Biophysical Properties of Proteins in Solution

And the Effects of Protein Glycosylation

PhD Thesis

Heidi Louise Bagger

Department of Science, Systems and Models
Roskilde University, Denmark, 2007
Biophysical Properties of Proteins in Solution

And the Effects of Protein Glycosylation

By Heidi Louise Bagger

Department of Science, Systems and Models,
Roskilde University, Denmark
Preface

This PhD thesis, entitled “Biophysical properties of proteins in solution, and the effects of Protein Glycosylation” is submitted to meet the requirements for attaining the Danish PhD degree at Roskilde University. The PhD work was initiated April 2003 and has been carried out in collaboration with Novozymes A/S as an industrial PhD project. The industrial PhD initiative is aimed at enhancing research and development in the Danish business sector and is administered by the Ministry of Science, Technology and Innovation. The experimental work of this project has primarily been carried out at Department of Science, Systems and Models, Roskilde University. Additional experimental work is accomplished in the business unit Protein Chemistry, Novozymes A/S; at the Faculty of Life Sciences, University of Copenhagen; and at the Institute for Storage Ring Facilities (ISA), University of Aarhus.

Acknowledgements. I am thankful to all the people who contributed with support and inspiration during this project. First of all, I thank my supervisors Peter Westh and Claus Crone Fuglsang for setting up a wide framework for my work and for continuous support. I am grateful to Søren Vrønning Hoffmann (University of Aarhus) for invaluable help in connection with my synchrotron radiation CD experiments and subsequent spectra analysis; and to Lars Holm Øgendal (University of Copenhagen) for very useful expert assistance with treatment and modulation of light scattering data. I will also thank my office mates over the years, Karin Liltorp, Anders Dybdal Nielsen, Lars Lundby Hansen and Søren Nymand Olsen for valuable discussions and good company. Finally I will thank my friends and family for being continuously supportive.

Heidi Louise Bagger
Copenhagen, April 2007
Successful industrial use of proteins requires sufficient functionality and stability under production, storage and application conditions. Addition of excipients (inactive substances used in formulation of active ingredients) is a universal way to modify the biophysical properties of proteins, such as the thermodynamic and kinetic stability. In this work we describe protein-solvent interactions and their effects on equilibrium stability and solubility, and furthermore discuss these interactions with respect to irreversible aggregation of denatured monomers. To elucidate the complex interrelation between preferential interactions and protein aggregation, we investigated the aggregate growth of BSA in the presence of different solutes by means of static light scattering (SLS) and size exclusion chromatography. It was found that heat induced BSA aggregation is a sensitive process towards solvent modifications, and that the non-specific interactions usually dominating the excipient induced stabilization of native proteins are of minor importance for the kinetics of irreversible BSA aggregation. For example, the investigated ionic compounds, Na$_2$SO$_4$ (strong kosmotropic anion) and NaSCN (strong chaotropic anion) both very effectively suppress the aggregation process.

Protein associations, e.g. aggregation and reversible precipitation, are effectively suppressed by protein glycosylation and by the attachment of synthetic polymers, and thus, such modifications are of growing industrial interest. One important purpose of this thesis is to evaluate the effects of glycosylation on the equilibrium and kinetic stability of proteins. The basis of this evaluation is a literature review and experimental investigations of different glyco forms of the enzyme phytase.

Vapor pressure osmometry and calorimetry have been used in this work to study molecular interactions in different phytase systems. It has been shown that relative to the peptide part, the glycans interact very differently with solvent components such as water, polyols and surfactants. For example, determination of preferential interactions in solutions of sorbitol and glycerol showed that deglycosylation of phytase is associated with increased preferential protein hydration. Additionally, hydration studies of dry and semi dry phytase samples implied increased water affinity of peptides relative to glycans. This suggest that the
increased solubility generally found for glycosylated proteins rely on steric restrictions of close packing rather than increased hydrophilicity upon glycosylation as often suggested in the literature. The glycan induced changes in protein-protein interactions are also reflected in SLS studies of heated phytase solutions. These studies showed that the kinetics of protein aggregation is significantly increased upon deglycosylation, and that aggregates of glycosylated proteins most probably hold a restricted number of monomers.

Despite more favorable interactions of glycans, relative to peptides, with the common excipients sorbitol and glycerol and less favorable interactions with the anionic surfactant sodium dodecyl sulfate (SDS), the glycan mantle of phytase reveals only minor effects on equilibrium structural transitions in these systems. This is concluded on the basis of different calorimetric methods and synchrotron radiation circular dichroism (SRCD) spectroscopy. Interestingly sorbitol and glycerol, which are preferentially excluded from the peptide part of phytase (and from peptides in general), have been shown to be preferentially bound to the carbohydrate moieties of this enzyme. This indicates that glycan-polyol interactions increase protein solubility and consequently hold the potential to increase the compatibility of proteins in systems containing such stabilizing components.

Glyco effects undoubtedly depend on the investigated systems. However, some important, and general, characteristics of glycosylated proteins can certainly be emphasized. On the basis of the present phytase studies and a literature review it is concluded that glycosylation causes increased steric repulsions between protein monomers and consequently a considerable increase in solubility and kinetic stability. However, the thermodynamic stability of proteins is usually not affected by glycosylation, which is interpreted as equal glycan-solvent interactions for native and denatured glycoproteins.
Dansk Resumé


Tendensen til at aggregere er markant lavere for glycoproteiner samt for proteiner der er overflademodificeret med f.eks. syntetiske polymere såsom polyetylen glycol (PEG). Overflademodifikationer er derfor af stor industriel interesse. Formålet med denne afhandling er bl.a. at evaluere effekterne af proteinglykosylering på ligevægtsstabilitet og irreversibel aggregering. Denne evaluering er baseret på et litteraturstudie og på eksperimentelle undersøgelser af det glykosylerede enzym phytase.

Damptryksosmometri og kalorimetri er brugt til at undersøge molekylære veckslvirkninger i forskellige phytase systemer. Det har vist sig, at relativt til peptiddelen af phytase, veckslvirker glykanerne signifikant anderledes med solventkomponenter såsom vand, polyoler og detergenter. Bestemmelse af præferentielle veckslvirkninger i opløsninger bestående af phytase og polyoler (sorbitol eller glycerol) viser, at deglykosylering af phytase er forbundet med øget præferentiel hydtering af proteinet. Endvidere viser hydrieringsstudier med delvist tørre proteinprøver, at peptiddelen af phytase har større vandaffinitet end glykanerne. Dette indikerer, at den øgede opløselighed som generelt findes for

På baggrund af forskellige kalorimetriske metoder og CD spektroskopi har vi vist, at trods glycaners favorable vekselvirkninger (relativt til peptider) med de stabiliserende solutter sorbitol og glycerol, og mindre favorable vekselvirkninger med den anioniske detergent sodium dodecyl sulfat (SDS), medfører deglykosylering af phytase ikke betydelig effekt på ligevægtstabiliteten i disse solventsystemer. Glycosylering øger dog potentiel kompatibiliteten af proteiner og visse stabiliserende excipienter, da favorable glycan-solut interaktioner vil medføre øget proteinopløselighed i sådanne systemer.

INTRODUCTION

A crucial factor in protein product development is preservation of functional peptide structures. Proteins and enzymes constitute a high potential in the development of industrial products such as pharmaceuticals and cosmetics as well as detergent -, textile - and food products [1]. To retain the function of proteins, the peptide structure has to be folded in the unique native conformation, which is only marginal stable and very sensitive towards external factors e.g. temperature, pH, and chemical composition of the solvent (or of the solid matrix). Therefore these molecules often have a tendency to both denature and undergo irreversible aggregation under various production, storage and application conditions. The requirement of industrial proteins to function under severe conditions is high and thus highlights the need for rational formulation strategies.

This thesis primarily focuses on proteins in solution. Within protein product development, there are also a number of pertinent problems regarding solid-state stability of proteins which will not be treated. The importance of e.g. water content and the composition of excipients (stabilizing components) in protein powders is well established, and it has generally been noted that the lower the residual water level, the more the stable protein product [2-4]. In powder mixtures the relative hygroscopicity of proteins and additives determines the peptide hydration, and thus together with temperature govern the stability of “dry” proteins. As with proteins in solution it is in principle the interactions between the present components and different protein states that determines the stability of solid state samples.

A simple and highly conserved way used by Nature to prevent denaturation of proteins is production and accumulation of organic solutes e.g. trehalose, glycerol and free amino acids. Particularly high concentrations of such solutes are found in organisms living at extreme temperatures or under osmotic stress [5-7]. This stabilizing strategy is mimicked in protein formulation protocols where addition of salts and non-ionic components (e.g. sugars and polyols) at relative high concentrations (>1M) is frequently used for stabilization. In
contrast aggregation of denatured proteins (*in vivo*) is primarily inhibited by complex interactions between many proteins with chaperone activity. These chaperones function either by supporting the correct folding of the peptide chain, or by preventing association between unfolded or misfolded proteins [8-11]. In spite of the indispensability of chaperones *in vivo* addition of such protein components is in general not applied for protein product development. Compared to small compatible solutes, chaperones are much more system specific and the costs-benefit balance using these proteins during product development are usually unfavorable. Aggregation may also be modulated by conventional stabilizing components such as sucrose and Na₂SO₄. However, the aggregation preventing properties of the frequently used excipients is usually not straight forward, which will be discussed in some detail in Part I “Molecular Interactions and Protein Properties” below and in Paper 4. Therefore in the attempt to prevent protein aggregation, surface modifications are of growing interest in biotechnology. For example, covalent addition of carbohydrates (chemical glycosylation) and addition of synthetic polymers (e.g. PEGylation) has revealed promising results, and has been suggested as one of the most promising approaches for stabilization towards denaturation and aggregation of proteins during industrial applications [12-15].

In this work it is important to differentiate between equilibrium protein stability and protein aggregation, which is a kinetically controlled process, usually far from equilibrium. Under equilibrium conditions the free energy of denaturation, \( \Delta G = 0 \). In contrast, denaturation is thermodynamically favored during aggregation, thus \( \Delta G < 0 \). The equilibrium stability of a protein (referred to as “stability” in this report) can be described by the difference in standard free energy (\( \Delta G^\circ \)) of the native (N) and denatured (D) states. The limited stability of proteins at room temperature is a consequence of a relative equal standard free energy of those different states: \( \Delta G^\circ, G^\circ_{D} - G^\circ_{N}, \sim 20 \text{ to } 65 \text{ kJ/mol} \) [16, 17]. A large number of relatively weak interactions are involved in the process of peptide folding. For comparisons, the typical enthalpy change for protein denaturation is \( \sim 500-1000 \text{ kJ/mol} \), this corresponds to the bond enthalpies of two C-C covalent bonds.

In biotechnology different approaches are used to increase protein stability, and generally they can be divided into two different modes: *internal* and *external stabilization*. External stabilization is based on changes in the composition of the surrounding medium, as mentioned above, by addition of stabilizing solutes. Such components belong to a chemically
divers group: Natural components (sugars, polyols, amino acids and different salts) which in general interact weakly with proteins and exert their function at relative high molar concentrations; and chemical component such as poly ethylene glycol (PEG) and other large and weakly interacting polymers. The mechanisms underlying external stabilization is described in Part I. Internal stabilization on the other hand covers changes in the primary structure of proteins e.g. site directed mutagenesis. This method is of course more complicated to carry out in practice and much more protein specific. Successful increase in protein stability by specific alterations in the amino acid sequence requires the three dimensional structure of the wild type protein, and the ability to predict the tertiary structure of the modified protein [18]. The aim of both of these approaches is to enhance the thermodynamic stability i.e. to displace the $N \leftrightarrow D$ equilibrium towards the left. For example, substitutions of core amino acids are often made to increase hydrophobic interactions [19]. This means that the average amino acid – solvent interactions are less favorable for the mutated protein than for the wild type protein which favors a more condensed conformation. Thus, such internal modifications primarily stabilize the native protein by increasing the free energy of unfolded states ($G_D^{\uparrow}$). Other amino acid substitutions are made to introduce disulfide bridges in the folded conformation [19], hence stabilizing the protein by decreasing the free energy of the native state ($G_N^{\downarrow}$). In contrast external modifications (solvent modifications) usually lead to an increase in the free energy of both the native($G_N^{\uparrow}$) and unfolded states ($G_D^{\uparrow\uparrow}$). The last two examples of stabilization are illustrated in Fig. 1.
Figure 1. Examples of internal (Green) and external (Blue) stabilization. The protein is stabilized because the difference in standard free energy ($\Delta G^o$) between the native and the denatured state is increased relative to $\Delta G^o$ of the reference system (wild type peptide, solution without excipients).

As mentioned above surface modifications are of growing interest in the attempt to suppress irreversible aggregation of proteins i.e. to increase the kinetic stability. More than thirty years ago Vegarud et al. suggested some biophysical advantages of conjugates between peptides and polysaccharides and predicted future commercial use of glycoproteins [20]. These predictions have been partially fulfilled, and there is a still increasing industrial interest in protein glycosylation (both chemical and in vivo addition of carbohydrates). With respect to therapeutic agents, including non-peptide components, PEGylation is suggested as the most important modification [14]. This is primarily due to the in vivo compatibility of PEG groups and their positive influence on many kinds of product degradation processes. PEGylation will not be treated in this thesis, for review in this subject see [14, 15]. The interest in glycosylation (and PEGylation) is motivated by the positive biological and biophysical effects of the carbohydrate part found for naturally glycosylated proteins, and by the requirement to enhance the biotechnological potential of proteins and peptides as well as non-peptide products [12, 15, 21, 22]. The biophysical effects of protein glycosylation are discussed in Part II “Effects of Protein Glycosylation” below, and some aspects of these effects are elucidated in Paper 1, 2 and 3.
Outline and aim of the thesis

The main focus in the entire PhD work is equilibrium and kinetic stability of proteins in solution. In the first part of this report (Part I) we introduce some relevant thermodynamic and kinetic theories, while Part II reviews the effects of surface attached carbohydrates on protein properties. The scope of the thesis is somewhat broader than its experimental work, which is based on investigations of bovine serum albumin (BSA) and of the glycosylated enzyme *Peniophora Lycii* phytase. The results and conclusions from the studies on *Peniophora Lycii* phytase are presented in Paper 1-3, and as a part of the review on biophysical effects of glycosylation (in Part II). Another important scope of this work is to elucidate the general aspects of the relation between protein properties (conformational transitions and aggregation) and molecular interactions between solution components (Part I). This is done in part by summing up some theories and conclusions obtained within this field, and in part by experimental investigations of solute effects on BSA aggregation. The content of the thesis is shortly described in the outline below, paper 1-4 is found in the appendix.

- **Part I**
  The first part of this report, relates protein-solvent interactions to biophysical protein properties such as equilibrium stability and solubility, and it is further discussed how these interactions may affect irreversible protein aggregation. The latter approach, which remains mainly unexplored in the literature, is illustrated by the investigations of heat induced BSA aggregation in different solvent systems.

