagmann, M. C. Manning, J. F. Carpenter, The effects of Tween 20 and sucrose on the stability of anti-L-selectin during lyophilization and reconstitution, J. Pharm. Sci. 90 (2001) 1466–1477.

[54] J. F. Carpenter, M. J. Pikal, B. S. Chang, T. W. Randolph, Rational design of stable lyophilized protein formulations: Some practical advice, Pharm. Res. 14 (1997) 969–975.

[55] S. Politis, P. Colombo, G. Colombo, D. M. Rekkas, Design of experiments (DoE) in pharmaceutical development, Drug. Dev. Ind. Pharm. 43 (2017) 889–901.

[56] Y. Grant, P. Matejtschuk, C. Bird, M. Wadhwa, P. A. Dalby, Freeze drying formulation using microscale and design of experiment approaches: A case study using granulocyte colony-stimulating factor, Biotechnol. Lett. 34 (2012) 641–648.

[57] A. Hottot, S. Vessot, J. Andrieu, Freeze-drying of pharmaceuticals in vials:

Influence of freezing protocol and sample configuration on ice morphology and freeze-dried cake texture, Chem. Eng. Process. 46 (2007) 666–674.

[58] J. C. Kasper, W. F. Friess, The freezing step in lyophilization: Physico-chemical fundamentals, freezing methods and consequences on process performance and quality attributes of biopharmaceuticals, Eur. J. Pharm. Biopharm. 78 (2011) 248–263.

[59] J. A. Searles, J. F. Carpenter, T. W. Randolph, The ice nucleation temperature determines the primary drying rate of lyophilization for samples frozen on a temperature-controlled shelf, J. Pharm. Sci. 90 (2001) 860–871.

[60] L. C. Capozzi, R. Pisano, Looking inside the 'black box': Freezing engineering to ensure the quality of freeze-dried biopharmaceuticals., Eur. J. Pharm. Biopharm. 129 (2018) 58–65.

[61] B. S. Bhatnagar, M. J. Pikal, H. B. Robin, Study of the individual contributions of ice formation and freeze-concentration on isothermal stability of lactate dehydrogenase during freezing., J. Pharm. Sci. 97 (2008) 798–814.

[62] R. Pisano, Alternative methods of controlling nucleation in freeze drying, in: K. R. Ward, P. Matejtschuk (Eds.), Lyophilization of Pharmaceuticals and Biologicals: New Technologies and Approaches, Springer New York, New York, NY, 2019, pp. 79–111.

[63] R. Geidobler, G. Winter, Controlled ice nucleation in the field of freeze-drying: Fundamentals and technology review, Eur. J. Pharm. Biopharm. 85 (2013) 214–222.

[64] R. Pisano, A. Arsiccio, K. Nakagawa, A. A. Barresi, Tuning, measurement and prediction of the impact of freezing on product morphology: A step toward improved design of freeze-drying cycles, Dry. Technol. 37 (2019) 579–599.

[65] R. Geidobler, I. Konrad, G. Winter, Can controlled ice nucleation improve freeze-drying of highly-concentrated protein formulations?, J. Pharm. Sci. 102 (2013) 3915 – 3919.

[66] R. Fang, K. Tanaka, V. Mudhivarthi, R. H. Bogner, M. J. Pikal, Effect of controlled ice nucleation on stability of lactate dehydrogenase during freeze-drying, J. Pharm. Sci. 107 (2018) 824–830.

[67] S. N. Singh, S. Kumar, V. Bondar, N. Wang, R. Forcino, J. Colandene, D. Nesta, Unexplored benefits of controlled ice nucleation: Lyophilization of a highly concentrated monoclonal antibody solution, Int. J. Pharm. 552 (2018) 171 – 179.

[68] F. Franks, Solid aqueous solutions, Pure Appl. Chem. 65 (1993) 2527–2537.

[69] B. S. Chang, C. S. Randall, Use of subambient thermal analysis to optimize protein lyophilization, Cryobiology 29 (1992) 632–656.

[70] C. C. Hsu, A. J. Walsh, H. M. Nguyen, D. E. Overcashier, H. Koning-Bastiaan, R. C. Bailey, S. L. Nail, Design and application of a low-temperature peltier-cooling microscope stage, Chem. Eng. Process 85 (1996)

70–74.

[71] A. I. Kim, M. J. Akers, S. L. Nail, The physical state of mannitol after freeze-drying: Effect of mannitol concentration, freezing rate, and a noncrystallizing cosolute., J. Pharm. Sci. 87 (1998) 931–935.

[72] M. J. Akers, N. Milton, S. R. Byrn, S. L. Nail, Glycine crystallization during freezing: The effect of salt form, pH, and ionic strength, Pharm. Res. 12 (1995) 1457–1461.

[73] M. C. Heller, J. F. Carpenter, T. W. Randolph, Application of a thermodynamic model to the prediction of phase separations in freeze-concentrated formulations for protein lyophilization, Arch. Biochem. Biophys. 363 (1999) 191–201.

[74] K. Izutsu, S. Yoshioka, T. Terao, Effect of mannitol crystallinity on the stabilization of enzymes during freeze-drying, Chem. Pharm. Bull. (Tokyo) 42 (1994) 5–8.

[75] K. Izutsu, S. Yoshioka, T. Terao, Decreased protein-stabilizing effects of cryoprotectants due to crystallization, Pharm. Res. 10 (1993) 1232–1237.

[76] J. A. Searles, J. F. Carpenter, T. W. Randolph, Annealing to optimize the primary drying rate, reduce freezing-induced drying rate heterogeneity, and determine T'_g in pharmaceutical lyophilization, J. Pharm. Sci. 90 (2001) 872–887.

[77] D. Q. Wang, J. M. Hey, S. L. Nail, Effect of collapse on the stability of freeze-dried recombinant factor VIII and α-amylase, J. Pharm. Sci. 93 (5) (2004) 1253–1263.

[78] S. Luthra, J.-P. Obert, D. S. Kalonia, M. J. Pikal, Investigation of drying stresses on proteins during lyophilization: Differentiation between primary and secondary-drying stresses on lactate dehydrogenase using a humidity controlled mini freeze-dryer, J. Pharm. Sci. 96 (1) (2007) 61 – 70.

[79] K. Schersch, O. Betz, P. Garidel, S. Muehlau, S. Bassarab, G. Winter, Systematic investigation of the effect of lyophilizate collapse on pharmaceutically relevant proteins I: Stability after freeze-drying, J. Pharm. Sci. 99 (5) (2010) 2256 – 2278.

[80] L. A. Daukas, E. H. Trappler, Assessing the quality of lyophilized parenterals, Pharm. Cosmetic Quality 2 (1998) 21–25.

[81] S. Yoshioka, Y. Aso, S. Kojima, Dependence of the molecular mobility and protein stability of freeze-dried γ-globulin formulations on the molecular weight of dextran, Pharm. Res. 14 (1997) 736–741.

[82] I. Oddone, R. Pisano, R. Bullich, P. Stewart, Vacuum-induced nucleation as a method for freeze-drying cycle optimization, Ind. Eng. Chem. Res. 53 (2014) 18236–18244.

[83] I. Oddone, P.-J. Van Bockstal, T. De Beer, R. Pisano, Impact of vacuum-induced surface freezing on inter- and intra-vial heterogeneity, Eur. J. Pharm. Biopharm. 103 (2016) 167–178.

[84] I. Oddone, A. A. Barresi, R. Pisano, Influence of controlled ice nucleation on the freeze-drying of pharmaceutical products: The secondary drying step, Int. J. Pharm. 524 (2017) 134–140.

[85] U.S. Food and Drug Administration, Department of Health and Human Services, Pharmaceutical CGMPs for the 21st Century: A Risk-Based Approach, Silver Spring, MD, September 2004.