- **Part II**
  The second part of this thesis concentrates on protein glycosylation. On the basis of a literature review and the current investigations of *Peniophora Lycii* phytase, the effects of glycosylation on equilibrium and kinetic stability is evaluated and reviewed. Due to considerable differences between investigated systems, distinctions are made between naturally glycosylated protein, mutated glycoproteins and chemically glycosylated proteins. Part II initially gives a short description of the *in vivo* glycosylation process, and the different glyco-variants of *Peniophora Lycii* phytase that are used in the experimental work.
• **Paper 1**, (Appendix I)


The preferential interactions between two glyco variants of *Peniophora Lycii* phytase and the solutes sorbitol and glycerol were determined by vapor pressure osmometry. It was found that both solutes (but particularly sorbitol) are preferentially excluded from the surfaces of the deglycosylated peptides, and conversely preferentially bound to the glycan moieties. In spite of the distinct differences between peptide and carbohydrate groups, glycosylation has no effect on the stabilizing action provided by glycerol and sorbitol. Based on this it was concluded that the carbohydrate mantle of native phytase is fully accessible to the solvent, and thus, that its interactions with compatible solutes has little or no effect on the conformational equilibrium.

Preliminary work of this paper was published by Bagger, H.L. (2002), Master thesis “Molekylære mekanismer for solut induceret protein stabilisering: En undersøgelse af glykosyleringens effekt på præferentielle vekselvirkninger”.

• **Paper 2**, (Appendix II)


The hygroscopic character of *Peniophora Lycii* phytase in its glycosylated and deglycosylated form was investigated by a two-channel isothermal calorimetric method. The free energy and enthalpy of hydration at 25°C was determined. Comparisons of the two glyco-variants indicated small differences between hydration enthalpies, but revealed that peptide has higher water affinity than glycans. Thus, the solubilizing and aggregation preventing properties of glycosylation is unlikely to be due to particularly favorable water-glycan interactions.
• **Paper 3**, (Appendix III)
  *Biophysical Chemistry* 129:251-258

  The interactions of sodium dodecyl sulfate (SDS) and respectively the glycosylated and deglycosylated form of *Peniophora Lycii* phytase were investigated. Effects, at 24°C, of titrating SDS to Phy and dgPhy was studied by isothermal titration calorimetry (ITC) and synchrotron radiation circular dichroism (SRCD) spectroscopy. The SRCD spectra suggest that both the native and the SDS-denatured states of the two variant are mutually similar and hence that the denaturation process is structurally equivalent in the two cases. It was found that the glycans create only a small increase in the resistance towards SDS induced denaturation. Furthermore, analysis of the relative affinity of the glycan and peptide moieties suggested that the carbohydrates bind much less surfactant (at saturation, glycans adsorbed about half as much SDS (in g/g) as the peptide part of phytase).

• **Paper 4**, (Appendix IV)
  *Article in press, Biophysical Chemistry, doi:10.1016/j.bpc.2007.06.014*

  Temperature induced aggregation of BSA was investigated by light scattering and size exclusion chromatography to elucidate the interrelationships between preferential interactions and protein aggregation. The effects on growth rate and aggregate size were studied for preferentially excluded solutes (sorbitol, sucrose and Na$_2$SO$_4$) and preferentially bound solutes (urea and NaSCN). These studies have shown that non-specific “Hofmeister-type” interactions usually dominating the excipient induced stabilization of native proteins are of minor importance for the kinetics of irreversible BSA aggregation. For example it was found that the two salts (containing respectively a strong kosmotrope, SO$_4^{2-}$, and a strong chaotropic...
SCN\(^-\)) suppress the aggregation to about the same extend, and this, very effectively compared to the non-electrolytes. The substantial suppression of aggregation induced by salts was suggested to depend on modulation of electrostatic forces in the transition state.

- **Concluding remarks**
  The most important results of this thesis are extracted, and the effects of glycosylation are elucidated in an industrial perspective with respect to preservation of native and functional peptide structures.
Part I

Molecular Interactions and Protein Properties

Solvent composition and equilibrium stability

A considerable amount of work has addressed the effects of additives on the equilibrium transition between the native and a denatured state of proteins. The outcome of this, among many things, has been the discovery of some generality in the solute-effects, and thus the ability to classify solutes as general stabilizers or destabilizers. Additionally, it has become clear that external stabilization (and destabilization) is based on the relative affinity of water and solute for the protein in its different stable states (explained below). The theoretic fundamentals of this relation can be found in the thermodynamic description of unequal distribution of molecules in multi-component, non-ideal systems [23-26].

When a protein is dissolved in an aqueous solution it may interact differently with solvent components, and some interactions are more favorable than others. This is called preferential interactions. To explain the relation between molecular interactions and protein stability we consider a simple three-component system consisting of water (component 1), protein (component 2) and a solute (component 3). Preferential interactions give rise to concentration differences of solutes near the protein and in the bulk phase. If the protein surface (on average) interacts more favorable with water than with solutes, it will result in preferential exclusion of solutes. This situation, also refereed to as preferential hydration of the protein, is illustrated in Fig. 2A. Preferential exclusion can in principle be due to unfavorable solute-protein interactions, favorable protein-water (or solute-water) interactions, or due to steric hindrance and size differences between solute and water molecules. If
conversely the protein interacts more favorably with the solute than with water, there will be a preferential accumulation of solutes near the protein, Fig. 2B.

Figure 2. Consequences of unequal solute/water distribution between the solvent in the bulk and the solvent in the vicinity of the protein are illustrated. A: Preferential exclusion and B: Preferential accumulation (preferential binding). The dotted line represents the boundary between bulk and the volume close to the protein (yellow). The black spheres represent solutes and the gray spheres represent water molecules.

Preferential interactions in 3-component protein solutions can be quantified by the preferential binding parameter (\(\Gamma_{\mu_3}\)), defined in eq. 1. The preferential binding parameter is a measure of how the molality of the solute (\(m_3\)) must be changed to keep its chemical potential (\(\mu_3\)) constant when the molality of the protein (\(m_2\)) is changed at fixed temperature and pressure.

\[
\Gamma_{\mu_3} \equiv \left( \frac{dm_3}{dm_2} \right)_{T,P,\mu_3} = -\left( \frac{d\mu_2}{d\mu_3} \right)_{T,P,m_2}
\]  

(1)

If the protein concentration is increased its chemical potential (\(\mu_2\)) is consequently raised. Depending on the existing molecular interactions this can either increase \(\mu_3\)
(unfavorable protein-solute interactions) or decrease $\mu_3$ (favorable protein-solute interactions). If $\mu_3$ is fixed, the perturbation induced by an increase in $m_2$ can be adjusted by a change in $m_3$. Consequently $\Gamma_{\mu_3}$ values are positive in the case of preferential accumulation (binding) and negative when the solutes are preferentially excluded (“negative binding”). Preferential binding parameters can be obtained experimentally by equilibrium dialyses for examples see [27-32], and by vapor pressure osmometry [33-36], and Paper 1.

As the solvent exposed surface of proteins is changed upon denaturation the preferential interaction coefficients quantifying solute-protein interactions can be different for the two states N and D ($\Gamma^N_{\mu_3} \neq \Gamma^D_{\mu_3}$). For many protein-solute systems the difference between $\Gamma^N_{\mu_3}$ and $\Gamma^D_{\mu_3}$ almost exclusively depends on the extent of surface expansion subsequent to denaturation, and thus, not on the character of solvent exposed peptides. This is referred to as non-specific interactions. Explained by the Wyman linkage theory (eq.2) the difference between $\Gamma^N_{\mu_3}$ and $\Gamma^D_{\mu_3}$ leads to solvent induced changes in the stability of proteins [37, 38]

$$\left(\frac{d \ln K}{d \ln a_3}\right)_{T,P,a(x)} = \Gamma^D_{\mu_3} - \Gamma^N_{\mu_3} = \Delta \Gamma_{\mu_3} \quad (2)$$

where $K$ is the equilibrium constant of the folding reaction ($N \xrightleftharpoons[\Delta \Gamma_{\mu_3}]{K} D$), $a_3$ denotes the thermodynamic activity of the solute. $\Gamma_{\mu_3}$ is the binding parameter (eq. 1) reflecting the effective interaction between solutes and the protein molecules in its different states (N and D) and $a_{3(x)}$ is the activity of any component ($x \neq \text{component 3}$).

The Wyman linked function states that the presence of a solute in an aqueous solution can shift the equilibrium between different stable states (reactant and product) by interacting more or less favorable with those different states. It fits any biological process or biochemical reaction which is influenced by interactions with a solvent component, and it considers both specific and non-specific interactions [39, 40].

Consequences of specific interactions: Components which bind specifically to the native structure of a protein will, at low concentration, increase the stability because the
binding site (or sites) exists on the native protein only, hence $\Gamma_{\mu_3}^N > \Gamma_{\mu_3}^D$, an example is ligand induced stabilization. However, some solutes that bind strongly to specific sites on N also interacts favorably with peptides in general (preferential accumulation, Fig. 2B) and thus, at higher concentrations potentially destabilize the protein because of increased binding to the denatured, and more expanded, conformation ($\Gamma_{\mu_3}^N < \Gamma_{\mu_3}^D$). This is often found for protein-detergent mixtures. For example the anionic detergent sodium dodecyl sulfate (SDS) significantly increases the denaturation temperature ($T_d$) of BSA at low concentrations because of specific binding of a few detergent molecules. At molar ratios $\sim 10$ SDS/BSA $T_d$ starts decreasing [41, 42].

Consequences of non-specific interactions: Non-specific binding of solutes generally destabilizes proteins because the character of binding does not change when the peptide chain denatures, and the extent of binding therefore increases upon surface expansion ($\Delta \Gamma_{\mu_3} > 0$). Non-specific binding reduces the chemical potential of both N ($\mu^N$) and D ($\mu^D$), and hence increase the solubility of both states. Examples of generally favorable interacting (and perturbing) components are urea, SCN$^-$ and GuHCl (guanidine hydrochloride). In contrast stabilizing solutes are usually preferentially excluded from the surface of proteins. This is a thermodynamically unfavorable situation, which can be improved by contraction of the protein conformation and by precipitation. Examples of such solutes are betaine, sucrose, glycerol, sorbitol, SO$_4^{2-}$ and PEG. In other words typical excipients generally favor the native conformation but do also favor higher order native multimers or larger molecular assembles (reversible precipitation). Thus, in a protein formulation perspective, there is a conflict between stability and solubility effects of non-specific interacting solutes, which is illustrated in Fig. 3. This relation, between molecular interactions and protein stability and solubility, is extensively investigated and experimentally supported for many different protein systems [27, 29, 31, 34-36, 39, 40, 43-63].
Figure 3. Relation between molecular interactions and the standard chemical potential of proteins (shown for both the N and D conformations). Modified after [39]

Figure 3 shows the general picture of solute effects due to changes in the standard chemical potential of a protein in its different states. The dashed red line illustrates that adding a solute which interacts favorably with the native protein decreases the chemical potential of the protein ($\mu^\circ_N$), and hence increases the solubility. In contrast the dashed blue line illustrates the precipitating properties of preferentially excluded solutes (increased $\mu^\circ_N$). The effects on protein stability are indicated by the length of the solid lines (the difference between the chemical potentials of respectively the native and denatured states). When $\Delta\Delta\mu^\circ$ ($\Delta\mu^\circ_{\text{solut}} - \Delta\mu^\circ_{\text{ref.}}$) < 0 the protein is destabilized and when $\Delta\Delta\mu^\circ$ > 0 the protein is stabilized.

Solvent composition and protein aggregation

As described above proteins can be stabilized through the addition of excipients. However, under some conditions denaturation does not progress reversibly for example because of aggregation of denatured monomers. In these cases, it is challenging to establish a framework for the understanding of solute effects. Irreversible aggregation is, as mentioned in the introduction, a significant problem in the development of protein products and even after many years of scientific investigations there is still great demand for a better
understanding of this process. Generally, to achieve reversible unfolding conditions, the temperature has to be moderate and the proteins have to be small (< 30 kDa), dilute (< 1 mg/ml) and highly charged [64], (additionally protein glycosylation effectively suppress irreversible aggregation, which will be discussed in Part II of this report). Thus, in many cases protein denaturation is irreversible and kinetic parameters are used as measures of stability (the kinetic stability).

In this section protein aggregation will be discussed with respect to solvent composition, and the molecular picture of the process will be interpreted as a two-steps process (illustrated in equation 3a and 3b) originally described by Lumry and Eyring [65]. The first step (eq. 3a) is the reversible transition from the native conformation to a denatured (aggregation competent) conformation. The second step (eq. 3b) is a growth step where denatured monomers adsorb to already existing aggregates consisting of n monomers, A_n.

Transition step: Native (N) $\xrightarrow{\text{\rightleftharpoons}}$ Denatured (D) \hspace{1cm} (3a)

Growth step: D + A_{(n)} \rightarrow A_{(n+1)} \hspace{1cm} (3b)

Many observations suggest that protein aggregation is not controlled by random molecular associations but by specific interactions, usually $\beta$-sheet formations [66-71]. This means that aggregation is controlled by the amino acid sequence, and that aggregates of a given protein probably have relative similar structures regardless of the factors promoting the aggregation process [66].

It is well established that protein aggregation is strongly dependent on physical factors such as solvent composition and temperature. Aggregation rates are increased at high protein concentrations because of increased probability of molecular collisions [72, 73] and at high temperatures, primarily because of increased concentration of denatured proteins [74-76]. Furthermore aggregation is promoted in solutions with pH values close to pI, because of decreased electrostatic repulsions between monomers [75, 77, 78]. Solutes (stabilizing and perturbing agents) also affect the kinetics of protein aggregation [73, 79-84]. From most studies it is however unclear if the effects of solutes primarily rely on the equilibrium shift (eq. 3a), or on the aggregate growth step (eq. 3b). As it is usually only denatured (unfolded or partially unfolded) proteins which associate to form aggregates [66], stabilizing solutes
might impede aggregation by lowering the concentration of denatured proteins (i.e. the reactant in eq. 3b). However, the stabilizing effect of non-specifically interacting solutes is driven by thermodynamically unfavorable interactions between solutes and proteins (preferential exclusion of solutes) which also favor the formation of inter protein contacts (as discussed above). In other words, the unfavorable protein-solute interactions will destabilize the monomer D state because it has the highest surface area (giving rise to a larger number of unfavorable interactions). This will stabilize N over D, but also promote association of denatured molecules once formed. In contrast perturbing agents favor the monomer denatured state. In the light of this suggested dual effects of solutes, also discussed in Paper 4, studies on possible interrelationships between preferential interactions and aggregation kinetics appear to be of interest.

To elucidate potential effects of solutes on the aggregation kinetics, the transitions state theory may be useful. This theory states that in a kinetically controlled reaction there is a transition state (activated complex) which has the highest free energy and which is in equilibrium with the reactants. Thus, if the growth step (eq. 3b) is assumed to be rate limiting, it may be rewritten as in eq. 4, where $D-A_{(n)}^{\ddagger}$ denotes the transition state.

$$
D + A_{(n)} \xrightarrow{\ddagger} D-A_{(n)}^{\ddagger} \rightarrow A_{(n+1)}
$$

As for all equilibrium processes the equilibrium constant and the standard free energy can be expressed as follows: $K^{\ddagger} = [D-A_{(n)}^{\ddagger}]/[A_{(n)}][D]$ and $\Delta G^{\ddagger} = -RT \cdot \ln(K^{\ddagger})$

At any given temperature it is only the concentration of activated complexes that determines the reaction rate [85]. Therefore the effect of solutes on aggregation rates depends on their changes in $K^{\ddagger}$, and hence in standard free energy, $\Delta G^{\ddagger}$. Large $\Delta G^{\ddagger}$ values slow down the process and rapid reactions correspond to small $\Delta G^{\ddagger}$ values. Figure 4 illustrates how the standard free energy ($G^{\ddagger}_{\text{system}}$) changes with the reaction coordinate of the aggregation process. $G^{\ddagger}_{\text{system}}$ is represented by a red line (Fig. 4), and the dashed blue line illustrates how addition of a solute may suppress the rate of the aggregation process (increasing $G^{\ddagger}_{\text{system}}$). If for example the solvent exposed surface area is larger for $D-A_{(n)}^{\ddagger}$ than
for \((D + A_{(n)})\) addition of a preferentially excluded solute may change the standard free energy curve from the black to the blue example in Fig. 4.

**Figure 4.** The standard free energy of the system (black curve) is illustrated as a function of the reaction coordinate according to the transition state theory. The change in standard free energy, \(\Delta G^{\ddagger}\), between the reactants \((D \text{ and } A_{(n)})\) and the transition state \((D-A_{(n)}^{\ddagger})\), is represented by the red line. The dashed blue line illustrates an increased \(\Delta G^{\ddagger}\) corresponding to kinetically stabilization e.g. by a solute.

To elucidate the interrelationship between preferential interactions and protein aggregation we have investigated heat induced aggregation of BSA in solutions containing sorbitol, sucrose and \(\text{Na}_2\text{SO}_4\) (preferentially excluded at room temperature [39, 47-51]) as well as in solutions of urea and \(\text{NaSCN}\) (preferentially bound at room temperature [39, 48, 60, 86]). This study is presented in Paper 4 with focus on aggregation mechanisms and kinetics. Here only part of the experimental work is presented. To single out the solute effects on the growth step (eq. 3b), all experiments are made at the denaturation temperature of BSA \((T_d K = 1\), where \(K\) is the equilibrium constant for eq. 3a). Thereby temperature differences compensate for the solute effect on the equilibrium constant of the transition step (eq. 3a). The \(T_d\) values of BSA in the eleven different solvents \((T_d \sim 68^\circ\text{C}-74^\circ\text{C})\) were determined by means of differential scanning calorimetry (DSC) as described in Paper 4.
Due to varying experimental temperatures one limitation of this work is that the normal Arrhenius effect (of ~6°C) on the rate is neglected.

As the intensity of light scattered from a solution depends on the molecular weight (and size) of solubilized particles the kinetics of molecular associations can be followed by monitoring the light scattering (LS) as a function of time. All protein species contribute to the scattering intensity as described in eq. 5.

\[
I = k \cdot ([N] \cdot mw^2 + [D] \cdot mw^2 + [A] \cdot (n \cdot mw)^2) \quad (5a) \\
I = k \cdot mw^2 ([N] + [D] + [A] \cdot n^2) \quad (5b)
\]

where \( I \) is the scattering intensity at one given angle, \( k \) is a constant specific to the investigated system and experimental details. \([N], [D] \) and \([A] \) respectively denotes molar concentrations of native monomers, denatured monomers and aggregates, \( n \) is the aggregation number and \( mw \) is the molecular weight of the protein.

The protein concentration was identical in all LS trials, \([\text{BSA}] = 1.04 \, \text{mg/ml}\). The solvents contained 20 mM succinate buffer, pH 7.0 and varying solutes and solute concentrations. The scattering intensities for the eleven investigated BSA solutions was normalized according to initial (\( t = 0 \)) scattering (in Fig. 5) and thus represent the relative increase in LS as a function of time. The Normalized Scattering Intensity (NSI) is determined as:

\[
\text{NSI}_t = \frac{I_t - I_b}{I_0} 
\]

where \( I_t \) and \( I_0 \) are the scattering intensity from the protein solution at time \( t \) and at \( t = 0 \) respectively, and \( I_b \) is the background intensity (scattering from pure solvent).
Figure 5. Normalized Scattering Intensities (NSI values) as function of time for BSA in different solutions. The added solutes are referred to by different colors. Broad lines denote the highest concentrations: Urea (green) 0.74M and 0.37M, sorbitol (red) 0.62M and 0.31M, sucrose (pink) 0.58M and 0.29M, Na$_2$SO$_4$ (blue) 0.44M and 0.22M, NaSCN (yellow) 0.73M and 0.36M. The black curve represents BSA in pure buffer. BSA was dissolved in 20 mM succinate buffer, pH 7.0 to a weight concentration of 1.04 mg/ml in all solutions. The experimental temperature was varied to equal the $T_d$ values (68-74°C). The wavelength of the laser light was 532 nm, and only the photon count rates recorded at 90° scattering angle were used. The measurements were performed using a home built instrument on the Faculty of Life Sciences, University of Copenhagen, Denmark.

As light scattering depends on the aggregate growth, the NSI curves in Fig. 5 clearly show an effect of additives on this process. Different colors refer to different solutes and broad lines denote the highest concentration. The solute concentrations are: Urea 0.74M and 0.37M, sorbitol 0.62M and 0.31M, sucrose 0.58M and 0.29M, Na$_2$SO$_4$ 0.44M and 0.22M, NaSCN 0.73M and 0.36M. Figure 5 shows a time interval (lag time) of about ~2-3 minutes where no particle growth can be detected (NSI ~ 1). This is suggested to account for the time required to heat the protein solutions and to initiate aggregate nucleation (particle formation). Further, Fig. 5 shows that all additives limit the increase in NSI values. Thus, the scattering data may indicate that all solutes suppress BSA aggregation. However as the intensity of
scattered light scales with the square of the aggregation number (eq. 5b) large aggregates will contribute much more to I_t (eq. 6) than smaller ones, and therefore, NSI values may not always be good for comparisons of aggregation extents in different solutions (particularly not far from initial values t >> 0). For example, solutes favoring many small aggregates may, by this method, seem more effective in preventing aggregation than solutes favoring a smaller number of larger aggregates.

According to the transition state theory solute effects on aggregation rates can be quantified due to their effects on $\Delta G^{\ddagger}$ (for example as illustrated by the dashed blue line in Fig. 4). To quantify the effects of the different solutes in this work, determine $\Delta \Delta G^{\ddagger}$ values, we use the initial slopes (subsequent to the lag time) of the curves in Fig. 5. It is assumed that the process follows first order kinetics, which has been found for temperature induced aggregation of other proteins [68, 87-90], and that initial $d$NSI$/dt$ values are proportional to particle growth rate. Thus the aggregation rate ($r$) can be expressed as follows:

$$ r = - \frac{d[D]}{dt} = - k \cdot [D] \approx \frac{d$NSI$}{dt} $$ \hspace{1cm} (7)

where $k$ represent the rate constant. $\Delta \Delta G^{\ddagger}$ can be determined from the relation between rate constants as shown in eq. 8, where $k_0$ and $k_S$ represents the rate constants of the aggregation processes in respectively the pure buffer and in the solute system. As the aggregation is induced at $T_d$ the concentration of denatured protein is suggested to be the same in all solutions at $t \sim 0$. Therefore the initial $d$NSI$/dt$ values (eq. 7) are used directly in eq. 8 to determine $\Delta \Delta G^{\ddagger}$. An analogous approach to the analysis of solute effects on protein aggregation has previously been used by Eronina et al. 2005 [83].

$$ \Delta \Delta G^{\ddagger} = -RT \cdot \ln(k_S/k_0) \approx RT \cdot \ln \left( \frac{(d$NSI$/dt)_0}{(d$NSI$/dt)_S} \right) $$ \hspace{1cm} (8)

The change in standard free activation energies are shown as a function of the solute concentration in Fig. 6. Noteworthy is the large positive $\Delta G^{\ddagger}$ values for the two salts compared to the non-electrolytes. $SO_4^{2-}$ (a strong kosmotrope) and SCN$^-$ (a strong chaotrope)
have opposite effects with respect to changes of surface potential in air-water interfaces, protein solubility and equilibrium stability [91]. Nevertheless, they show relatively equal effects on the kinetics of BSA aggregation. This may indicate that the transition state formation is controlled by electrostatic interactions (affected by increasing ionic strength), and that the “Hofmeister effects” on protein solubility are of minor importance in this respect.

The effects of sucrose, sorbitol and urea on the kinetics of BSA aggregation corresponds to positive $\Delta \Delta G^{\ddagger}$ values, but these are much weaker than the effects of salts (Fig. 6). Relative similar effects are found for the three non ionic solutes. The results indicate, however, that preferentially excluded additives repress aggregation rates more at high concentrations, whereas urea has a more pronounced effect at the low concentration. One suggestion concerning similarities between the effects of the three different non ionic components could be that sucrose and sorbitol slightly compress the conformation of the D-state (by being preferentially excluded [92]), and that the preferential accumulation of urea disturbs intermolecular peptide interaction. Both such perturbations can increase the standard free energy of the transition state.

![Figure 6](image.png)

**Figure 6.** The difference in standard free activation energy changes $\Delta \Delta G^{\ddagger}$ of BSA aggregation is shown as a function of the solute concentration for sucrose, urea, sorbitol, $\text{Na}_2\text{SO}_4$ and $\text{NaSCN}$. The reference process is BSA aggregation in pure buffer, and the values are determined according to eq. 8.
It appears from the present BSA studies (DSC and SLS) that solutes affect both steps in the aggregation process (eqs. 3a and 3b). However, these studies (which are more thoroughly treated in Paper 4) further show that non-specific “Hofmeister-type” interactions can be of minor importance for the irreversible step in the aggregation process. Compared to the effects of solutes on equilibrium stability and solubility, generalities are very difficult to establish for protein aggregation. This may be explained by the many facets of protein aggregation, and the potential opposing effects of solutes on the two different steps in the aggregation process (denaturation and molecular association).
Part II

Effects of Protein Glycosylation

Protein glycosylation *in vivo*

Protein glycosylation is the most common posttranslational modification of proteins in eukaryotic cells. This covalent addition of sugars to the peptide chain takes place in the endoplasmic reticulum (ER) or Golgi apparatus. In ER preformed oligosaccharides, consisting of two N-acetylglucosamine residues and several sugar residues primarily mannose (GlnNAc₂-Manₓ), are attached to asparagines at specific glycosylation sites with the sequences Asn-X-Ser or Asn-X-Thr. These glycans are called asparagines linked (or N-linked) high mannose oligosaccharides, and are co-translationally attached to the growing polypeptide chain. The two GlnNAc groups create the bridge between the peptide part and the sugar residues of glycoproteins. N-linked high-mannose oligosaccharides can be modified to more complex oligosaccharides in the Golgi apparatus. In addition O-linked glycosylation (carbohydrates linked to OH groups of selected serine or threonine side chains) occurs in the Golgi apparatus. [93, 94]

The glycosylation pathway is complex and highly conserved. The biological significance of protein glycosylation clearly appears from the very broad and numerous physiological effects of underglycosylation, which in worst cases is lethal [95]. As each glycoprotein represents a unique system the molecular mechanisms underlying the numerous biological purposes of glycosylation are not specifically defined, but some overall effects have been accentuated. One important role of the glycans appears to be molecular recognition (e.g. cell-cell recognition) and sorting such as extracellular transportation. In general cytoplasmic proteins are non-glycosylated (or limited glycosylated) while secreted and membrane proteins are glycosylated. Many proteins therefore occur *in vivo* both in a
glycosylated (external) and a non-glycosylated (internal) form in the same organism [93, 94]. Furthermore the presence of these carbohydrate chains on the protein surface prevents protein aggregation [66, 96] (in vitro aggregation is discussed below), and effectively reduces the degradation by proteolytic enzymes [94, 97, 98].

Phytase and the investigated glyco variants

Phytases (myo-inositol hexakisphosphate phosphohydrolase) are found naturally in plants and microorganisms, particularly fungi. Many different phytases of fungal origin have been cloned, engineered, expressed and purified, and some of them also formulated into industrial products. Except if expressed in bacteria phytases are glycosylated with glycosylation patterns depending on the host organism. The industrial interest in this enzyme relates to the catalysis of the hydrolysis of phosphomonoester bonds of phytic acid. Phytic acid is the major storage form of phosphorus in plants. Monogastric animals such as pigs and poultry are unable to degrade this compound, and hence unable to extract free phosphate from e.g. seeds, cereal and legumes in animal food. Addition of phytase in animal feed preparations therefore reduces the need for supplement of inorganic phosphate and hence reduces the environmental pollution [99-103].

For our studies we use different glyco variant of Peniophora lycii phytase. One variant, which in this work is referred to as Phy, is expressed in the fungus Aspergillus oryzae. This enzyme was supplied from Novozymes subsequent to purification according to procedures described by Lassen et al. 2001 [100]. Phy is composed of a 47.5 kDa peptide part (439 amino acids) and a carbohydrate part of about 18 kDa, which consists of high-mannose oligosaccharides distributed among 10 N-glycosylation sites. Thus, this phytase (Phy) consists of ~ 30% (w/w) carbohydrates. Another variant is the deglycosylated form of Phy, which is referred to as dgPhy. To produce dgPhy the enzyme (Phy) is enzymatically deglycosylated by Endo F₁ (Endo-β-N-acetylglucosaminidase; EC 2.2.1.96). Endo F₁ hydrolyses the glycosidic bond between the first and the second N-acetylglucosamin groups on each glycan. Hence deglycosylation gives a phytase which consists of the same peptide part as Phy and about 10 GlnNac residues remaining at the glycosylation sites after the deglycosylation (~4% carbohydrate). This deglycosylated enzyme has the same activity as the glycosylated form, Paper 1.
Additionally we have purified an over glycosylated variant of *Peniophora lycii* phytase expressed in yeast *Saccharomyces Cerevisiae*. This glycoprotein is referred to as **ogPhy**. The culture broth (∼ 5L), which was supplied from Novozymes, was filtered through filtration cloth and then through filter plates. The filtrate was concentrated by ultrafiltration and then analyzed for phytase activity. The concentrate, containing phytase, was dialysed to reduce the ionic strength (conductivity ∼ 1.7 ms/cm) and adjusted to pH 6.5 by NaOH. The enzyme solution was then purified according to the procedures described by Lassen *et al.* 2001 [100]. SDS page on the purified heavily glycosylated phytase from yeast revealed a broad band corresponding to a molecular weight of ∼100-130 kDa. This corresponds to a glycosylated phytase (ogPhy) containing ∼ 50-60% carbohydrates. The obtained amount of ogPhy (dialysed and lyophilized) was 8 mg.

**Biophysical effects of protein glycosylation**

The presence of carbohydrates on protein surfaces affects self association as well as interactions with other solvent components, which leads to changes in biochemical and biophysical protein properties. Such glyco effects have been subject to numerous investigations during the last thirty years. Analyzing the literature within this field reveals some important differences between different glycoprotein systems, which make comparisons of glyco effects difficult. The glycoprotein systems can in principle, and for practical reasons, be divided into three groups:

1. **Naturally glycosylated protein**
   (The wild type protein holds glycosylation site/sites)

2. **Mutated glycoproteins**
   (The peptide chain are mutated to hold glycosylation site/sites)

3. **Chemically glycosylated proteins**
   (Carbohydrates or synthetic “glycans” are chemically attached to protein surfaces)

In the attempt to find generalities in the effects of glycosylation, distinctions between these groups may be necessary. Such distinctions have to our knowledge never been pointed out, and might be of importance in relation to the industrial application of surface modification. In the following the biophysical effects of glycosylation will be reviewed on
the basis of investigations of different glycoprotein systems. This includes our studies of the naturally glycosylated protein *Peniophora Lycii* phytase. Molecular interactions, equilibrium stability and protein aggregation are the main focus in this section.