[86] L. X. Yu, Pharmaceutical quality by design: Product and process development, understanding, and control., Pharm. Res. 25 (2008) 781–791.

[87] L. X. Yu, G. Amidon, M. A. Khan, S. W. Hoag, J. Polli, G. K. Raju, J. Woodcock, Understanding pharmaceutical quality by design, AAPS J. 16 (2014) 771–783.

[88] D. Frenkel, B. Smit, Understanding Molecular Simulation, Academic., New York, 2002.

[89] M. P. Allen, D. J. Tildesley, Computer Simulation of Liquids, Oxford University Press., Oxford, 2017.

[90] P. K. Weiner, P. A. Kollman, AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions., J. Comput. Chem. 2 (1981) 287–303.

[91] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, J. Phys. Chem. B 102 (1998) 3586–3616.

[92] W. R. P. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger, W. F. van Gunsteren, The GROMOS biomolecular simulation program package, J. Phys. Chem. A 103 (1999) 3596–3607.

[93] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, J. Am. Chem. Soc. 118 (1996) 11225–11236.

[94] D. Shukla, C. Shinde, B. L. Trout, Molecular computations of preferential interaction coefficients of proteins, J. Phys. Chem. B 113 (2009) 12546–12554.

[95] P. Ganguly, N. F. A. van der Vegt, Convergence of sampling kirkwood-buff integrals of aqueous solutions with molecular dynamics simulations, J. Chem. Theory Comput. 9 (2013) 1347–1355.

[96] J. Barnoud, L. Monticelli, Coarse-grained force fields for molecular simulations, Methods Mol. Biol. 1215 (2015) 125–149.

[97] C. Abrams, G. Bussi, Enhanced sampling in molecular dynamics using metadynamics, replica-exchange, and temperature-acceleration, Entropy 16 (2014) 163–199.

[98] O. Valsson, P. Tiwary, M. Parrinello, Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint, Ann. Rev. Phys. Chem. 67 (2016) 159–184.
[99] U. H. Hansmann, Parallel tempering algorithm for conformational studies of biological molecules, Chem. Phys. Lett. 281 (1997) 140–150.
[100] Y. Sugita, Y. Okamoto, Replica-exchange molecular dynamics method for protein folding, Chem. Phys. Lett. 314 (1999) 141–151.
[101] B. M. Eckhardt, J. Q. Oeswein, T. A. Bewley, Effect of freezing on aggregation of human growth hormone, Pharm. Res. 8 (1991) 1360–1364.
[102] M. J. Pikal, K. M. Dellerman, M. L. Roy, R. M. Riggin, The effects of formulation variables on the stability of freeze-dried human growth hormone, Pharm. Res. 8 (1991) 427–436.
[103] H. R. Costantino, K. G. Carrasquillo, R. A. Cordero, M. Mumenthaler, C. C. Hsu, K. Griebenow, Effect of excipients on the stability and structure of lyophilized recombinant human growth hormone, J. Pharm. Sci. 87 (1998) 1412–1420.
[104] A. M. Abdul-Fattah, D. Lechuga-Ballesteros, D. S. Kalonia, M. J. Pikal, The impact of drying method and formulation on the physical properties and stability of methionyl human growth hormone in the amorphous solid state, Biotechnology 97 (2008) 163–184.
[105] M. S. Salnikova, C. R. Russell Middaugh, J. H. Rytting, Stability of lyophilized human growth hormone, Int. J. Pharm. 358 (2008) 108–113.
[106] M. R. Kasimova, S. J. Milstein, E. Freire, The conformational equilibrium of human growth hormone, J. Mol. Biol. 277 (1998) 409–418.
[107] M. R. DeFelippis, M. A. Kilcomons, M. P. Lents, K. M. Youngman, H. A. Havel, Acid stabilization of human growth hormone equilibrium folding intermediates, Biochim. Biophys. Acta 1247 (1995) 35–45.
[108] M. R. DeFelippis, L. A. Alter, A. H. Pekar, H. A. Havel, D. N. Brems, Evidence for a self-associating equilibrium intermediate during folding of human growth hormone, Biochemistry 32 (1993) 1555–1562.
[109] M. M. Andersson, R. Hatti-Kaul, Protein stabilising effect of polyethyleneimine, J. Biotechnol. 72 (1999) 21–31.
[110] K. Izutsu, S. Yoshioka, S. Kojima, Increased stabilizing effects of amphiphilic excipients on freeze-drying of lactate dehydrogenase (LDH) by dispersion into sugar matrices, Pharm. Res. 12 (1995) 838–843.
[111] S. J. Prestrelski, T. Arakawa, J. F. Carpenter, Separation of freezing- and drying-induced denaturation of lyophilized proteins using stress-specific stabilization. II. Structural studies using infrared spectroscopy, Arch. Biochem. Biophys. 303 (1993) 465–473.
[112] S. Jiang, S. L. Nail, Effect of process conditions on recovery of protein activity after freezing and freeze-drying, Eur. J. Pharm. Biopharm. 45 (1998) 249–257.

[113] K. Kawai, T. Suzuki, Stabilizing effect of four types of disaccharide on the enzymatic activity of freeze-dried lactate dehydrogenase: Step by step evaluation from freezing to storage, Pharm. Res. 24 (2007) 1883–1890.

[114] S. Nema, K. E. Avis, Freeze-thaw studies of a model protein, lactate dehydrogenase, in the presence of cryoprotectants, J. Parenter. Sci. Technol. 47 (1992) 76–83.

[115] A. Arsiccio, R. Pisano, Stability of proteins in carbohydrates and other additives during freezing: The human growth hormone as a case study, J. Phys. Chem. B 121 (2017) 8652–8660.

[116] A. Arsiccio, R. Pisano, Water entrapment and structure ordering as protection mechanisms for protein structural preservation, J. Chem. Phys. 148 (2018) 055102.

[117] A. Arsiccio, A. Paladini, F. Pattarino, R. Pisano, Designing the optimal formulation for biopharmaceuticals: A new approach combining molecular dynamics and experiments, J. Pharm. Sci. 108 (2019) 431–438.

[118] A. Arsiccio, R. Pisano, The preservation of lyophilized human growth hormone activity: How do buffers and sugars interact?, Pharm. Res. 35 (2018) 131.

[119] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, GROMACS: Fast, flexible, free, J. Comput. Chem. 26 (2005) 1701–1718.

[120] A. M. de Vos, M. Ultsch, A. A. Kossiakoff, Human growth hormone and extracellular domain of its receptor: Crystal structure of the complex, Science 255 (1992) 306–312.

[121] B. A. Beaupre, J. V. Roman, M. R. Hoag, K. M. Meneely, N. R. Silvaggi, A. L. Lamb, G. R. Moran, Ligand binding phenomena that pertain to the metabolic function of renalase, Arch. Biochem. Biophys. 612 (2016) 46–56.

[122] A. K. Malde, L. Zuo, M. Breeze, M. Stroet, D. Poger, P. C. Nair, C. Oostenbrink, A. E. Mark, An automated force field topology builder (ATB) and repository: Version 1.0, J. Chem. Theory Comput. 7 (2011) 4026–4037.

[123] N. Schmid, A. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. Mark, W. van Gunsteren, Definition and testing of the GROMOS force-field versions 54A7 and 54B7, Eur. Biophys. J. 40 (2011) 843–856.

[124] H. J. C. Berendsen, J. R. Grigera, T. P. J. Straatsma, The missing term in effective pair potentials, J. Phys. Chem. 91 (1987) 6269–6271.

[125] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity-rescaling, J. Chem. Phys. 126 (2007) 014101.

[126] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, A smooth particle mesh ewald method, J. Chem. Phys. 103 (1995) 8577–8593.

[127] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, J. Appl. Phys. 52 (1981) 7182–7190.

[128] B. Hess, Determining the shear viscosity of model liquids from molecular dynamics simulations, J. Chem. Phys. 116 (2002) 209–217.
[129] T. Arakawa, Y. Kita, J. F. Carpenter, Protein-solvent interactions in pharmaceutical formulations, Pharm. Res. 8 (1991) 285–291.
[130] T. Y. Lin, S. N. Timasheff, On the role of surface tension in the stabilization of globular proteins, Protein Sci. 5 (1996) 372–381.
[131] J. H. Crowe, L. M. Crowe, J. F. Carpenter, Preserving dry biomaterials: The water replacement hypothesis, part I, Biopharm 6 (1993) 28–33.
[132] J. H. Crowe, L. M. Crowe, J. F. Carpenter, Preserving dry biomaterials: The water replacement hypothesis, part II, Biopharm 6 (1993) 40–43.
[133] J. Genotelle, Expression of the viscosity of sugar solutions, Ind. Alim. Agric. 95 (1978) 747–755.
[134] M. P. Longinotti, H. R. Corti, Viscosity of concentrated sucrose and trehalose aqueous solutions including the supercooled regime, J. Phys. Chem. Ref. Data 37 (2008) 1503–1515.
[135] S. Yoshioka, Y. Aso, Correlations between molecular mobility and chemical stability during storage of amorphous pharmaceuticals, J. Pharm. Sci. 96 (2007) 960–981.
[136] S. Rossi, M. P. Buera, S. Moreno, J. Chirife, Stabilization of the restriction enzyme EcoRI dried with trehalose and other selected glass-forming solutes, Biotechnol. Prog. 13 (1997) 609–616.
[137] P. Davidson, W. Q. Sun, Effect of sucrose/raffinose mass ratios on the stability of co-lyophilized protein during storage above the T_g, Pharm. Res. 18 (2001) 474–479.
[138] B. Wang, S. Tchessalov, M. T. Cicerone, N. W. Warne, M. J. Pikal, Impact of sucrose level on storage stability of proteins in freeze-dried solids: II. Correlation of aggregation rate with protein structure and molecular mobility, J. Pharm. Sci. 98 (2009) 3145–3166.
[139] M. T. Cicerone, J. F. Douglas, β-relaxation governs protein stability in sugar-glass matrices, Soft Matter 8 (2012) 2983–2991.
[140] M. T. Cicerone, M. J. Pikal, K. K. Qian, Stabilization of proteins in solid form, Adv. Drug. Deliv. Rev. 93 (2015) 14–24.
[141] M. T. Cicerone, C. L. Soles, Fast dynamics and stabilization of proteins: Binary glasses of trehalose and glycerol, Biophys. J. 86 (2004) 3836–3845.
[142] M.-C. Bellissent-Funel, J. Teixeira, Structural and dynamic properties of bulk and confined water, in: L. Rey, J. C. May (Eds.), Freeze-Drying/Lyophilization of Pharmaceutical and Biological Products, Marcel Dekker, New York, 1999, pp. 29–51.
[143] C. M. Sudrik, T. Cloutier, N. Mody, H. A. Sathish, B. L. Trout, Understanding the role of preferential exclusion of sugars and polyols from native state IgG1 monoclonal antibodies and its effect on aggregation and reversible self-association, Pharm. Res. 36 (2019) 109.

[144] T. J. Zbacnik, R. E. Holcomb, D. S. Katayama, B. M. Murphy, R. W. Payne, R. C. Coccaro, G. J. Evans, J. E. Matsuura, C. S. Henry, M. C. Manning, Role of buffers in protein formulations, J. Pharm. Sci. 106 (2017) 713 – 733.

[145] L. Weng, G. D. Elliott, Distinctly different glass transition behaviors of trehalose mixed with Na_2HPO_4 or NaH_2PO_4: Evidence for its molecular origin, Pharm. Res. 32 (2015) 2217–2228.

[146] I. Skarmoutsos, E. Guardia, J. Samios, Hydrogen bond, electron donor-acceptor dimer, and residence dynamics in supercritical CO_2-ethanol mixtures and the effect of hydrogen bonding on single reorientational and translational dynamics: A molecular dynamics simulation study, J. Chem. Phys. 133 (2010) 014504.

[147] S. Ohtake, C. Schebor, S. P. Palecek, J. J. de Pablo, Effect of pH, counter ion, and phosphate concentration on the glass transition temperature of freeze-dried sugar-phosphate mixtures, Pharm. Res. 21 (2004) 1615–1621.

[148] F. Franks, T. Auffret, Freeze-Drying of Pharmaceuticals and Biopharmaceuticals, Royal Society of Chemistry, 2008.

[149] E. P. Kets, P. J. IJpelaar, F. A. Hoekstra, V. H., Citrate increases glass transition temperature of vitrified sucrose preparations, Cryobiology 48 (2004) 46–54.

[150] J. K. Kaushik, R. Bhat, A mechanistic analysis of the increase in the thermal stability of proteins in aqueous carboxylic acid salt solutions, Protein Sci. 8 (1999) 222–233.

[151] R. Gillespie, T. Nguyen, S. Macneil, L. Jones, S. Crampton, S. Vunnum, Cation exchange surface-mediated denaturation of an aglycosylated immunoglobulin (IgG1), J. Chromatogr. A 1251 (2012) 101–110.

[152] B. L. Chen, T. Arakawa, E. Hsu, L. O. Narhi, T. J. Tressel, S. L. Chien, Strategies to suppress aggregation of recombinant keratinocyte growth-factor during liquid formulation development, J. Pharm. Sci. 83 (1994) 1657–1661.

[153] L. C. D. Novaes, P. G. Mazzola, A. Pessoa, P. T. C. V, Citrate and phosphate influence on green fluorescent protein thermal stability, Biotechnol. Prog. 27 (2011) 269–272.

[154] K. Mizutani, Y. Chen, H. Yamashita, M. Hirose, A. S, Thermostabilization of ovotransferrin by anions for pasteurization of liquid egg white, Biosci. Biotechnol. Biochem. 70 (2006) 1839–1845.

[155] X. M. Cao, Z. Y. Wang, Y. W. Liu, C. X. Wang, T. Y, Effect of additive on the thermal denaturation of lysozyme analyzed by isoconversional method, Acta Chim. Sinica 68 (2010) 194–198.

[156] V. Joshi, T. Shivach, V. Kumar, N. Yadav, A. Rathore, Avoiding antibody aggregation during processing: establishing hold times, Biotechnol. J. 9 (2014) 1195–1205.

[157] N. Harnkarnsujarit, M. KNakajima, K. Kawai, M. Watanabe, T. Suzuki, Thermal properties of freeze-concentrated sugar-phosphate solutions, Food