- Effects of glycosylation on equilibrium stability

As described in the introduction native proteins are only marginally stable with limited lifetimes due to different mechanisms of structural changes. Glycosylation are known to stabilize proteins towards proteolytic digestion [97] and irreversible aggregation (see below). Relationships of glycosylation and equilibrium protein stability however, appear to depend on the specific system.

*Mutated and chemically glycosylated proteins.* For proteins which are not naturally glycosylated the stability changes induced by glycosylation dependent on the glycan position and on the degree of glycosylation. For example, glycans induced via peptide mutations might interfere sterically with protein folding and thereby induce large conformational changes or just slightly destabilize the native structure. Conversely some glycosylated mutants have been shown to be somewhat more stable than their non-glycosylated parent molecules. For example, introduction of N-glycosylation signal sequences in lysozyme (and following glycosylation of this enzyme) has shown to increase its resistance towards denaturants [104]. It has, however, not been established whether the stabilizing effect of the glycans reflects adsorption of the denaturants or an actual intrinsic stabilization of the peptide conformation. In general this aspect has not been discussed by workers addressing the stabilizing effects of glycosylation, which counts reports of different glycoprotein systems. In some cases introduction of oligosaccharides via peptide mutations have been shown to reduce the thermal stability of proteins [105, 106]. However single site mutated proteins with increased thermal stability has been found [107]. In general the larger number of induced glycosylation sites the less thermodynamically stable proteins. To our knowledge thermodynamic stabilization by introduction of more than one glycosylation site has not been reported. Another problematic effect of imperfect glycosylation is loss of protein functionality because of interference between glycans and for example enzyme-substrate interactions [104, 106, 107]. Such activity reducing effects of glycans have also been found for naturally and chemically glycosylated enzymes [21, 97].
For chemically glycosylated proteins increased thermodynamic stability has been reported [12, 108]. Most successful chemical glycosylation occur under mild conditions (moderate pH values and low temperatures) where proteins remain native during the process, and hence glycosylation of buried amino acids is avoided. For protein-lactose and protein-dextran conjugates it has been shown that the stabilization is more pronounced with increasing glycan-content [12]. However, for other protein-carbohydrate conjugates conformational stabilization has been less clear [109]. Even though mechanistic details of possible stabilizing effects is unclear, chemical glycosylation has been suggested as a promising approaches for thermodynamic stabilization of proteins [12, 13]. As mutated glycoproteins are often found to be less stable than the un-glycosylated wild type we suggest that peptide changes are a critical point in the attempt to increase thermodynamic stability of proteins by glycosylation.

**Naturally glycosylated proteins.** In contrast to mutated proteins, and perhaps not surprising in an evolutionary perspective, the glycans of naturally glycosylated proteins do not exhibit negative effects on protein equilibrium stability. Nevertheless the relationships of glycosylation and equilibrium protein stability for this group of glycoproteins also seem to depend on the investigated system. Some studies address stabilizing effects of glycosylation [110-114] and some address no, or only limited effects on the protein stability [115-122]. Judging from the effects of glycosylation on protein aggregation (discussed in the next section), glycans seem to disfavor protein-protein interactions. Such unfavorable interactions (preferential exclusion) might increase protein stability, particular at high concentrations. From papers addressing stabilizing effects of glycosylation it is often unclear if irreversible denaturation can be neglected under the experimental conditions, and hence, what appears to be an increased thermodynamic stability may be a reflection of the aggregation preventing effects of glycans.

Molecular interpretations of proposed stability enhancements (induced by glycans) are generally missing in the literature. But one suggested explanation is the presence of hydrogen bonds between surface amino acids and glycans, which has been found for crystallized RNase B [123] and glucose oxidase [124]. Such hydrogen bonds have the potential to increase the stability of proteins (if for example they exists only for the native protein conformation), and have been suggested as a more general phenomenon for naturally glycosylated proteins [114]. However these bonds might dependent on physical factors such as temperature and water content. It is for example questionable if they are of significant
strength in the presence of excess water, where the acceptor and donor of the broken bond can interact equally with $H_2O$.

We investigated three different glyco-variants of Peniophora lycii phytase by means of DSC. Panels A, B and C in Fig. 7 demonstrate the thermal denaturation of respectively Phy (~30% carbohydrate), ogPhy (~60% carbohydrate) and the deglycosylated form, dgPhy (~4% carbohydrate). It appears that glycosylation of this enzyme also confers limited effects on the denaturation temperature ($T_d$). $T_d$ values of respectively Phy, ogPhy and dgPhy are: 61.9°C, 61.3°C and 60.8°C. In previous studies of Phy and dgPhy we found slightly different $T_d$ values and slightly larger differences between the denaturation temperatures of Phy and dgPhy, Paper 1. Thus, we can conclude that the difference between the three glyco variant is small compared to effects specific to factors such as buffer concentration and pH. The limited effects of glycans on the thermodynamics of protein folding are further indicated in Fig. 7D. The enthalpy of denaturation, which is given by the area under the transition peak, is practically equal for the three glyco variants. $\Delta H_d$~850 kJ/mol for phy, and slightly lower for ogPhy and dgPhy. No significant difference in heat capacity changes ($\Delta C_p$) is obtained from these DSC scans ($\Delta C_p$~ 20 kJ-mol$^{-1}$-K$^{-1}$). The heat capacity change is considered to be a sensitive measure of hydration effects, particularly of unpolar groups. Thus the very similar $\Delta C_p$ values between the different glyco variants reveals that the glycans do not contribute to measurable changes in the protein-solvent interactions upon denaturation of phytase. Equal $\Delta C_p$ values of glycosylated and deglycosylated proteins is also reported for the unfolding of three different heavily glycosylated proteins [110]. In general enthalpy and heat capacity changes are highly sensitive to changes in intermolecular interactions (much more so than $T_d$), and it is remarkable that the massive covalent changes do not bring about more significant changes. These observations, thus strongly suggest that glycan-peptide and glycan-solvent interactions remain practically unchanged upon denaturation of glycoproteins. Although no direct evidence is available the most straight forward interpretation of this is that the carbohydrate moieties are fully exposed to the solvent in the native and denatured states, and that they do not affect the peptide structure of these conformational states. This will be discussed later as an important explanation of the limited effects of glycans on the thermodynamic stability of glycoproteins.
Figure 7. DSC traces for the thermal denaturation of A: Phy (~30% carbohydrate), B: ogPhy (~60% carbohydrate) and C: dgPhy (~4% carbohydrate). The three glyco-variants were dissolved in 20 mM Na-acetate buffer, pH 5.0 to a weight concentration of 1.2 mg/ml. The scan rate of all scans was 2°C/min, and a pure buffer scan is subtracted the presented $C_p$ curves. In Panel D the $C_p$ curves are modified relative to the molar concentration. The DSC measurements were performed using a SCAL-1 micro calorimeter (Puschino, Russia).

The decrease in heat capacity subsequent to denaturation of the deglycosylated enzyme (Fig 7, Panel C), is due to the exothermic aggregation of denatured proteins, which is prevented by the presence of glycans (discussed in next section). In a number of DSC trials
conducted as in Fig. 7 we did not observe refolding within the time course of these experiments. This DSC instrument does not allow for holds in the temperature scanning circles, thus after fast cooling from 90°C to 4°C heating is immediately continued. In other DSC studies (using another type of instrument) we found that denaturation of Phy by heating to 80°C was reversible to some degree (~80% native proteins was obtained after 2 hours at 10°C), whereas refolding of dgPhy was not detected by rescanning (data not shown).

In accordance with results for many naturally glycosylated proteins [115-122] this DSC study of phytase suggests that glycosylation is relatively ineffective with respect to thermodynamic stabilization of proteins. Structural investigations of naturally glycosylated proteins (including phytase, Paper 3) reveal that the peptide structures are not affected by the carbohydrates [97, 110, 114, 123, 125], and hence, that the folding information is contained exclusively in the peptide moieties. In the light of the importance of protein-solvent interactions for the thermodynamic stability (see Part I), it is quite remarkable that the stability of glycoproteins are not significantly different from un-glycosylated forms as protein-solvent interactions are markedly changed upon glycosylation [126], Paper 1, Paper 3. For example, analysis of the relative affinity of the glycan and peptide moieties of phytase suggested that the carbohydrates bind much less surfactant, Paper 3. However, it was also found that deglycosylation of Phy brings about only a small decrease in the resistance towards SDS induced denaturation. Other studies reveal that the deglycosylation has also limited consequence for the effects induced by stabilizing components, Paper 1. In this work two common stabilizers (sorbitol and glycerol) were found to be preferentially bound to the glycans of phytase (and excluded from the peptide part of the enzyme). The respective stabilizing effect (increased $T_d$ values) of sorbitol and glycerol is, however, identical for Phy and dgPhy.

We therefore suggest that the limited stabilizing effect of glycosylation is a consequence of equivalent solvent exposure of glycans in respectively the N and the D state. Most likely, this reflects that the carbohydrates are fully exposed to the solvent in either state. If, for all interacting solvent component, the glycans contributes equally to $\Gamma^N_{\mu 3}$ and $\Gamma^D_{\mu 3}$ (see Wyman theory eq. 2) the carbohydrate moiety will contribute equally to the chemical potential of the native and the denatured state and hence, will not affect the thermodynamic stability. In this case glycosylation will also not alter the stability changes induced by solutes (stabilizers or perturbing agents).
Interestingly it has been suggested that the dynamic stability of proteins is increased by glycosylation. This means that the frequency and extent of molecular fluctuations is suppressed by the glycan mantle. The exchange rates of partially buried amid proton with solvent (D\textsubscript{2}O) determined by NMR are markedly higher for RNase A (non glycosylated) than for the single site glycosylated form, RNase B [97, 125, 127]. As no structural changes can be observed upon deglycosylation of this enzyme, the decreased hydrogen-deuterium exchange rates indicate a glycan induced dampening effect on molecular fluctuations. Increased dynamic stability has been suggested as a possible general phenomenon for glycosylated proteins [114, 117] as well as for chemical glycosylated proteins [108]. Thus even though the folding information is contained exclusively in the peptide moiety the intramolecular interactions within the peptide structure might be affected by the presence of the surface attached carbohydrates. The relationship between dynamic and thermodynamic stability is not straight forward since reduced fluctuations will change both the enthalpic and entropic contributions to \(\Delta G^\circ\). Nevertheless the dampening effects of glycans on molecular fluctuations are suggested to be correlated to the increased stability observed for some glycoproteins [114]. Considering the stabilizing effects of drying, freezing and addition of preferential excluded solutes (increasing the surface tension in the protein solvent interface), which also reduces molecular fluctuations, this suggestion might be plausible. However, Tams and Welinder has suggested that the dampening effect of glycosylation can reduce the refolding rate of denatured proteins to the same extend as the unfolding rate, and thus induce no effect on the equilibrium stability [118].

- **Effects of glycosylation on protein aggregation**

As described in Part I denatured proteins may associate into aggregates, which make the denaturation process irreversible. Hence, the equilibrium protein stability is not always appropriate to describe critical aspects of protein applications, and the kinetic stability (e.g. the aggregation rate) is therefore a frequently used stability parameter. In contrast to the effect on equilibrium stability the carbohydrate part of glycoproteins has a pronounced effect on the solubility, and very effectively suppresses irreversible aggregation and surface adsorption [12, 78, 105, 106, 108-112, 117, 126, 128-134]. Some results have shown that the most distinctive effects are found when the glycosylation sites are placed
adjacent to hydrophobic sequences [111, 135]. However the increased solubility and decreased aggregation tendency brought about by glycans might not be a result of a particularly hydrophilic nature of glycoproteins. For example, preferential binding parameters (eq. 1) reveal an increased preferential hydration of dgPhy, relative to Phy, in solutions containing glycerol and sorbitol, Paper 1. This shows that water, relative to those polyols, interacts more favorable with the peptide moiety than with the glycans of this enzyme. Another interesting aspect of these results is that the precipitating effects of stabilizing solutes (Fig. 3, Part I) might be counteracted by protein glycosylation. We found that glycerol and sorbitol are preferentially excluded from the peptide part, and that they stabilized Phy and dgPhy towards heat induced denaturation. In addition our results suggested preferential binding of those solutes to the glycan moieties of Phy. Such favorable interactions between solutes and glycans increase the solubility of the protein. Thus glycosylation is suggested to increase the compatibility of proteins in solutions containing stabilizing components by counteracting the precipitating property of these solutes.

To clarify the hygroscopicity of glycans relative to peptides we investigated dry samples of Phy and dgPhy, Paper 2. The enzymes were exposed to vapor of varying relative humidity (RH: 0-98%) and the water uptake and binding enthalpy was quantified simultaneously. Binding isotherms of Phy and dgPhy (presented in Paper 2) reveal that the mass specific water uptake is identical for Phy and dgPhy in the RH interval ~ 0-50 %. Increasing humidity showed increasing difference between Phy and dgPhy, and at RH ~ 70-98 % the hygroscopicity is significantly higher for the deglycosylated enzyme. This shows that compared to peptides there are no particularly favorable interactions between water and glycans. Suggesting that we can compare fully hydrated proteins with those in solution, the results indicate that the solubilizing effect of glycosylation is unlikely to be due to particularly hydrophilicity of the glycans. This is in accord with second virial coefficients ($B_{22}$) of Phy and dgPhy determined by static light scattering (SLS) [78]. The authors found larger $B_{22}$ values for dgPhy, which indicates that deglycosylation leads to more favorable protein-solvent interactions. We therefore suggest that protein glycosylation most probably reduce molecular associations due to steric hindrance of close packing (and not due to a particular hydrophilic nature of glycans). An analogous mechanism is suggested to explain reduced aggregation of protein-dextran conjugates [12].
To investigate the effects of glycans on protein aggregation we have studied Phy, dgPhy and ogPhy at elevated temperatures by static light scattering. From these studies (which are not included in the attached manuscripts) it is found that glycosylation clearly suppresses the kinetics of heat induced protein aggregation (Fig. 8). At time t=0 the cold (~4°C) phytase samples are placed in the preheated water bath. The experimental temperature corresponds to the denaturation temperature of the investigated glyco variant: 61.9°C, 61.3°C and 60.8°C for Phy, ogPhy and dgPhy respectively (obtained by DSC, see Fig. 7). In the first minute no increase in the scattering intensity can be detected, this period is referred to as the lag time. The lag time for dgPhy aggregation is about 1 minute, for Phy and ogPhy it is ~ 2 minutes. Our studies show that aggregation rates are very sensitive to glycosylation and it is therefore difficult to investigate the process under equivalent conditions. In these studies the molar protein concentration is ~13 μM, the temperature = T_d and pH = 5.0 (pI for Phy and dgPhy is 3.6 [78], not determined for ogPhy). The time course of the experimental trials was however significantly varied. The dgPhy aggregation process was stopped by rapid cooling after 10 minutes, because the LS detection indicated large particles and beginning precipitation. In contrast Phy and ogPhy trials were continued for 50 minutes; nevertheless the scattering intensities in these trials were radically lower compared to scattering from the dgPhy solution. In Fig. 8 the scattering data is represented by the Normalized Scattering Intensities (eq. 6).