Biophys. 9 (2014) 213–218.
[158] K. Kawai, T. Suzuki, Effect of tetrasodium tripolyphosphate on the freeze-concentrated glass-like transition temperature of sugar aqueous solutions, CryoLetters 27 (2006) 107–114.
[159] K. Izutsu, S. Kadoya, C. Yomota, T. Kawanishi, E. Yonemachi, K. Terada, Stabilization of protein structure in freeze-dried amorphous organic acid buffer salts, Chem. Pharm. Bull. 57 (2009) 1231–1236.
[160] K. Izutsu, S. Kadoya, C. Yomota, T. Kawanishi, E. Yonemachi, K. Terada, Freeze-drying of proteins in glass solids formed by basic amino acids and dicarboxylic acids, Chem. Pharm. Bull. 57 (2009) 43–48.
[161] W. F. Wolkers, H. Oldendorf, F. Tablin, J. H. Crowe, Preservation of dried liposomes in the presence of sugar and phosphate, Biochim. Biophys. Acta 1661 (2004) 125–134.
[162] Y. Goto, I. J. Calciano, A. L. Fink, Acid-induced folding of proteins, Proc. Natl. Acad. Sci. U.S.A. 87 (1990) 573–577.
[163] G. H. Zerze, R. G. Mullen, Z. A. Levine, J.-E. Shea, J. Mittal, To what extent does surface hydrophobicity dictate peptide folding and stability near surfaces?, Langmuir 31 (2015) 12223–12230.
[164] A. E. Garcia, K. Y. Sanbonmatsu, Exploring the energy landscape of a β hairpin in explicit solvent, Proteins 42 (2001) 345–354.
[165] V. Muñoz, P. A. Thompson, J. Hofrichter, W. A. Eaton, Folding dynamics and mechanism of β-hairpin formation, Nature 390 (1997) 196–199.
[166] D. De Sancho, J. Mittal, R. B. Best, Folding kinetics and unfolded state dynamics of the GB1 hairpin from molecular simulation, J. Chem. Theory Comput. 9 (2013) 1743–1753.
[167] R. B. Best, J. Mittal, Microscopic events in β-hairpin folding from alternative unfolded ensembles, Proc. Natl. Acad. Sci. 108 (2011) 11087–11092.
[168] R. B. Best, J. Mittal, Free-energy landscape of the GB1 hairpin in all-atom explicit solvent simulations with different force fields: Similarities and differences, Proteins 79 (2010) 1318–1328.
[169] V. S. Pande, D. S. Rokhsar, Molecular dynamics simulations of unfolding and refolding of a β-hairpin fragment of protein g, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 9062–9067.
[170] R. Zhou, B. J. Berne, R. Germain, The free energy landscape for β hairpin folding in explicit water, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 14931–14936.
[171] P. G. Bolhuis, Transition-path sampling of β-hairpin folding, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 12129–12134.
[172] D. Paschek, A. E. García, Reversible temperature and pressure denaturation of a protein fragment: A replica exchange molecular dynamics simulation study, Phys. Rev. Lett. 93 (2004) 238105.
[173] V. Muñoz, E. R. Henry, J. Hofrichter, W. A. Eaton, A statistical mechanical

model for β-hairpin kinetics, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 5872–5879.

[174] S. Honda, N. Kobayashi, E. Munekata, Thermodynamics of a β-hairpin structure: Evidence for cooperative formation of folding nucleus, J. Mol. Biol. 295 (2000) 269 – 278.

[175] D. Du, Y. Zhu, C.-Y. Huang, F. Gai, Understanding the key factors that control the rate of β-hairpin folding, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 15915–15920.

[176] D. K. Chou, R. Krishnamurthy, T. W. Randolph, J. F. Carpenter, M. C. Manning, Effects of Tween 20® and Tween 80® on the stability of albutropin during agitation, J. Pharm. Sci. 94 (2005) 1368 – 1381.

[177] P. Garidel, C. Hoffmann, A. Blume, A thermodynamic analysis of the binding interaction between polysorbate 20 and 80 with human serum albumins and immunoglobulins: A contribution to understand colloidal protein stabilisation, Byophis. Chem. 143 (2009) 70–78.

[178] A. Arsiccio, R. Pisano, Surfactants as stabilizers for biopharmaceuticals: An insight into the molecular mechanisms for inhibition of protein aggregation, Eur. J. Pharm. Biopharm. 128 (2018) 98–106.

[179] A. Arsiccio, J. McCarty, R. Pisano, J.-E. Shea, Effect of surfactants on surface-induced denaturation of proteins: Evidence of an orientation-dependent mechanism, J. Phys. Chem. B 122 (2018) 11390–11399.

[180] B. S. Kendrick, T. Li, B. S. Chang, Physical stabilization of proteins in aqueous solution, Springer, Boston, MA, 2002, pp. 61–84.

[181] L. Liu, W. Qi, D. K. Schwartz, T. W. Randolph, J. F. Carpenter, The effects of excipients on protein aggregation during agitation: An interfacial shear rheology study, J. Pharm. Sci. 102 (2013) 2460 – 2470.

[182] L. Krielgaard, L. S. Jones, T. W. Randolph, S. Frokjaer, J. M. Flink, M. C. Manning, J. F. Carpenter, Effect of Tween 20 on freeze-thawing- and agitation-induced aggregation of recombinant human factor XIII, J. Pharm. Sci. 87 (1998) 1597 – 1603.

[183] J. Möller, M. A. Schroer, M. Erlkamp, S. Grobelny, M. Paulus, S. Tiemeyer, F. J. Wirkert, M. Tolan, R. Winter, The effect of ionic strength, temperature, and pressure on the interaction potential of dense protein solutions: From nonlinear pressure response to protein crystallization, Biophys. J. 102 (2012) 2641–2648.

[184] G. M. Torrie, J. P. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, J. Comput. Phys. 23 (1977) 187–199.

[185] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, P. A. Kollman, The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method., J. Comput. Chem. 13 (1992) 1011 – 1021.

[186] J. S. Hub, B. L. de Groot, D. van der Spoel, g_wham - A free weighted

histogram analysis implementation including robust error and autocorrelation estimates, J. Chem. Theory Comput. 6 (2010) 3713–3720.

[187] X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, A. E. Mark, Peptide folding: When simulation meets experiment, Angew. Chem. Int. 38 (1999) 236–240.

[188] B. Derjaguin, A theory of interaction of particles in presence of electric double-layers and the stability of lyophobe colloids and disperse systems, Prog. Surf. Sci. 43 (1993) 1–14.

[189] B. Derjaguin, L. Landau, Theory of the stability of strongly charged lyophobic sols and of the adhesion of strongly charged particles in solutions of electrolytes, Prog. Surf. Sci. 43 (1993) 30–59.

[190] E. J. W. Verwey, J. T. G. Overbeek, Theory of the Stability of Lyophobic Colloids, Elsevier, Amsterdam, 1948.

[191] R. P. Joosten, T. A. H. te Beek, E. Krieger, M. L. Hekkelman, R. W. W. Hooft, R. Schneider, C. Sander, G. Vriend, A series of PDB related databases for everyday needs, Nucleic Acids. Res. 39 (2011) D411–D419.

[192] W. Kabsch, C. Sander, Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.

[193] N. Chennamsetty, V. Voynov, V. Kayser, B. Helk, B. L. Trout, Design of therapeutic proteins with enhanced stability, Proc. Natl. Acad. Sci. U.S.A. 106 (2009) 11937–11942.

[194] N. de Groot, I. Pallares, F. Aviles, J. Vendrell, S. Ventura, Prediction of "hot spots" of aggregation in disease-linked polypeptides, BMC Struct. Biol. 5 (2005) 18.

[195] F. Chiti, N. Taddei, F. Baroni, C. Capanni, M. Stefani, G. Ramponi, C. M. Dobson, Kinetic partitioning of protein folding and aggregation, Nat. Struct. Mol. Biol. 9 (2002) 137–143.

[196] S. Ventura, J. Zurdo, S. Narayanan, M. P. no, R. Mangues, B. Reif, F. Chiti, E. Giannoni, C. M. Dobson, F. X. Aviles, L. Serrano, Short amino acid stretches can mediate amyloid formation in globular proteins: The Src homology 3 (SH3) case., Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 7258–7263.

[197] K. K. Frousios, V. A. Iconomidou, C. M. Karletidi, S. J. Hamodrakas, Amyloidogenic determinants are usually not buried, BMC Struct. Biol. 9 (2009) 44.

[198] M. J. Abraham, T. Murtola, R. Schulz, S. Pall, J. C. Smith, B. Hess, E. Lindahl, Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1-2 (2015) 19 – 25.

[199] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, G. Bussi, Plumed 2: New feathers for an old bird, Comput. Phys. Commun. 185 (2014) 604 – 613.

[200] C. Oostenbrink, A. Villa, A. E. Mark, W. F. Van Gunsteren, A biomolecular

force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6, J. Comput. Chem. 25 (2004) 1656–1676.

[201] P. F. J. Fuchs, H. S. Hansen, P. H. Hunenberger, B. A. C. Horta, A GROMOS parameter set for vicinal diether functions: Properties of polyethyleneoxide and polyethyleneglycol, J. Chem. Theory Comput. 8 (2012) 3943–3963.

[202] A. M. Gronenborn, D. R. Filpula, N. Z. Essig, A. Achari, M. Whitlow, P. T. Wingfield, G. Marius Clore, A novel highly stable fold of the immunoglobulin binding domain of streptococcal protein G, Science 253 (1991) 657–661.

[203] X. Tang, K. J. Huston, R. G. Larson, Molecular dynamics simulations of structure-property relationships of Tween 80 surfactants in water and at interfaces, J. Phys. Chem. B 118 (2014) 12907–12918.

[204] M. E. Mahmood, D. A. F. Al-koofee, Effect of temperature changes on critical micelle concentration for Tween series surfactant, Global. J. Sci. Front. Res. Chem. 13 (2013) 1–7.

[205] M. Masakazu, Y. Takuma, T. Hideki, GenIce: Hydrogen-disordered ice generator, J. Comput. Chem. 39 (2018) 61–64.

[206] S. Das, B. H. Lee, R. T. H. Linstadt, K. Cunha, Y. Li, Y. Kaufman, Z. A. Levine, B. H. Lipshutz, R. D. Lins, J.-E. Shea, A. J. Heeger, B. K. Ahn, Molecularly smooth self-assembled monolayer for high-mobility organic field-effect transistors, Nano Lett. 16 (2016) 6709–6715.

[207] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.

[208] A. Laio, M. Parrinello, Escaping free-energy minima, Proc. Natl. Acad. Sci. 99 (2002) 12562–12566.

[209] J. Pfaendtner, M. Bonomi, Efficient sampling of high-dimensional free-energy landscapes with parallel bias metadynamics, J. Chem. Theory Comput. 11 (2015) 5062–5067.

[210] A. Prakash, K. Sprenger, J. Pfaendtner, Essential slow degrees of freedom in protein-surface simulations: A metadynamics investigation, Biochem. Biophys. Res. Commun. 498 (2018) 274–281.

[211] M. Deighan, J. Pfaendtner, Exhaustively sampling peptide adsorption with metadynamics, Langmuir 29 (2013) 7999–8009.

[212] K. G. Sprenger, J. Pfaendtner, Strong electrostatic interactions lead to entropically favorable binding of peptides to charged surfaces, Langmuir 32 (2016) 5690–5701.

[213] K. G. Sprenger, A. Prakash, G. Drobny, J. Pfaendtner, Investigating the role of phosphorylation in the binding of silaffin peptide R5 to silica with molecular dynamics simulations, Langmuir 34 (2018) 1199–1207.

[214] T. L. Ogorzalek, S. Wei, Y. Liu, Q. Wang, C. L. Brooks, Z. Chen, E. N. G. Marsh, Molecular-level insights into orientation-dependent changes in the

thermal stability of enzymes covalently immobilized on surfaces, Langmuir 31 (2015) 6145–6153.

[215] Z. A. Levine, M. V. Rapp, W. Wei, R. G. Mullen, C. Wu, G. H. Zerze, J. Mittal, J. H. Waite, J. N. Israelachvili, J.-E. Shea, Surface force measurements and simulations of mussel-derived peptide adhesives on wet organic surfaces, Proc. Natl. Acad. Sci. 113 (2016) 4332–4337.

[216] P. Raiteri, A. Laio, F. L. Gervasio, C. Micheletti, M. Parrinello, Efficient reconstruction of complex free energy landscapes by multiple walkers metadynamics, J. Phys. Chem. B 110 (2006) 3533–3539.

[217] A. Barducci, G. Bussi, M. Parrinello, Well-tempered metadynamics: A smoothly converging and tunable free-energy method, Phys. Rev. Lett. 100 (2008) 020603.

[218] Z. A. Levine, S. A. Fischer, J.-E. Shea, J. Pfaendtner, Trp-cage folding on organic surfaces, J. Phys. Chem. B 119 (2015) 10417–10425.

[219] A. Prakash, M. D. Baer, C. J. Mundy, J. Pfaendtner, Peptoid backbone flexibilility dictates its interaction with water and surfaces: A molecular dynamics investigation, Biomacromolecules 19 (2018) 1006–1015.

[220] F. Pietrucci, A. Laio, A collective variable for the efficient exploration of protein beta-sheet structures: Application to SH3 and GB1, J. Chem. Theory Comput. 5 (2009) 2197–2201.

[221] F. Baftizadeh, X. Biarnes, F. Pietrucci, F. Affinito, A. Laio, Multidimensional view of amyloid fibril nucleation in atomistic detail, J. Am. Chem. Soc. 134 (2012) 3886–3894.

[222] P. Tiwary, M. Parrinello, A time-independent free energy estimator for metadynamics, J. Phys. Chem. B 119 (2015) 736–742.

[223] L. Hovgaard, S. Frokjaer, M. van de Weert, Pharmaceutical Formulation Development of Peptides and Proteins, Second Edition, Taylor & Francis, 2012.

[224] K. A. Dill, D. O. V. Alonso, K. Hutchinson, Thermal stabilities of globular proteins, Biochemistry 28 (1989) 5439–5449.

[225] M. Davidovic, C. Mattea, J. Qvist, B. Halle, Protein cold denaturation as seen from the solvent, J. Am. Chem. Soc. 131 (2009) 1025–1036.

[226] C. R. Babu, V. J. Hilser, A. J. Wand, Direct access to the cooperative substructure of proteins and the protein ensemble via cold denaturation, Nat. Struct. Mol. Biol. 11 (2004) 352–357.

[227] J. Jonas, L. Ballard, D. Nash, High-resolution, high-pressure NMR studies of proteins, Biophys. J. 75 (1998) 445–452.

[228] B. A. Patel, P. G. Debenedetti, F. H. Stillinger, P. J. Rossky, A water-explicit lattice model of heat-, cold-, and pressure-induced protein unfolding, Biophys. J. 93 (2007) 4116–4127.

[229] B. A. Patel, P. G. Debenedetti, F. H. Stillinger, P. J. Rossky, The effect of sequence on the conformational stability of a model heteropolymer in explicit