From these scattering experiments we find that deglycosylated enzymes associate very fast relative to glycosylated variants of the same protein. After 2 minutes the intensity of scattered light from the dgPhy solution is higher than any detected scattering from solutions of the glycosylated proteins (Phy and ogPhy). The initial slope (~ first minutes after the lag time) of NSI curves indicates that the rate of the particle growth is about 130 times faster for dgPhy than for ogPhy and about 30 times faster than for Phy. Using the same interpretations as for the solute effects on BSA aggregation (see Part I) these values corresponds to an increased standard free activation energy (ΔΔG°‡) of about ~ 9 kJ/mol upon addition of the glycan amount of Phy, and about 14 kJ/mol upon addition of the ogPhy glycans. During these studies we frequently found conditions where no detectable aggregation of Phy and ogPhy was observed within hours. In contrast significant aggregation of dgPhy was observed under all investigated condition (e.g. at 40°C and pH = 7.0). Our SLS results, which show that glycosylation significantly decreases the rate of aggregation, are
consistent with other investigations of *Peniophora Lycii* phytase. Høiberg et al. investigated the effects of pH changes on the heat induced aggregation of Phy and dgPhy by SLS [78]. They found strong pH dependencies, particularly on the aggregation of dgPhy. Their results thus suggested that specific interactions between charged amino acids are involved in the aggregation process and that the glycans might sterically hinder this mechanism.

**Figure 8.** Normalized scattering intensities as function of time for the three glyco variants: Phy (black), ogPhy (green) and dgPhy (blue). The left and right figures respectively show different time intervals of the light scattering detection. The three proteins were dissolved in 20 mM Na-acetate buffer, pH 5.0 to a weight concentration of 0.84 mg/ml for Phy, 1.3 mg/ml ogPhy and 0.61 mg/ml for dgPhy (corresponding to about the same molar concentration of the three solutions). The experimental temperature was varied to equal the $T_d$ values: 61.9°C, 61.3°C and 60.8°C for respectively Phy, ogPhy and dgPhy. The wavelength of the laser light was 532 nm, and only the photon count rates recorded at 90° scattering angle were used. The measurements were performed using a home built instrument on the Faculty of Life Sciences, University of Copenhagen, Denmark.

The light scattering data for Phy and ogPhy (Fig. 8) show that there is a slow increase in NSI values, which is further slowed down after about the first 5-10 minutes. This is due to a very slow particle growth. To clarify if this is caused by effective preservation of monomers, or if it is due to restrictions of the particle size (number of monomers in each aggregate) we studied the aggregated samples by combined size exclusion chromatography.
and refractive index (RI) measurement. The chromatographic resolution of these trials is not optimal, possible due to interactions between the proteins and the column material. The RI signals are shown in Fig. 9A for the native samples (fresh made and not exposed to high temperatures). The retention time is equal for Phy and dgPhy (∼26 minutes), and somewhat lower for ogPhy (∼23 minutes) which shows that the relation between molecular weight and retention time differ between the different glyco variants. The RI chromatograms of the heated samples are shown in Fig. 9B. The two peaks representing respectively aggregates and monomers are not completely separated on the column, indicating interactions with the column material and perhaps some degree of aggregate polydispersity. In spite of these limitations, the chromatographic data does elucidate the depletion of monomeric protein upon heat exposure. Comparisons of the RI signals in Fig. 9A and 9B reveal a significant decrease in the monomer concentration of all glyco variants due to aggregation. As expected from the light scattering data (Fig. 8), the reduction in monomer concentration is more pronounced for the dgPhy solution. As the dgPhy sample was heated for only 10 minutes, compared to 50 minutes for the glycosylated enzymes, this clearly indicates the potential of glycans to suppress aggregation. However, the chromatographic data reveals significant reduction of monomeric Phy and ogPhy after 50 minutes at T_d (Fig. 9). Thus, the considerable differences between NSI values of glycosylated and deglycosylated phytases (Fig. 8) indicate restriction in aggregation number of glycosylated proteins. We suggest that glycans due to steric restrictions of close packing, and hence, hindrance of strong intermolecular interactions, effectively reduces the number of monomers in each aggregate. To determine actual aggregation number of these glyco variant more studies will be necessary.
Figure 9. Size exclusion of native (A) and aggregated (B) samples of Phy (black), dgPhy (blue) and ogPhy (green). 500 μL of each sample, were injected to the column. To enable comparisons between the different glyco variants the RI, signals are normalized with respect to the peak area. The measurements were performed using a superdex 200 column (Amersham Pharmacia) 300×10 mm i.d. operated at a flow rate of 0.5 mL/min. The elution buffer was 50 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$, pH 7.0, 15 mM NaCl. The species that eluted from the column were passed through a differential refractometer (RID-10A, Shimadzu, Japan) for concentration determination. The measurements were done at the Faculty of Life Sciences, University of Copenhagen, Denmark.

Structural studies, using fourier transform infrared (FTIR) spectroscopy and circular dichroism (CD) spectroscopy, have shown that heat induced aggregation of different non glycosylated proteins is associated with intermolecular β-sheet formations [67-71]. Depending on their position, glycans might strongly interfere with such structural formations and thereby prevent a thermodynamic strong driving force for aggregate formation. Structural studies of aggregated glycosylated proteins including parallel investigations of deglycosylated variants will be required to clarify this proposal.

Both the literature and the present phytase study concurrently have shown that the kinetics of protein aggregation is effectively suppressed by glycans (as well as by other covalently attached polymers e.g. PEG). We suggest that this reduction in aggregation rates is caused by two factors: 1) a sterically induced disfavor of intermolecular interactions which may for example hinder intermolecular β-sheet formations and which increases the standard
free energy of the transition state ($\Delta \Delta G^\ddagger > 0$) and 2) a restriction in the number of monomers (n) in each aggregate (i.e. $\Delta \Delta G^\ddagger$ increases with n). This picture is illustrated in figure 10, showing the two step aggregation process (eqs. 3a and 3b) for a non-glycosylated protein (A) and a glycosylated protein (B). The rate constants in the transition step for folding and unfolding are referred to as $k_f$ and $k_u$, whereas $k_A$ is the rate constant of the aggregate growth step. Generally only small changes in equilibrium stability is found (discussed above), thus $k_u/k_f$ is suggested to be unaffected by glycosylation. Glycosylation reduces the growth rate, and $k_A$, which is indicated by the dotted arrow in illustration B. Just for simplicity low aggregation numbers are shown for both the glycosylated (n = 2), and the non-glycosylated (n = 5) protein in Fig. 10. However, our aggregation studies of dgPhy indicate the potential for significantly higher n values.

![Transition step and Growth step diagram](image)

**Figure 10.** A two step aggregation process for a non-glycosylated protein (A) and a glycosylated protein (B) is illustrated. The rate constants for denaturation, folding and aggregation are referred to as $k_u$, $k_f$ and $k_A$ respectively. The number of monomers in each aggregate, n, and $k_A$ are reduced upon glycosylation. This is indicated by a dotted arrow in B.

Comparisons of our aggregation studies of dgPhy and BSA indicate some significant differences between these proteins. The time dependent increase in NSI values (Fig. 5 and Fig. 8) indicates more pronounced particle growth in the dgPhy solution. The
experimental conditions are assessed to be comparable as molar protein concentrations and differences between solvent pH and protein pI are about the same in the BSA and dgPhy trials. Each experiment was carried out at the protein denaturation temperature which is \( \sim 61^\circ C \) for dgPhy and \( \sim 69^\circ C \) for BSA. Thus, these results indicate that the deglycosylated phytase has a stronger aggregation tendency than the naturally non glycosylated protein (BSA). In an evolutionary perspective this potentially signify an increased ability of glycosylated proteins to have surface exposed peptide sequences, which favor aggregation, and accordingly, perhaps a larger diversity between the amino acids forming the peptide surface. Similar interpretation of increased aggregation tendency of deglycosylated protein, relative to naturally non-glycosylated, has previously been put forward by Jafari-Aghdam et al. [128].

- **Effects of glycosylation on protein-surfactant interactions**

  Industrial application of proteins often requires some compatibility with surface active components. It follows that the usefulness of glycosylation to some degree depends on glycan-surfactant interrelationships. Interactions of this type, however, have only been sporadically investigated [136, 137]. We have used Phy and dgPhy to study the effects of protein glycosylation on structure- and stability perturbations by the anionic surfactant sodium dodecyl sulfate (SDS). It appears that glycosylation confers a limited effect on the resistance towards SDS-denaturation at 24°C. Thus, the SDS concentration required to denature half of the protein molecules was less than 1 mM higher for Phy than for dgPhy (respectively 6.9 mM and 6.2 mM SDS), Paper 3.

  DSC studies (not included in Paper 3) of phytase-SDS systems indicate a slight stabilizing effect of glycans against thermal denaturation. Figure 11 shows the denaturation temperature, \( T_d \) (left panel) and enthalpy, \( \Delta H_d \) (right panel) of Phy and dgPhy as a function of the total SDS concentration. \( T_d \) is generally higher for Phy (black symbols) than for dgPhy (blue). For Phy both the denaturation temperature and enthalpy are relatively constant in the \( \sim 0-1 \) mM SDS interval. For dgPhy \( T_d \) seems to decrease more constantly in the entire [SDS] interval. As the free SDS concentration ([SDS]_{total} − [SDS]_{bound}) in these studies is not known, an explanation of a protecting glycan effect could be favorable SDS-glycan interactions at high temperatures, which will decrease the activity of SDS in solutions of Phy relative to
solutions of dgPhy. However affinity studies at room temperature reveals weak SDS-glycan interactions, Paper 3. These studies showed that glycans adsorbed about half as much SDS (in g/g) as the peptide part at saturation.

The calculated denaturation enthalpies for Phy and dgPhy decreases strongly with SDS, and no significant differences between the two glyco-variants can be detected (Fig. 11, right panel). The data indicates a large decrease in $\Delta H_d$ around 1 mM SDS, rather than a gradual decrease in the entire [SDS] interval. Supported by SRCD (synchrotron radiation circular dichroism) investigations of these systems, this is suggested to be due to different temperature induced structural transitions at respectively high and low SDS (discussed below). At [SDS] > 5mM no denaturation can be identified by means of DSC. The same tendency has been reported for other SDS protein systems [41, 138]. Accordingly it appears, that at higher SDS concentrations $T_d$ is below the investigated temperature interval or, that $\Delta H_d$ is very small (and undetectable by DSC) in the presence of SDS > 5mM. Comparative analysis of SRCD data for phytase in 5.25 mM SDS and in pure water (Fig. 14) show that the temperature induced structural changes in the protein-detergent complex is gradual and less sensitive to temperature increase. For an equilibrium reaction (e.g. $N \rightleftharpoons D$) which has a small transition enthalpy the equilibrium constant only changes slightly with temperature, cf. the van’t Hoff equation. Thus, our studies suggest that small transition enthalpies explain the missing DSC signal in the surfactant containing systems.

Comparisons of temperature and SDS denaturation of phytase show that the denatured forms are different: Increasing the temperature seems to be associated with larger structural changes and an increased amount of unordered peptide, Paper 3. Along with that, heat induced changes in the secondary structures, investigated by SRCD, have been shown to depend strongly on the presence of SDS (data not included in Paper 3). CD spectra of Phy (Fig. 12) and dgPhy (Fig. 13) are recorded at varying temperatures (24-84°C) respectively in pure water (upper panels of Figs. 12 and 13) and in 5.25 mM SDS (lower panels). In pure water significant spectral changes of Phy and dgPhy are observed when the temperature is increased from 24°C (black curves) to 84°C (yellow curves). These spectra show an isodichroic point at ~ 205 nm, indicating a two state temperature induced transition of Phy and dgPhy. In contrast the same temperature increase in the presence of 5.25 mM SDS is correlated with minor spectral changes. Thus, these measurements clearly indicate that the
The secondary structure of SDS-protein complex at 5.25 mM SDS is not as temperature sensitive as the native (SDS free) proteins.

**Figure 11.** Denaturation temperatures (left panel) of Phy (black) and dgPhy (blue) and enthalpy change associated with the denaturation (right panel) are shown as a function of the total SDS concentration. The protein concentration in these trials was ~1.5 mg/ml for all Phy and dgPhy solutions, and the scan rate was 1.5°C/min. *The instrument details are described in figure 7.*

The protein concentration in the four SRCD temperature trial was 1.5 mg/ml for Phy and 1.13 mg/ml for dgPhy, which corresponds to about the same peptide concentration of the two glyco variants. Comparing spectra of Phy and dgPhy, however, reveals that the CD signals are higher for Phy than dgPhy, particular in SDS free solutions. As the spectra are superimposed when CD signals of dgPhy are multiplied by a scaling factor (1.3 in the pure water system), it is unlikely that the spectral variations are caused by differences in secondary structure. Therefore, it is suggested that decreased CD signals for dgPhy relies on increased surface adsorption of this protein during handling and a concomitant overestimation of the protein concentration in the cuvette.
Figure 12. Far UV (176-275 nm) SRCD spectra of Phy at nine different temperatures from 24°C (black) to 84°C (yellow). The gray spectrum at 24°C is recorded after heating to 84°C. The upper panel shows Phy spectra recorded in pure water, and the lower panel shows spectra recorded in 5.25 mM SDS. The protein concentration is 1.5 mg/ml in these trials, and a pure solvent baseline collected with the same cell was subtracted all spectra. SRCD spectra were recorded at the UV1 beam line on the storage ring ASTRID, Institute for Storage Ring Facilities (ISA), University of Aarhus, Denmark. Instrument details are described in Paper 3.
**Figure 13.** Far UV (176-275 nm) SRCD spectra of dgPhy at nine different temperatures from 24°C (black) to 84°C (yellow). The gray spectrum at 24°C is recorded after heating to 84°C. Upper and lower panels show respectively dgPhy in pure water and in 5.25 mM SDS. The protein concentration is 1.13 mg/ml in all trials, and a pure solvent baseline collected with the same cell was subtracted all spectra. *Instrument details are described in Paper 3.*
The limited temperature sensitivity of the SDS-protein complexes, relative to “native” Phy and dgPhy, is further illustrated by analysis of CD signals at selected bands as a function of temperature (Fig. 14 A and B). For spectra recorded in SDS free solutions the CD signal vs. temperature curves is s-shaped, Fig. 14A. The inflection point of each of these curves correspond to the denaturation temperature of Phy and dgPhy: $T_d(\text{Phy}) \sim 64^\circ C$ and $T_d(\text{dgPhy}) \sim 60^\circ C$, in reasonable accordance with the DSC data (Figs. 7 and 11). For spectra recorded in the presence of 5.25 mM SDS the CD signal vs. temperature curves (Fig. 14B) implies limited and gradual structural changes. It can therefore be concluded that there is no cooperative thermal transition of either protein in the presence of 5.25 mM SDS.