water, J. Chem. Phys. 128 (2008) 175102.
[230] C. L. Dias, T. Ala-Nissila, M. Karttunen, I. Vattulainen, M. Grant, Microscopic mechanism for cold denaturation, Phys. Rev. Lett. 100 (2008) 118101.
[231] K. A. Pikal-Cleland, N. Rodriguez-Hornedo, G. L. Amidon, J. F. Carpenter, Protein denaturation during freezing and thawing in phosphate buffer systems: Monomeric and tetrameric β-galactosidase, Arch. Biochem. Biophys. 384 (2000) 398–406.
[232] M. C. Heller, J. F. Carpenter, T. W. Randolph, Effects of phase separating systems on lyophilized hemoglobin, J. Pharm. Sci. 85 (1996) 1358–1362.
[233] M. L. Scalley, Q. Yi, H. Gu, A. McCormack, J. R. Yates, D. Baker, Kinetics of folding of the IgG binding domain of peptostreptoccocal protein L, Biochemistry 36 (1997) 3373–3382.
[234] M. L. Scalley, D. Baker, Protein folding kinetics exhibit an Arrhenius temperature dependence when corrected for the temperature dependence of protein stability, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 10636–10640.
[235] K. W. Plaxco, D. Baker, Limited internal friction in the rate-limiting step of a two-state protein folding reaction, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 13591–13596.
[236] T. Cellmer, R. Douma, A. Huebner, J. Prausnitz, H. Blanch, Kinetic studies of protein L aggregation and disaggregation, Biophys. Chem. 125 (2007) 350–359.
[237] D. E. Kim, C. Fisher, D. Baker, A breakdown of symmetry in the folding transition state of protein L, J. Mol. Biol. 298 (2000) 971 – 984.
[238] Q. Yi, M. L. Scalley, K. T. Simons, S. T. Gladwin, D. Baker, Characterization of the free energy spectrum of peptostreptococcal protein L, Fold. Des. 2 (1997) 271–280.
[239] J. Karanicolas, C. L. Brooks, The origins of asymmetry in the folding transition states of protein L and protein G, Protein Sci. 11 (2002) 2351–2361.
[240] E.-H. Yap, N. L. Fawzi, T. Head-Gordon, A coarse-grained α-carbon protein model with anisotropic hydrogen-bonding, Proteins 70 (2008) 626–638.
[241] A. Arsiccio, R. Pisano, Clarifying the role of cryo- and lyo-protectants in the biopreservation of proteins, Phys. Chem. Chem. Phys. 20 (2018) 8267–8277.
[242] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.
[243] J. L. Abascal, E. Sanz, R. G. Fernandez, C. Vega, A potential model for the study of ices and amorphous water: TIP4P/Ice, J. Chem. Phys. 122 (2005) 234511.
[244] R. G. Fernandez, J. L. Abascal, C. Vega, The melting point of ice Ih for common water models calculated from direct coexistence of the solid-liquid interface, J. Chem. Phys. 124 (2006) 144506.

[245] W. Damm, A. Frontera, J. Tirado-Rives, W. L. Jorgensen, OPLS all-atom force field for carbohydrates, J. Comput. Chem. 18 (1997) 1955–1970.
[246] M. Wikstroem, T. Drakenberg, S. Forsen, U. Sjoebring, L. Bjoerck, Three-dimensional solution structure of an immunoglobulin light chain-binding domain of protein L. Comparison with the IgG-binding domains of protein G, Biochemistry 33 (1994) 14011–14017.
[247] D. Kony, W. Damm, S. Stoll, W. F. Van Gunsteren, An improved OPLS-AA force field for carbohydrates, J. Comput. Chem. 23 (2002) 1416–1429.
[248] P. Tiwary, M. Parrinello, From metadynamics to dynamics, Phys. Rev. Lett. 111 (2013) 230602.
[249] M. Salvalaglio, P. Tiwary, M. Parrinello, Assessing the reliability of the dynamics reconstructed from metadynamics, J. Chem. Theory Comput. 10 (2014) 1420–1425.
[250] D. Pramanik, Z. Smith, A. Kells, P. Tiwary, Can one trust kinetic and thermodynamic observables from biased metadynamics simulations?: Detailed quantitative benchmarks on millimolar drug fragment dissociation, J. Phys. Chem. B 123 (2019) 3672–3678.
[251] P. Tiwary, B. Berne, Spectral gap optimization of order parameters for sampling complex molecular systems, Proceedings of the National Academy of Sciences 113 (2016) 2839–2844.
[252] Y. Wang, E. Papaleo, K. Lindorff-Larsen, Mapping transiently formed and sparsely populated conformations on a complex energy landscape, Elife 5 (2016) e17505.
[253] D. Branduardi, F. L. Gervasio, M. Parrinello, From A to B in free energy space, J. Chem. Phys. 126 (2007) 054103.
[254] J. McCarty, M. Parrinello, A variational conformational dynamics approach to the selection of collective variables in metadynamics, J. Chem. Phys. 147 (2017) 2014109.
[255] Z. F. Brotzakis, M. Parrinello, Enhanced sampling of protein conformational transitions via dynamically optimized collective variables, J. Chem. Theory Comput. 15 (2019) 1393–1398.
[256] Y. Wang, O. Valsson, P. Tiwary, M. Parrinello, K. Lindorff-Larsen, Frequency adaptive metadynamics for the calculation of rare-event kinetics, J. Chem. Phys. 149 (2018) 072309.
[257] R. B. Russell, G. J. Barton, Multiple protein sequence alignment from tertiary structure comparison: Assignment of global and residue confidence levels, Proteins 14 (1992) 309–323.
[258] I. D. Kuntz, T. S. Brassfield, G. D. Law, G. V. Purcell, Hydration of macromolecules, Science 163 (1969) 1329–1331.
[259] M. G. Usha, R. J. Wittebort, Orientational ordering and dynamics of the hydrate and exchangeable hydrogen atoms in crystalline crambin, J. Mol. Biol. 208 (1989) 669–678.

[260] A. B. Siemer, K.-Y. Huang, A. E. McDermott, Protein–ice interaction of an antifreeze protein observed with solid-state NMR, Proc. Natl. Acad. Sci. U.S.A. 107 (2010) 17580–17585.

[261] K. Tompa, P. Bánki, M. Bokor, P. Kamasa, G. Lasanda, P. Tompa, Interfacial water at protein surfaces: Wide-line NMR and DSC characterization of hydration in ubiquitin solutions, Biophys. J. 96 (2009) 2789–2798.

[262] B. Bhatnagar, B. A. Zakharov, A. S. Fisyuk, X. Wen, F. Z. Karim, K. Lee, Y. V. Seryotkin, M. Mogodi, A. Fitch, E. Boldyreva, A. Kostyuchenko, E. Shalaev, Protein/ice interaction: High-resolution synchrotron X-ray diffraction differentiates pharmaceutical proteins from lysozyme, J. Phys. Chem. B 123 (2019) 5690–5699.

[263] C. A. Knight, A. L. DeVries, Ice growth in supercooled solutions of a biological "antifreeze", AFGP 1–5: an explanation in terms of adsorption rate for the concentration dependence of the freezing point, Phys. Chem. Chem. Phys. 11 (2009) 5749–5761.

[264] S. M. Marks, A. J. Patel, Antifreeze protein hydration waters: Unstructured unless bound to ice, Proc. Natl. Acad. Sci. 115 (2018) 8244–8246.

[265] P.-L. Chau, A. J. Hardwick, A new order parameter for tetrahedral configurations, Mol. Phys. 93 (1998) 511–518.

[266] D. van der Spoel, P. J. van Maaren, P. Larsson, N. Timneanu, Thermodynamics of hydrogen bonding in hydrophilic and hydrophobic media, J. Phys. Chem. B 110 (2006) 4393–4398.

[267] A. Luzar, D. Chandler, Hydrogen-bond kinetics in liquid water, Nature 379 (1996) 55–57.

[268] A. Luzar, Resolving the hydrogen bond dynamics conundrum, J. Chem. Phys. 113 (2000) 10663–10675.

[269] G. D. Rose, P. J. Fleming, J. R. Banavar, A. Maritan, A backbone-based theory of protein folding, Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 16623–16633.

[270] A. Arsiccio, A. A. Barresi, T. De Beer, I. Oddone, P.-J. Van Bockstal, R. Pisano, Vacuum induced surface freezing as an effective method for improved inter- and intra-vial product homogeneity, Eur. J. Pharm. Biopharm. 128 (2018) 210–219.

[271] I. Oddone, A. Arsiccio, C. Duru, K. Malik, J. Ferguson, R. Pisano, P. Matejtschuk, Vacuum induced surface freezing for the freeze-drying of the human growth hormone: How does nucleation control affect protein stability?, J. Pharm. Sci. In press (2019) https://doi.org/10.1016/j.xphs.2019.04.014.

[272] J. Liu, T. Viverette, M. Virgin, M. Anderson, P. Dalal, A study of the impact of freezing on the lyophilization of a concentrated formulation with a high fill depth, Pharm. Dev. Technol. 10 (2005) 261–272.

[273] A. Margaritis, A. Bassi, Principles and biotechnological applications of bacterial ice nucleation, Crit. Rev. Biotechnol. 11 (1991) 277–295.

[274] W. Lindong, N. Shannon, S. Anisa, L. Shannon, T. Mehmet, Controlled ice nucleation using freeze-dried *Pseudomonas syringae* encapsulated in alginate beads, Cryobiology 75 (2017) 1–6.

[275] N. Cochet, P. Widehem, Ice crystallization by *Pseudomonas syringae*, Appl. Microbiol. Biotechnol. 54 (2000) 153–161.

[276] W. Rau, Eiskeimbildung durch Dielektrische Polarisation, Zeitschrift fur Naturforschung A 6 (1951) 649–657.

[277] A. Petersen, H. Schneider, G. Rau, B. Glasmacher, A new approach for freezing of aqueous solutions under active control of the nucleation temperature, Cryobiology 53 (2006) 248–257.

[278] T. Inada, X. Zhang, A. Yabe, Y. Kozawa, Active control of phase change from supercooled water to ice by ultrasonic vibration, part 1: Control of freezing temperature, Int. J. Heat Mass Transfer 44 (2001) 4523–4531.

[279] X. Zhang, T. Inada, A. Yabe, S. Lu, Y. Kozawa, Active control of phase change from supercooled water to ice by ultrasonic vibration, part 2: Generation of ice slurries and effect of bubble nuclei, Int. J. Heat Mass Transfer 44 (2001) 4533–4539.

[280] M. Saclier, R. Peczalski, J. Andrieu, A theoretical model for ice primary nucleation induced by acoustic cavitation, Ultrason. Sonochem. 17 (2010) 98–105.

[281] M. Saclier, R. Peczalski, J. Andrieu, Effect of ultrasonically induced nucleation on ice crystals size and shape during freezing in vials, Chem. Eng. Sci. 65 (2010) 3064–3071.

[282] K. Nakagawa, A. Hottot, S. Vessot, J. Andrieu, Influence of controlled nucleation by ultrasounds on ice morphology of frozen formulations for pharmaceutical proteins freeze-drying, Chem. Eng. Process. 45 (2006) 783–791.

[283] S. Passot, I. C. Trelea, M. Marin, M. Galan, G. J. Morris, F. Fonseca, Effect of controlled ice nucleation on primary drying stage and protein recovery in vials cooled in a modified freeze-dryer, J. Biochem. Eng. 131 (2009) 074511.

[284] S. Rambhatla, R. Ramot, C. Bhugra, M. J. Pikal, Heat and mass transfer scale-up issues during freeze drying: II. Control and characterization of the degree of supercooling, AAPS PharmSciTech 5 (2004) 54–62.

[285] S. M. Patel, C. Bhugra, M. J. Pikal, Reduced pressure ice fog technique for controlled ice nucleation during freeze-drying, AAPS PharmSciTech 10 (2009) 1406–1411.

[286] R. Geidobler, S. Mannschedel, G. Winter, A new approach to achieve controlled ice nucleation of supercooled solutions during the freezing step in freeze-drying, J. Pharm. Sci. 101 (2012) 4409–4413.

[287] M. Umbach, Freeze drying plant, EP Patent 3093597 B1 (December 2017).

[288] W. Ling, Controlled nucleation during freezing step of freeze drying cycle using pressure differential ice crystals distribution from condensed frost, US Patent 8875413 B2 (November 2014).

[289] T. H. Gasteyer, R. R. Sever, B. Hunek, N. Grinter, M. L. Verdone, Lyophilization system and method, US Patent 9651305 B2 (May 2017).

[290] R. Bursac, R. Sever, B. Hunek, A practical method for resolving the nucleation problem in lyophilization, Bioproc. Int. 7 (2009) 6672.

[291] A. K. Konstantinidis, W. Kuu, L. Otten, S. L. Nail, R. R. Sever, V. Bons, D. Debo, M. J. Pikal, Conrolled nucleation in freeze-drying: Effects on pore size in the dried product layer, mass transfer resistance, and primary drying rate, J. Pharm. Sci. 100 (2011) 3453–3470.

[292] M. Kramer, B. Sennhenn, G. Lee, Freeze-drying using vacuum-induced surface freezing, J. Pharm. Sci. 91 (2002) 433–443.

[293] S. M. Patel, D. Takayuki, M. J. Pikal, Determination of end point of primary drying in freeze-drying process control, AAPS PharmSciTech 11 (2010) 73–84.

[294] S. A. Velardi, V. Rasetto, A. A. Barresi, Dynamic parameters estimation method: Advanced manometric temperature measurement approach for freeze-drying monitoring of pharmaceutical, Ind. Eng. Chem. Res. 47 (2008) 8445–8457.

[295] S. A. Velardi, A. A. Barresi, Method and system for controlling a freeze drying process, patent US8800162 B2 (August 2014).

[296] J. Gearing, K. P. Malik, P. Matejtschuk, Use of dynamic mechanical analysis (DMA) to determine critical transition temperatures in frozen biomaterials intended for lyophilization, Cryobiology 61 (2010) 27–32.

[297] Council of Europe European Pharmacopoeia Commission, European Pharmacopoeia 9.0, Strasbourg (2017).

[298] T. Tanaka, R. P. Shiu, P. W. Gout, C. T. Beer, R. L. Noble, H. G. Friesen, A new sensitive and specific bioassay for lactogenic hormones: Measurement of prolactin and growth hormone in human serum, J. Clin. Endocrinol. Metab 51 (1980) 1058–1063.

[299] T. B. L. Kirkwood, Predicting the stability of biological standards and products, Biometrics 33 (1977) 736–742.

[300] T. B. L. Kirkwood, M. S. Tydeman, Design and analysis of accelerated degradation tests for the stability of biological standards II. A flexible computer program for data analysis, J. Biol. Stand. 12 (1984) 207–214.

[301] T. B. L. Kirkwood, Design and analysis of accelerated degradation tests for the stability of biological standards III. Principles of design, J. Biol. Stand. 12 (1984) 215–224.

[302] Council of Europe European Pharmacopoeia Commission, European Pharmacopoeia 8.2: Monograph 2.7.4 assay of human coagulation factor VIII, Strasbourg (2014).

[303] A. Burger, J. O. Henck, S. Hetz, J. M. Rollinger, A. A. Weissnicht, H. Stottner, Energy/temperature diagram and compression behavior of the polymorphs of D-mannitol., J. Pharm. Sci. 89 (2000) 457–468.

[304] M. Mehta, S. P. Bhardwaj, R. Suryanarayanan, Controlling the physical form of mannitol in freeze-dried systems, Eur. J. Pharm. Biopharm. 85 (2013) 207–213.

[305] M. G. Cares-Pacheco, G. Vaca-Medina, R. Calvet, F. Espitalier, J.-J. Letourneau, A. Rouilly, E. Rodier, Physicochemical characterization of D-mannitol polymorphs: The challenging surface energy determination by inverse gas chromatography in the infinite dilution region, Int. J. Pharm. 475 (2014) 69–81.

[306] M. L. Williams, R. F. Landel, J. D. Ferry, The temperature dependence of relaxation mechanisms in amorphous polymers and other glass-forming liquids, J. Am. Chem. Soc. 77 (1995) 3701–3707.

[307] I. Vollrath, W. Friess, A. Freitag, A. Hawe, G. Winter, Does controlled nucleation impact the properties and stability of lyophilized monoclonal antibody formulations?, Eur. J. Pharm. Biopharm. 129 (2018) 134–144.

[308] J. M. Sarciaux, S. Mansour, M. J. Hageman, S. L. Nail, Effects of buffer composition and processing conditions on aggregation of bovine IgG during freeze-drying, J. Pharm. Sci. 88 (1999) 1354–1361.