**Figure 14.** Panel A: The average CD signals measured at 186-192 nm for Phy and dgPhy in SDS free solutions is shown as function of the temperature. The inflection points correspond to $T_d(\text{Phy}) \sim 64^\circ C$ and $T_d(\text{dgPhy}) \sim 60^\circ C$. Panel B: The CD signals (average 188-194 nm) for Phy and dgPhy in 5.25 mM SDS is shown as function of the temperature.

The CD spectra recorded at 24°C reveal equal secondary structure of protein-SDS complexes before and after heating to 84°C (Fig. 12 and 13, lower panels) which is in contrast to the results obtained in pure water (Fig. 12 and 13, upper panels). This implies fully reversible conformational changes of Phy and dgPhy in 5.25 mM SDS under the experimental conditions and may well be explained by the prevention of aggregation by
anionic surfactant (electrostatic repulsion between protein-SDS complexes). Thus, in spite of the denaturing action of anionic surfactants, they may have a potential in a formulation perspective as they preclude the irreversible loss through aggregation and thus permit refolding into the active conformation upon reestablishment of favorable conditions.

Comparisons of CD spectra, at 24°C, for Phy and dgPhy after heating in pure water indicate, that glycosylation only to a very limited extend increases the recovery (refolding) after heating (Fig. 12 and 13, upper panels, gray curves). This shows, that under the experimental conditions aggregation preventing effect of glycans are very limited compared to the effects of SDS binding. However, the slight decrease in thermo stability (decreased T_d values) upon deglycosylation of phytase found in the present study (Figs. 11 and 14 A) and in DSC experiments published in Paper 1, may still be explained by the favoring of protein-protein interaction upon deglycosylation.

SRCD spectroscopy at 24°C and varying SDS concentration (∼0-20 mM SDS) as well as isothermal titration calorimetry (∼0-20 mM SDS, 24°C) implies that Phy and dgPhy react relatively equal to SDS, and that the stability towards surfactants (at room temperature) are as good as identical for the two glyco variants, Paper 3. In general our SDS studies propose that detergents compatibility of proteins is not particularly increased by glycosylation, which indicates that SDS-peptide interactions are not affected by glycosylation. Further, the limited effect of glycans, on detergent induced denaturation, suggests equivalent solvent exposure of glycans in respectively the N and the D state as discussed previously.

Summary of the biophysical effects of glycosylation

Under equilibrium conditions naturally glycosylated proteins are usually not more stable than their deglycosylated variants; for some systems no stability changes can be observed upon removal of even very large amount of carbohydrates. Our phytase studies imply a slight stabilizing effect of the glycan mantle when the enzyme is exposed to elevated temperature, and very limited effects on detergent compatibility. In contrast, proteins which are mutated prior to glycosylation are often less stable than their non-glycosylated parent molecules because of inexpedient position of glycosylation sites. Even successful glycosylation (no detectable change in the peptide structure) rarely increase the
thermodynamic stability of proteins. Better stabilizing results have been obtained for chemically glycosylated proteins, when the conjugates are made under mild temperature and pH conditions. The most straightforward interpretation of this is that introduction of glycosylation sequences in the peptide chain is a critical point in the attempt to increase the thermodynamic stability.

For naturally glycosylated proteins, as well as for proteins modified to hold glycans, pronounced effects on solubility and kinetic stability (e.g. prevention of aggregation) have been shown to be the clearest consequence of glycosylation. Based on our phytase investigations it is suggested that reduced aggregations rates of glycosylated proteins, at least in part, is due to a restriction in the number of monomers in each aggregate. The varying differences, and similarities, between glycosylated and non-glycosylated proteins might all be due to the same mechanism namely steric exclusion of glycans from interfaces. Steric hindrance of glycan-glycan and glycan-peptide interactions favors full solvation of carbohydrates in both the native and in denatured protein states. Therefore, glycans do not contribute significantly to changes in the free energy of structural transitions, and thus, do not affect the thermodynamic stability. On the other hand, the steric repulsions between the bulky glycans and surfaces suppress aggregation, surface adsorption as well as proteolytic degradation. This very unspecific mechanism is supported by very similar effects of surface attached carbohydrates and synthetic polymers.
Proteins are only marginal stable, and in aqueous solutions they usually have a tendency to both denature and undergo irreversible aggregation. Preservation of functional peptide structures is to a large extent improved when compatible solutes are added at high concentrations and/or when the peptide surfaces are glycosylated. This is primarily due to the denaturation preventing properties of preferentially excluded solutes, and aggregation preventing properties of glycans. The change in standard free energy upon structural transitions between native and denatured states is in general not significantly affected by glycosylation. Thus under equilibrium conditions glycosylated proteins and their deglycosylated variants are generally equally stable. The molecular mechanism, by which the attached carbohydrates (as well as synthetic polymers) improve the kinetic stability of proteins in solution, is suggested to be increased steric repulsions between monomers. This disturbs eventually intermolecular structure formations and the ability to form large molecular assemblies. Most importantly, from an industrial perspective, protein glycosylation hold the potential to successfully increase the shelf life of protein solutions. Therefore, attachment of carbohydrates and other polymers is a possible tool to overcome some requirements in protein product development.

This thesis only focuses on proteins in solution. However, the stability of protein powders might be increased by glycosylation. Comparative hydration studies of glycosylated and deglycosylated forms of phytase indicate the presence of hydrogen bonds between peptides and glycans in semi dry protein samples, Paper 2. Thus, glycans may hold the potential to increase the stability during drying in a similar way as suggested for some excipients (e.g. trehalose). Protein aggregation during lyophilization is a frequent problem, which is usually associated with the formation of intermolecular β-sheet formations [139, 140]. Due to the interruption of intermolecular peptide-peptide interactions glycans may prevent drying induced aggregation. This area, however, remains practically unexplored.
References


4. Ahern, T.J. and M.C. Manning, *Stability of Protein Pharmaceuticals*.


Paper 1
*Biochemistry, 42*:10295-10300
Preferential Binding of Two Compatible Solutes to the Glycan Moieties of
Peniophora lycii Phytase†
Heidi L. Bagger,‡ Claus C. Fuglsang,§ and Peter Westh*‡

Department of Life Sciences and Chemistry, Roskilde University, Building 18.1, P.O. Box 260, DK-4000 Roskilde, Denmark, and Novozymes A/S, Krogshøjvej 36, DK-2880 Bagsværd, Denmark

Received April 30, 2003; Revised Manuscript Received July 8, 2003

ABSTRACT: Regulation of hydration behavior, and the concomitant effects on solubility and other properties, has been suggested as a main function of protein glycosylation. In this work, we have studied the hydration of the heavily glycosylated Peniophora lycii phytase in solutions (0.15–1.1 m) of the two compatible solutes glycerol and sorbitol. Osmometric measurements showed that glycerol preferentially binds to phytase (i.e., glycerol–glycoprotein interactions are more favorable than water–glycoprotein interactions resulting in a preferential accumulation of glycerol near the protein interface), while sorbitol is preferentially excluded from the hydration sphere (water–glycoprotein interactions are the more favorable). To assess contributions from carbohydrate and peptide moieties, respectively, we compared phytase (Phy) and a modified, yet enzymatically active form (dgPhy) in which 90% of the glycans had been removed. This revealed that both polyols showed a pronounced and approximately equal degree of preferential binding to the carbohydrate moiety. This preferential binding of polyols to glycans is in contrast to the exclusion from peptide interfaces observed here (for dgPhy) and in numerous previous reports on nonglycosylated proteins. Despite the distinct differences between peptide and carbohydrate groups, glycosylation had no effect on the stabilizing action provided by glycerol and sorbitol. On the basis of this, it was concluded that the carbohydrate mantle of Phy is equally accessible in the native and thermally denatured states, respectively (most likely fully accessible in both), and thus that its interactions with compatible solutes have little or no effect on conformational equilibria of the glycoprotein. For solubility and aggregation equilibria, on the other hand, the results suggest a polyol-induced stabilization of monomeric forms.

The interaction of polyhydroxy alcohols (polyols) and proteins has attracted considerable research interest over decades. For example, polyols belong to the evolutionary conserved group of so-called compatible solutes, which are accumulated in a variety of cell types in response to high external osmotic pressures (1–3). While the primary role of this buildup is the colligative decrease in intracellular water activity and the concomitant regulation of cellular volume, insight into biomolecule–polyol interactions is necessary to elucidate in vivo effects of these compounds. For example, this type of information may elucidate the remarkable noneffect of polyols, even in molar concentrations, on metabolic processes (2). Furthermore, recent work has suggested that at least some compatible solutes are not merely inert effectors of cellular volume but provide a necessary protection of macromolecular structure in the low-water environment that prevails even after accumulation of the organic solutes (4–6). This aspect, which has led to terms such as “compensatory solutes” for the compounds (5, 7, 8) and “osmophobic effects” for their stabilizing mechanism (9), further stresses the relevance of molecular information on biomolecule–polyol interactions.

A main approach of elucidating polyol–protein interactions has its roots in the theory of preferential interactions (see, e.g., refs 10 and 11). Experimental work along these lines has shown that polyols are in general moderately to strongly excluded from the protein–solvent interface (12–14). This implies that the local concentration of the polyol in the protein hydration sphere is lowered compared to the bulk. By use of the so-called Wyman relationships (15), changes in the extent of exclusion between different states can be directly related to the effect of the polyol on protein equilibria, such as thermal unfolding or aggregation. This connection has motivated the use of arguments based on preferential interaction theory in discussions of mechanisms underlying numerous processes within biochemistry, medicine, and biotechnology. Some examples include adaptive responses to hyperosmotic environments mentioned above (4, 5, 16, 17), stresses associated with freezing (18), gelation processes (19) and the well-known in vitro stabilization provided by polyols (14, 20, 21).

In this work we have used dew point osmometry (12, 22) to measure the preferential interactions of a heavily glycosylated phytase and two polyols, glycerol and sorbitol. The work was motivated by several factors. First, glycoproteins have not previously been directly investigated in this respect, and insight into the properties of glycosylated interfaces is clearly required in discussions of polyol effects in vivo. Second, polyol–glycoprotein interrelationships may be of...
particular interest since glycosylation covalently modifies physical properties such as solubility, thermal stability, and aggregation behavior (23–32). These same properties can also be regulated by polyols through noncovalent (solvent) effects (33), and the interplay of these two mechanisms appears to be of interest. Third, the fact that phytase (like several other glycosylated enzymes) can be readily deglycosylated without significant changes in the enzymatic activity (and hence presumably in polypeptide structure) may provide an approach to single out separate effects of peptide and carbohydrate and to assess driving forces underlying the measured preferential interactions. Recent investigations of the latter problem (34–37) have focused on the balance between soft and hard interactions (i.e., respectively “soft” intermolecular binding such as van der Waals forces or hydrogen bonding and “hard” steric repulsions). Most of these works have highlighted the importance of the hard contributions to polyol–protein interactions. Due to the pronounced differences in chemical structure and steric properties of the interface of phytase (Phy)1 and deglycosylated phytase (dgPhy), respectively, a comparative analysis of the Phy/dgPhy system may provide an experimental approach to this discussion.

MATERIALS AND METHODS

Glycerol (99%) and D-sorbitol (>98%) were purchased from Sigma-Aldrich (St. Louis, MO) and used as supplied. *Peniophora lycii* phytase was expressed in *Aspergillus oryzae* and purified according to previously published procedures (38). SDS–PAGE showed >95% purity and a molecular mass of about 65 kDa. The purified protein solution was stored at −25 °C. When needed, the samples were thawed and extensively dialyzed (MW cutoff 8 kDa) against milliQ water.

The systematic name of phytase (EC 3.1.3.26) is myo-inositol hexakisphosphate phosphohydrolase, and it catalyzes the hydrolysis of phosphomonoester bonds of phosphorylated myo-inositol, phytate (myo-inositol 1,2,3,4,5,6-hexakisphosphate). The products are lower forms of myo-inositol phosphates and inorganic phosphates (38). The enzyme is N-glycosylated at 10 sites with heterogeneous glycan chains dominated by structures of two N-acetylgalactosamine units connecting the protein and a high mannose moiety (38). To produce deglycosylated phytase (dgPhy), 6.3 mL of the purified sample (23.7 mg/mL) was mixed with a 1.6 mL solution (0.093 mg/mL) of the enzyme Endo F1 (a kind gift from Hoffmann-La Roche, Basel, Switzerland) and 32 mL of 0.15 M sodium acetate and 0.01% (v/v) Tween-20, pH 5.5. The enzyme was further diluted 26-fold into a preincubated (37 °C) substrate solution (5 mM sodium phytate in 0.1 M sodium acetate and 0.01 vol % Tween-20, pH 5.5). After 30 min at 37 °C, the reaction was stopped by adding an equal volume of 10% trichloroacetic acid. Free inorganic phosphate was measured by the addition of an equal volume of molybdate reagent (7.3 g of FeSO4, 1.0 g of (NH4)6Mo7O24·4H2O, and 3.2 mL of H2SO4 diluted in milliQ water to 100 mL). Absorbance was measured at 750 nm (Vmax microplate reader; Molecular Devices, Sunnyvale, CA), and the phosphate concentration was determined relative to prepared standard solutions. The assay was repeated in substrate solutions containing 2.2 M glycerol or sorbitol.

The water activity (osmolality) of two-component (water + polyol) and three-component (water + polyol + Phy or water + polyol + dgPhy) solutions was measured using a Vaprop 5520 osmometer (Wescor Inc., Logan, UT). This instrument measures the thermodynamic activity of water in a solution by recording the difference between the temperature of a liquid and the dew point of its equilibrated gas phase. In this work, the osmolality of two- and three-component systems was measured as a function of the polyol concentration in the 0.15–1.1 m range. All experimental trials were initiated by calibrating the osmometer using standard solutions (100, 290, and 1000 mOsm, respectively) provided by Wescor. If standard measurements did not provide reproducible readings, the osmometer thermocouple was cleaned with an ammonium hydroxide solution (Wescor “cleaning solution”) and deionized water. The calibration was frequently tested (at least once for every 10 readings) against the standard solutions. If the reading deviated more than 2 mOsm from the nominal value, the osmometer was recalibrated. All data reported here were recorded at 25 ± 1 °C and a laboratory relative humidity of 20–30%. Courtenay et al. (12) have found that measurements on this instrument become less reproducible at an ambient humidity exceeding 42%. The quality of the data is crucially sensitive to the exact control of the sample composition. Hence, the following gravimetric procedure for sample preparation was adopted. First, two-component batch solutions of glycerol and sorbitol, respectively, were prepared gravimetrically from milliQ water and newly delivered polyol. Subsequently, 1000 µL of the dialyzed samples of Phy and dgPhy, respectively, was transferred to tared plastic tubes, weighed, and lyophilized in the tubes for 48 h. Immediately after removal from the lyophilizer, the dry protein was dissolved in ∼200 µL of the two-component batch solution of polyol added directly to the plastic tube. The sample was sealed, weighed, and stored at 5 °C. For the osmometric measurements, 7 µL aliquots of the two-component batch solution and protein sample (three-component solution), respectively, were ana-

1 Abbreviations: Phy, *Peniophora lycii* phytase (myo-inositol hexakisphosphate phosphohydrolase); dgPhy, deglycosylated *P. lycii* phytase; Endo F1, endo-β-N-acetylgalactosaminidase; DSC, differential scanning calorimetry.
Families of curves, the two ordinates are displaced by 100 mOsm. (triangles). To facilitate reading of the two nearly superimposed capacity trace was recorded during heating from 30 to 90 °C. The thermal stability was measured at about 4% (w/w) (weight loss after lyophilization in samples dried to constant mass over P2 O5 at 80 °C). The slopes of the curves are illustrated in Figure 1. On the basis of the reproducibility of polyol solutions and Phy and dgPhy preparations were conducted.