[309] M. J. Pikal, S. Rambhatla, R. Ramot, The impact of the freezing stage in lyophilization: Effects of the ice nucleation temperature on process design and product quality, Am. Pharm. Rev. 5 (2002) 48–53.

[310] L. Lewis, R. E. Johnson, M. E. Oldroyd, S. S. Ahmed, L. Joseph, I. Saracovan, S. Sinha, Characterizing the freeze-drying behavior of model protein formulations, AAPS PharmSciTech 11 (2010) 1580–1590.

[311] S. J. Shire, Z. Shahrokh, J. Liu, Challenges in the development of high protein concentration formulations, J. Pharm. Sci. 93 (2004) 1390–1402.

[312] A. Arsiccio, A. C. Sparavigna, R. Pisano, A. A. Barresi, Measuring and predicting pore size distribution of freeze-dried solutions, Dry. Technol. 37 (2019) 435–447.

[313] A. Arsiccio, A. A. Barresi, R. Pisano, Prediction of ice crystal size distribution after freezing of pharmaceutical solutions, Cryst. Growth Des. 17 (2017) 4573–4581.

[314] A. Arsiccio, R. Pisano, Application of the quality by design approach to the freezing step of freeze-drying: Building the design space, J. Pharm. Sci. 107 (2018) 1586–1596.

[315] A. Arsiccio, P. Giorsello, L. Marenco, R. Pisano, Considerations on protein stability during freezing and its impact on the freeze drying cycle: A design space approach, J. Pharm. Sci.

[316] R. P. Singh, D. R. Heldman, Introduction to Food Engineering, Academic Press, San Diego, U.S.A., 2014.

[317] W. Kurz, D. J. Fischer, Fundamentals of Solidification, Trans Tech Publications, Switzerland, 1992.

[318] W. Bald, Food Freezing: Today and Tomorrow, Springer-Verlag, New York/London, 1991.
[319] J. L. Bomben, C. J. King, Heat and mass transport in the freezing of apple tissue, Int. J. Food Sci. Technol. 17 (1982) 615–632.
[320] K. Nakagawa, A. Hottot, S. Vessot, J. Andrieu, Modeling of freezing step during freeze-drying of drugs in vials, AIChE J. 53 (2007) 1362–1372.
[321] M. Kochs, C. Körber, B. Nunner, I. Heschel, The influence of the freezing process on vapour transport during sublimation in vacuum-freeze-drying, Int. J. Heat Mass Transfer 34 (1991) 2395 – 2408.
[322] S. L. Nail, M. J. Akers, Development and Manufacture of Protein Pharmaceuticals, Kluwer Academic/Plenum Publisher, New York, U.S.A., 2002.
[323] R. Pisano, L. C. Capozzi, Prediction of product morphology of lyophilized drugs in the case of vacuum induced surface freezing, Chem. Eng. Res. Des. 125 (2017) 119–129.
[324] M. Matsumoto, S. Saito, I. Ohmine, Molecular dynamics simulation of the ice nucleation and growth process leading to water freezing, Nature 416 (2002) 409–413.
[325] T. Nishinaga, K. Nishioka, J. Harada, A. Sasaki, H. Takei, Advances in the Understanding of Crystal Growth Mechanisms, Elsevier Science, Amsterdam, 2012.
[326] N. Hannay, Changes of State -Treatise on Solid State Chemistry, Springer Science & Business Media, New York, U.S.A., 2013.
[327] D. Fissore, R. Pisano, Computer-aided framework for the design of freeze-drying cycles: Optimization of the operating conditions of the primary drying stage, Processes 3 (2015) 406–421.
[328] D. Fissore, R. Pisano, A. A. Barresi, Advanced approach to build the design space for the primary drying of a pharmaceutical freeze-drying process., J. Pharm. Sci. 100 (2011) 4922–4933.
[329] R. Pisano, D. Fissore, A. A. Barresi, P. Brayard, P. Chouvenc, B. Woinet, Quality by design: Optimization of a freeze-drying cycle via design space in case of heterogeneous drying behavior and influence of the freezing protocol, Pharm. Dev. Technol. 18 (2013) 280–295.
[330] G. Blond, D. Simatos, M. Catté, C. G. Dussap, J. B. Gros, Modeling of the water-sucrose state diagram below 0 °C, Carbohydr. Res. 298 (3) (1997) 139–145.
[331] F. E. Young, F. T. Jones, Sucrose hydrates. The sucrose-water phase diagram, J. Phys. Chem. 53 (1949) 1334–1350.
[332] R. Zwanzig, Two-state models of protein folding kinetics, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 148–150.
[333] T. E. Creighton, Toward a better understanding of protein folding pathways, Proc. Natl. Acad. Sci. U.S.A. 85 (1988) 5082–5086.

[334] E. Haas, E. Katchalski-Katzir, I. Z. Steinberg, Brownian motion of the ends of oligopeptide chains in solution as estimated by energy transfer between the chain ends, Biopolymers 17 (1978) 11–31.

[335] A. Ansari, C. M. Jones, E. R. Henry, J. Hofrichter, W. A. Eaton, The role of solvent viscosity in the dynamics of protein conformational changes, Science 256 (1992) 1796–1798.

[336] T. Kleinert, W. Doster, H. Leyser, W. Petry, V. Schwarz, M. Settles, Solvent composition and viscosity effects on the kinetics of CO binding to horse myoglobin, Biochemistry 37 (1998) 717–733.

[337] A. B. Anderson, C. R. Robertson, Absorption spectra indicate conformational alteration of myoglobin adsorbed on polydimethylsiloxane, Biophys. J. 68 (1995) 2091–2097.

[338] P. L. Privalov, Thermodynamics of protein folding, J. Chem. Thermodyn. 29 (1997) 447–474.

[339] R. H. M. Hatley, F. Franks, The cold-induced denaturation of lactate dehydrogenase at sub-zero temperatures in the absence of perturbants, FEBS Lett. 257 (1989) 171–173.

[340] W. Cao, Y. Xie, S. Krishnan, H. Lin, R. M, Influence of process conditions on the crystallization and transition of metastable mannitol forms in protein formulations during lyophilization, Pharm. Res. 30 (2013) 131–139.

[341] B. H. Peters, L. Staels, J. Rantanen, F. Molnar, T. De Beer, V. P. Lehto, J. Ketolainen, Effects of cooling rate in microscale and pilot scale freeze-drying - variations in excipient polymorphs and protein secondary structure, Eur. J. Pharm. Sci. 95 (2016) 72–81.

[342] A. Shosheva, M. Miteva, P. Christova, B. Atanasov, pH-dependent stability of sperm whale myoglobin in water-guanidine hydrochloride solutions, Eur. Biophys. J. 31 (2003) 617–625.

[343] X. C. Tang, M. J. Pikal, The effect of stabilizers and denaturants on the cold denaturation temperatures of proteins and implications for freeze-drying, Pharm. Res. 22 (2005) 1167–1175.

[344] M. A. Miller, M. A. Rodrigues, M. A. Glass, S. K. Singh, K. P. Johnston, J. A. Maynard, Frozen-state storage stability of a monoclonal antibody: Aggregation is impacted by freezing rate and solute distribution, J. Pharm. Sci. 12 (2013) 1194–1208.

[345] E. Cao, Y. Chen, Z. Cui, P. R. Foster, Effect of freezing and thawing rates on denaturation of proteins in aqueous solutions, Biotechnol. Bioeng. 82 (2003) 684–690.

[346] M. L. Roy, M. J. Pikal, Process control in freeze drying: Determination of the end point of sublimation drying by an electronic moisture sensor, J. Parenter. Sci. Technol. 43 (1989) 60–66.

This Ph.D. thesis has been typeset by means of the TEX-system facilities. The typesetting engine was pdfLATEX. The document class was `toptesi`, by Claudio Beccari, with option `tipotesi=scudo`. This class is available in every up-to-date and complete TEX-system installation.