The thermal stability of Phy and dgPhy was investigated by scanning calorimetry (MC 2-DSC; MicroCal, Northampton, MA). Dialyzed and lyophilized protein was added to the two-component batch solution to yield a total protein concentration of about 5 mg/mL. The calorimeter was loaded with the protein–polyol solution (cell) and a binary polyol solution of the same concentration (reference), and the heat capacity trace was recorded during heating from 30 to 90 °C at 45 °C/h. The thermal stability was measured at about 10 different polyol concentrations in the 0–1500 mOsm range.

RESULTS AND DATA TREATMENT

Representative raw data from the dew point osmometry are illustrated in Figure 1. On the basis of the reproducibility of the calibration measurements and the observation that lyophilized protein samples have a residual moist content up to 4% (w/w) (weight loss after lyophilization in samples dried to constant mass over P2 O5 at 80 °C), the experimental precision in Figure 1 is estimated to be ±5 mOsm. The slopes of the curves are 22 ± 4.1 m−1 (Phy) and −43 ± 3.6 m−1 (dgPhy) for sorbitol and 31 ± 2.6 m−1 (Phy) and −5 ± 2.4 m−1 (dgPhy) for glycerol. These results reflect protein–solvent interactions, but it has been argued (41) that a different preferential binding parameter, Γμj, provides a more meaningful approach to discussions of thermodynamic effects and their molecular interpretation. This latter binding parameter is defined analogously to Γμj (eq 1), the only difference being that the chemical potential of polyol (μj) rather than that of water (μ1) is kept constant in the partial derivative. Γμj is not readily accessible by experiment but can be estimated from Γμj, as described by Courtenay et al. (12). Using eq 12 of this work, we have calculated Γμj for the present systems and plotted it against the polyol concentration, m3, in Figure 2. The numeric difference between Γμj and Γμj is very small. Thus, the data for sorbitol (filled symbols) shows negative slopes, indicating that both Phy and dgPhy are preferentially hydrated in aqueous sorbitol. The slopes of the curves are −24 ± 6 and −45 ± 6 m−1, respectively, suggesting that the preferential hydration (or preferential exclusion of sorbitol) is stronger by a factor of 2 for the deglycosylated protein. For glycerol (Figure 2, open symbols) the Phy data show a distinctive positive slope (25 ± 5 m−1), while Γμj decreases slowly with m3 for dgPhy (Γμj/m3 = −6 ± 5 m−1). Hence, glycerol is preferentially bound to the deglycosylated protein while the peptide moiety shows a slight preference for water. Comparison of all data in Figure 2 reveals the general trend that glycans interact more favorably with the polyols than with water. Interestingly, the (negative) change in slope observed as a result of the deglycosylation is approximately the same for the two investigated solutes. This signifies favorable interactions of similar strength between the two polyols and the glycan moieties.

\[ \Gamma_{\mu_1} \approx \left( m_3 - m_3^A \right)/m_2 \]  

(2)

where \( m_3^A \) is the polyol molality of the two-component solution, which has the same water activity (osmolality) as the three-component solution. Hence, the numerator of eq 2 is the horizontal distance between a set of curves (with/without protein) in Figure 1. This difference was determined using each data point for the two-component systems and the isoosmotic composition of the three-component solution calculated from the regression line shown in Figure 1.

Plots of Γμj vs polyol concentration (not shown) revealed linear correlation. The slopes of the regression lines (±95% confidence limit) were respectively −22 ± 4.1 m−1 (Phy) and −43 ± 3.6 m−1 (dgPhy) for sorbitol and 31 ± 2.6 m−1 (Phy) and −5 ± 2.4 m−1 (dgPhy) for glycerol. These results show that the isoosmotic binding parameter specifies the number of polyol molecules which have to be added or removed to reestablish the chemical potential of water upon addition of one protein molecule to an aqueous polyol solution. Its rigorous relation to solution thermodynamics, other preferential binding parameters, and conventional binding theory as well as its molecular interpretation have been discussed in detail in a number of works (10, 11, 39–41) and will not be reviewed here. We only note that the sign of Γμj can be interpreted as indicating whether the small solute is accumulated in (preferential accumulation or “binding”, Γμj > 0) or depleted from (preferential exclusion, Γμj < 0) the zone near the protein interface.

For nonvolatile solutes, Γμj can be derived directly from dew point osmometry measurements (12)
The degree of stabilization brought about by sorbitol and glycerol, respectively, is practically identical for Phy and dgPhy (the lines in Figure 3 are pairwise parallel). Selected samples were taken through a second heating scan, and this suggested that the thermal denaturation of Phy and dgPhy was almost fully reversible.

The enzyme activity assays showed that deglycosylation brought about a small but systematic reduction in turnover rate. Thus, the activity of the deglycosylated samples fell within the 90–95% range of glycosylated protein from the same sample. The presence of 2.2 M glycerol or sorbitol in the reaction buffer did not change this behavior.

### DISCUSSION

Carbohydrates are posttranslationally attached to most secreted eucaryotic proteins and appear to serve a number of purposes (42), which may be coarsely categorized in two groups. First, glycosylation has been connected to molecular recognition with respect to, e.g., targeting, immunological properties, protease susceptibility, and interactions with chaperone proteins (42–45). Second, the glycans modify physical (and hence functional) properties of enzymes. In particular, glycosylation has been related to solubility, but other properties such as oligomeric aggregation behavior and conformational stability have also been discussed intensively (23–32).

The understanding of this latter group of effects relies on the elucidation of protein–solvent interactions, and the scope of the current work is to address this through measurements of preferential interactions in Phy—polyol and dgPhy—polyol systems, respectively.

The main result of the preferential interaction data in Figure 2 is that glycerol and sorbitol preferentially bind to the glycan groups of Phy. This effect is rather pronounced. Thus, the additional glycosylation of Phy (approximately 20 kDa oligosaccharides compared to approximately 2 kDa monosaccharides on dgPhy) increases the value of $\Gamma_{\mu}/m_3$ (i.e., the slope in Figure 2) by $20–30 \text{ m}^{-1}$ for both polyols.

To our knowledge this is the first direct quantification of such interactions. The favorable interactions of glycans and polyols are in contrast to their preferential exclusion from polypeptide interfaces found here (data for dgPhy in Figure 2) and in several previous reports on nonglycosylated proteins (12–14). Comparison with these works shows that the preferential exclusion of glycerol from dgPhy found here is similar to what has been reported for (nonglycosylated) proteins. Thus, the slope in Figure 2 ($-6 \pm 5 \text{ m}^{-1}$) for glycerol/dgPhy (MW 48 kDa) compares well with results for bovine serum albumin (66 kDa), chymotrypsinogen (26 kDa), $\alpha$-chymotrypsin (25 kDa), and RNase A (14 kDa), which fall in the $-10$ to $-2 \text{ m}^{-1}$ range according to their size (12).

For sorbitol the preferential exclusion ($-45 \pm 6 \text{ m}^{-1}$) found here for dgPhy is stronger than what has previously been measured for RNase A (approximately $-8 \text{ m}^{-1}$) (14). Even if the results are normalized with respect to the sizes of the proteins (i.e., surface areas assuming that this parameter is proportional to MW$^{2/3}$), the preferential exclusion of sorbitol from dgPhy is more than twice the value in RNase A solutions. The degree of preferential exclusion in dgPhy—sorbitol is comparable to that found for the BSA—betaine system, where it roughly corresponds to the complete exclusion of solute from a monolayer of water at the protein interface (12).
Preferential interactions can be related to the thermodynamic stability of protein conformations through the Wyman relationships (15). For a simple two-state (native ↔ denatured) equilibrium, for example, a solute will stabilize the conformation with the highest \( \Gamma_{\mu} \) (i.e., least negative \( \Gamma_{\mu} \) for preferentially excluded solutes). Since, in many cases, the preferential interactions are rather independent of the solution conditions and surface properties of the protein (11, 33), preferential exclusion of a solute from the native state is often tantamount to its stabilization. In that case, the degree of exclusion in the denatured state is larger, and hence the stability of the native state is increased, simply as a result of the enlargement of the interface upon denaturation. This type of stabilization is believed to occur for a range of compatible (or compensatory) solutes, and it has been suggested that “a major characteristic of the compensatory solute is that it is preferentially excluded from the protein surface and its immediate hydration sphere” (4). While the current data (Figures 2 and 3) suggest that this picture applies to dGPhy, it may not be readily applicable to glycoproteins, since the polyols showed a considerable preferential binding to the carbohydrate mantle of Phy. The extent of this effect may be illustrated by comparing the influence of the glycans on the slope \( \Gamma_{\mu}/m_3 \) (20–30 \( m^{-1} \) according to Figure 2) and the preferential binding to intact proteins of denaturants such as urea and guanidine hydrochloride. For urea the preferential binding to small globular proteins typically corresponds to \( \Gamma_{\mu}/m_3 \) values in the 5–15 \( m^{-1} \) range (11, 22, 46) although negative values have been observed (11). The affinity of serum albumin for guanidine hydrochloride is slightly higher, with a slope \( (\Gamma_{\mu}/m_3) \) of about 18 \( m^{-1} \) in dilute solution (47, 48). This comparison might seem puzzling with respect to the cellular compatibility of these solutes, but as illustrated in Figure 3, the stabilization conferred by the polyols is in fact independent of the extent of glycosylation. These observations can be reconciled if the oligosaccharides are equally exposed to the solvent in the native and thermally denatured states. Such lack of additional exposure upon unfolding most likely implies that the glycans are fully exposed in the more condensed (native) state. If so, their preferential interactions will not change as a result of structural changes such as thermal unfolding, and hence polyol–glycan interactions will not perturb conformational equilibria of the protein molecule. It is interesting to note that if this mechanism turns out to be general for glycosylated proteins, it may set a maximum limit for the tolerated load of carbohydrates on a protein. At very high levels of glycosylation, some degree of (temporary) steric restrictions will limit the solvent accessibility of the carbohydrates in the native state. When these restrictions are reduced in a less compact, unfolded state, the increased accessibility would promote preferential binding of the polyol and consequently stabilize this state over the native glycoprotein. It follows that, rather than being compensatory solutes, polyols would be denaturants of such heavily glycosylated proteins. On the basis of these considerations, we suggest that the oligosaccharides of Phy are fully hydrated in both the native and thermally denatured state. As a result, the polyol-induced stabilization of Phy is suggested to arise from unfavorable polyol–polypeptide interactions (preferential hydration of the peptide moiety), while glycanceroy interactions are of minor importance for the conformational stability of the protein.

As mentioned in the introduction, many reports have discussed glycosylation with respect to solubility (and changed oligomeric aggregation behavior). The current results suggest that, in an environment containing polyols, glycosylation will strongly promote solubility (and the stability of nonaggregated forms) and thus counteract the “salting-out” effect of polyols on peptides. This follows from the favorable glycanceroy interactions which will favor forms with larger accessible carbohydrate surface. Analogous arguments have been put forward by Tams et al. (49), who found that the solubility of a glycosylated, fungal peroxidase was enhanced by ammonium sulfate and reduced by acetone. They further showed that this effect on the solubility scaled with the carbohydrate content in a series of N-glycosylation mutants. Tams et al. suggested that the results for ammonium sulfate reflected favorable interactions, with the glycans, and this conclusion has an interesting analogy to the present observations. Thus, both polyols and sulfate generally show preferential exclusion from peptide interfaces (11, 33) but appear to preferentially bind to the carbohydrate moiety of glycoproteins. This distinct difference raises the question of what sources drive the interactions. For nonglycosylated proteins, the interaction with sulfate is rationalized through the strong structure-making (cosmotropic) properties of the ion. This (Hofmeister) effect of sulfate relates to its strong hydration and the concomitant affinity for bulk rather than interfacial positions (50). For polyols, analysis of the interaction through scaled particle theory suggests that steric restriction is a major contributor to the preferential exclusion (34, 36). However, both of the above mechanisms are independent of the chemical structure of the interface, and hence they cannot account for the observed difference between peptide and carbohydrate surfaces. Therefore, we suggest that preferential binding of polyols to glycans as well as the “non-Hofmeister” behavior of sulfate relies on direct (i.e., soft) interactions of carbohydrates and the solutes. For the polyols, this probably involves hydrogen bonding to the glycans, although this explanation is far from unequivocal. Thus, thermodynamic properties of most aqueous polyols and carbohydrates (51–53) show that interactions with water are favored over mutual solute interactions. It follows that there is no a priori reason why glycans should attract polyols stronger than they do water, and further insight into this mechanism will probably require systematic studies of model systems.

The results for polyol–glycan interactions presented here may also be discussed along the lines of the seminal work on glycerol–protein interactions by Gekko and Timasheff (13). These workers discussed surface-dependent interactions and suggested that glycerol was attracted to the polar regions of (nonglycosylated) proteins while nonpolar areas interacted more favorably with water. The current finding of a sizable preferential binding of glycerol to the polar surface of the oligosaccharides corroborates the suggestion of Gekko and Timasheff. More recent work in this field by Bolen and co-workers found that the interaction of amino acid side chains and compatible solutes was typically weakly attractive, while the compatible solutes were strongly expelled from the interface of the peptide backbone (9, 54, 55). The preferential binding of polyols to glycans found here may point toward
an additional interrelationship of surface chemistry and preferential interaction. Thus, the subtle balance of water and polyol interactions with the interface of glycoproteins seems to result in a relative affinity (or preference) for water, which scales as peptide backbone > side chains > glycans. Elucidation of the validity and importance of this suggestion will rely on further systematic studies of preferential interactions of glycosylated proteins.

To summarize, we have found that the compatible solutes glycerol and sorbitol preferentially bind to the carbohydrate moiety of the glycoprotein phytase, while they are preferentially excluded from the peptide surface of this protein. The thermal stability of Phy is slightly reduced upon deglycosylation, but the stabilizing action of the compatible solutes is identical for Phy and dgPhy. It is suggested that the polyols accumulate in the hydration sphere of the solutes is identical for Phy and dgPhy. It is suggested that compatibility of polyols and glycoproteins. Further studies of glycosylated proteins including parallel investigations of native and denatured forms will be required to clarify these proposals.

REFERENCES

Paper 2


Hydration of a glycoprotein: relative water affinity of peptide and glycan moieties

Received: 26 August 2005 / Revised: 7 November 2005 / Accepted: 15 November 2005 / Published online: 8 December 2005
© EBSA 2005

Abstract Glycosylation, the most prevalent post-translational modification of proteins, affects a number of physical properties including the interactions with the surrounding aqueous solvent. Such glycan–water interactions have been discussed with respect to the increased solubility generally observed for glycoproteins, but experimental support of this correlation remains sparse. We have applied a two-channel calorimetric method to measure the free energy and enthalpy of hydration at 25°C for the glycoprotein phytase (Phy) and a deglycosylated form (dgPhy) of the same protein. Comparisons of results for Phy and dgPhy show that the polypeptide moiety has a higher affinity for water than the glycans. In fact, at moderate hydration levels (~0.3 g water/g macromolecule) the water uptake appears to be entirely governed by adsorption to the peptide groups. We conclude that strengthened interaction with the solvent is unlikely to be the mechanism underlying the increased solubility and lowered propensity of aggregation often reported to result from the glycosylation of proteins.

Introduction

Glycosylation is the most common post-translational modification of proteins, and one of its roles appears to be the suppression of aggregation and increase of solubility (Narhi et al. 2001; Price et al. 2003; Schulke and Schmid 1988a, b; Song et al. 2001; Tams et al. 1999; Tams and Welinder 1995). The mechanisms underlying these effects of glycans remain poorly understood, but it has been suggested that they may rely on the “hydrophilic nature” of the sugar residues (Creighton 1996). This implies that the sugar moieties interact more favorably with water than the peptide chain does and that these favorable interactions stabilize the dissolved monomeric state of the proteins over crystallized or aggregated states. In this work we address the question of the relative hydrophilicity of glycans and peptides through a novel calorimetric method, which simultaneously measures the free energy (the binding isotherm) and the enthalpy of hydration. Our model system is the hydrolytic enzyme phytase (Phy) E.C.3.1.3.26. from Peniophora lycii. This N-glycosylated enzyme consists of 47.5 kDa polypeptide and around 17.5 kDa glycans, mainly mannose residues, distributed on ten N-glycosylation sites (Lassen et al. 2001). Phy can be enzymatically deglycosylated with no change in the catalytic properties (and hence in the peptide structure). It follows that comparative investigations of Phy and its deglycosylated form (dgPhy) has a particular potential to elucidate the effects of glycans on hydration and other physico-chemical properties.

The conspicuous solubility in water of some carbohydrates may be taken as an indication of strong interactions between water and glycans. However, we have found some evidence based on thermodynamic and light scattering measurements that water may in fact interact more favorably with the peptide moieties than with the glycans (Bagger et al. 2003) (Nielsen et al., in preparation). If indeed so, the solubility promoting effect of glycans cannot rely on favorable solvent interactions but must depend on other factors e.g., steric inhibition of protein–protein contacts and the concomitant poor packing of aggregates or crystals. To elucidate this, we have investigated the effects of glycans on the water uptake by dried samples. This process may be considered the most fundamental measure of protein hydration (Rupley and Careri 1991) and thus essential to the understanding of glycan–water interactions.
Materials and methods

*Peniophora lycii* Phy was expressed in *Aspergillus oryzae* and purified according to the previously published procedures (Lassen et al. 2001) and stored at −25°C in buffer solution (20 mM Tris-acetate, pH 6). The molecular mass of the glycosylated Phy (~65 kDa) and the purity of this enzyme (> 95%) were determined by SDS-PAGE. The enzymatic deglycosylation of Phy producing dgPhy was performed by Endo-β-N-acetylglucosaminidase F1 (Endo F1) E.C.3.2.1.96 purchased from Calbiochem (Darmstadt, Germany). The deglycosylation procedure has previously been described in detail (Bagger et al. 2003). To ensure that the deglycosylation was quantitative, the reaction mixture was investigated by SDS-PAGE. Phy and dgPhy was extensively dialyzed (MWCO 12–14 kDa) against milliQ water at 4°C to avoid any salts (other than counter-ions) in the hydration experiments.

To produce dry enzyme samples Phy and dgPhy were lyophilized at 0.05 torr for 2 days. After drying the pressure was equalized with dry nitrogen gas and the cell containing the enzyme was immediately sealed with a close-fitting lid. To assure complete drying test samples of lyophilized enzyme was placed at 80°C to gravimetrically confirm that no further water evaporated.

The calorimetric measurements were conducted on a multi-channel thermal activity monitor (TAM, 2277) isothermal calorimeter (Thermometric A/B, Järfälla, Sweden). The experimental principles have been described in detail elsewhere (albeit for a different application) (Hansen et al. 2004). Briefly, two calorimetric channels are serially perfused with N2(g) at a controlled flow rate and relative humidity. The exothermic heat flow associated with water adsorption to the surface of the calorimetric cell during the RH ramp.

To minimize the exposure of dry enzymes to (moist) air prior to the hydration experiment Phy and dgPhy were lyophilized in the calorimetric cell (cell A), which was directly connected to the perfusion unit. The enzyme was kept in the calorimeter under a flow of dry nitrogen gas until the hydration experiment was initiated. The total duration of an experiment was ~45 h. The first 2–6 h the relative humidity was set to 0% and kept there until the heat flow from channel A and B were stable for about 1 h. Subsequently, the RH was increased linearly from 0 to 98% over 20 h and kept at 98% for an additional ~20 h. This procedure was adopted from an earlier work (Liltorp 2003) which showed that two-channel calorimetry provided water binding isotherms and adsorption enthalpies for hen lysozyme similar to those reported by the other experimental approaches (Rupley and Careri 1991; Hnojewyj and Reyerson 1961; Smith et al. 2002).

Several reference experiments with an empty cell in channel A were conducted to measure the amount of water which adsorbs to the surface of the calorimetric cell during the RH ramp.

Results and discussion

Figure 1 exemplifies results from the calorimetric trials. Panel A shows the heat flow, HF_A, arising from the adsorption of water to 9.74 mg Phy when the RH is changed as illustrated by the solid line in the same graph (right hand ordinate). Panel A also shows the heat flow, for the protein sample (HF_p) and for the empty calorimetric cell (refHF_p). This panel also shows the ambient humidity in the calorimetric cell, RH_cell, calculated from Eq. 4. The amount of water adsorbed by the protein is illustrated by the data from channel B (b). The water removed from the purge gas through adsorption to the protein is quantified by the difference between HF_p and refHF_p. The water uptake, WH2O, specified by Eq. 1 in nanomole per second is shown in the inset.

![Fig. 1](image-url)
refHF_A generated by the adsorption of water to an empty calorimetric cell in a reference experiment with the same time-course of the RH. The amount of water adsorbed is illustrated in panel B by the (endothermic) heat flows associated with the saturation of the gas stream in the Phy trial (HF_B) and reference experiment (refHF_B). The amount of water taken up by the protein, W_{H2O}, (in moles/s) is specified by the difference between the two curves in Fig. 1b and may be written

\[ W_{H2O} = -\frac{H_{F_B} - refHF_B}{\Delta H_{evap}} \]

where \( \Delta H_{evap} \) is the heat of evaporation of water at 25°C. The total amount of adsorbed water, \( h \) (in g H\(_2\)O/g enzyme), can be expressed as a function of the experimental time, \( t \), according to Eq. 2.

\[ h = \frac{M_{H_2O}}{m_P} \int_0^t W_{H_2O} \, dt \]

where \( M_{H_2O} \) and \( m_P \) are, respectively, the molar mass of water and the mass of protein in cell A. Once the water uptake has been quantified the molar enthalpy of water adsorption, \( \Delta H_{ads} \) can be calculated as

\[ \Delta H_{ads} = \frac{H_{F_B} - refHF_A}{W_{H_2O}} + \Delta H_{evap}. \]

The raw data in Fig. 1a relates to the transition water(g) \( \rightarrow \) water(ads). However, since \( \Delta H_{evap} \) is added on the right-hand side of Eq. 3, \( \Delta H_{ads} \) specifies the enthalpy of transferring one mole of water from the pure liquid to the surface adsorbed state.

The actual RH in the calorimetric cell in channel A, RH_{cell}, is slightly lower than the humidity of the incoming gas, RH\(_{supply}\), due to the continuous uptake of water in the cell. However, RH_{cell}, which is the water activity governing the adsorption process, can be calculated from the calorimetric signal in cell B

\[ RH_{cell}(\%) = \frac{H_{F_B} - refHF_B(0)}{refHF_B(98) - refHF_B(0)} \times 98\% \]

where refHF_B(0) and refHF_B(98) are the heat flows from cell B in the reference experiment when RH\(_{supply}\) is 0 and 98%, respectively.

The binding isotherm, i.e., the total amount of water uptake, \( h \), plotted as a function of the water activity expressed by the relative humidity in cell A, RH_{cell}, is illustrated for Phy and dgPhy in Fig. 2. Each curve is calculated as the average of three runs and the error bars indicate the standard deviation. The figure shows the degree of hydration, \( h \), both in units of gram water/gram enzyme (panel A) and mol water/molar enzyme (panel B).

The results in Fig. 2 clearly show that the glycans of Phy do not bring about an increased uptake of water. Hence, no significant difference in the mass specific uptake (panel A), between Phy and dgPhy could be detected for RH_{cell} < 50%, and at higher humidity, the deglycosylated form adsorbs more than the glycoprotein. For RH_{cell} = 90%, for example, the polypeptide has an equilibrium water content of 0.34 g H\(_2\)O/g macromolecule while the glycosylated protein holds only 0.23 g H\(_2\)O/g macromolecule. If, for simplicity, the adsorption of water to the polypeptide and the glycan groups are considered independent, these numbers suggest that the glycans contributes negligibly to the total water adsorption at 90% RH. [Phy is \( \sim73\% \) (w/w) polypeptide corresponding to a contribution from the peptide of 0.73x0.34 = 0.25 g H\(_2\)O/g macromolecule to the water uptake of Phy. This value accounts for the full water adsorption of the glycoprotein and thus suggests an unimportant contribution of the glycans]. This is further illustrated in Fig. 2b, which shows the water uptake in mol/mol. It appears that at high humidities
(RH > 80%) the difference between the two curves is smaller than the experimental uncertainty. Hence, the number of water molecules taken up by one enzyme molecule is not increased by the glycans in the high RH range. In the lower RH range where the amount of water is less than one molecule per amino acid residue, the molar uptake of the glycosylated protein is slightly larger.

The adsorption isotherms in Fig. 2 quantify the free energy of the absorbed water molecules (with respect to an appropriate standard state). To illustrate this more clearly and to facilitate comparisons with Fig. 3 (below), the data were re-plotted (in the inset) to show the chemical potential of water as a function of the water content. Assuming that the adsorption process is practically equilibrated\(^1\) during the slow increase in RH, the chemical potential of adsorbed water may be written \(\mu_w = \mu_w^* + RT \ln\tfrac{RH}{100\%}\). In other words the lowering in the chemical potential of adsorbed water with respect to saturated vapor, \(\Delta \mu_w\) is \(\Delta \mu_w = RT \ln\tfrac{RH}{100\%}\). This function is shown in the inset of Fig. 2a, and will be discussed below.

The weak hydration of glycans suggested by the results in Fig. 2 strongly speak against a hydrophilic mechanism for the enhanced solubility and reduced propensity for aggregation generally observed as a result of glycosylation (see Introduction).

The hydration of glycans is further elucidated by the adsorption enthalpy, \(\Delta H_{\text{ads}}\), which is plotted as a function of \(h\) for Phy and dgPhy in Fig. 3.

It appears from this figure that water binds rather strongly to the enzymes at the lowest hydration levels. Thus, the enthalpy of adsorbed water is initially 20–30 kJ/mol lower than for the pure liquid. The pronounced positive slope of \(\Delta H_{\text{ads}}\) in this range, however, suggests that both forms of the enzyme only have few sites with this favorable binding enthalpy. A similar behavior of \(\Delta H_{\text{ads}}\) at low hydration levels, corresponding to one-water molecule for every 2–3 amino acid, has previously been reported for other proteins (Smith et al. 2002). While this behavior may be important for the general understanding of protein hydration we suggest that these very few strongly bound water molecules are of minor relevance in attempts to single out the effects of glycans on proteins in solution. At higher water contents, \(\Delta H_{\text{ads}}\) gradually increases from \(-10\) to \(-4\) kJ/mol. Comparisons of these results and the free energy data (inset of Fig. 2a) allows us to single out the enthalpic and entropic contributions to the net driving force of the hydration. At very low hydration we find, \(\Delta \mu_w \approx -10\) kJ/mol and \(\Delta H_{\text{ads}} \approx -25\) kJ/mol. In other words the entropic contribution \(\Delta S_w\) is about \(-15\) kJ/mol. This substantial entropic penalty appears intuitive for the transfer of water from the pure liquid to the adsorbed state. At intermediate hydration levels, the entropic contribution becomes rather small, and is only about \(-4\) kJ/mol at 0.2 g/g. A similar behavior has been observed for other proteins (Smith et al. 2002) and may rely on the “lubricating” effect of water on biomolecular movement (Barron et al. 1997). For the current discussion it is particularly important to notice that no clear differences between the two variants of Phy can be detected in Fig. 3. Also, the course of the \(\Delta H_{\text{ads}}\) function found in Fig. 3, resembles that reported for lysozyme in a recent high-sensitivity calorimetric study (Smith et al. 2002). Hence, the results show no signs of particularly favorable hydration energy of glycan moieties. We consistently found an exaggerated experimental noise at \(h\approx0.13–0.20\) for both Phy and dgPhy (see Fig. 3). The origin of this could not be established, but it may reflect lyotropic (hydration dependent) rearrangements in the freeze-dried protein.

---

\(^1\)The observation that the adsorption data did not depend on the amount of protein in the cell (3–16 mg) supports the assumption of equilibrium.
In conclusion we have applied a two-channel perfusion calorimetric method to the investigation of a glycoprotein and its carbohydrate-depleted variant. Comparison of these two-model systems shows that the peptide groups adsorb more water than the glycans. In other words, the free energy of interaction with water is more favorable for the peptide than the glycan moieties of Phy. Moreover, the adsorption energy, $\Delta H_{\text{ads}}$, is similar for the two variants and close to that observed for other model proteins. While specific hydration effects at still higher hydration levels (i.e., in dilute solution) cannot be ruled out, these observations suggest that particularly favorable interactions of water and glycans cannot be the main reason for the high solubility and limited aggregation propensity observed for glycosylated proteins.

Acknowledgements We are thankful to Dr. K. Liltorp and Dr. L. L. Hansen for important suggestions and contributions to this work. We would also like to acknowledge the suggestion by Dr. H. Binder and Dr. H. Heerklotz to implement the two-channel calorimetric technique. This work was supported financially by the Carlsberg Foundation, the Danish Technical Research Council, Novozymes A/S and the Danish National Research Council through the establishment of the MEMPHYS center of excellence.

References


Paper 3

Biophysical Chemistry, 129:251-258
Glycoprotein-surfactant interactions: A calorimetric and spectroscopic investigation of the phytase-SDS system

Heidi L. Bagger\textsuperscript{a,c}, Søren V. Hoffmann\textsuperscript{b}, Claus C. Fuglsang\textsuperscript{c}, Peter Westh\textsuperscript{a,*}

\textsuperscript{a} Dept. of Science, Systems and Models, Roskilde University, Building 18.1 P.O. Box 260, DK-4000 Roskilde, Denmark
\textsuperscript{b} Institute for Storage Ring Facilities, Aarhus University, Building 520, Ny Munkegade, DK-8000 Aarhus C, Denmark
\textsuperscript{c} Novozymes A/S, Kroghshøjvej 36, DK-2880 Bagsvaerd, Denmark

Received 2 April 2007; received in revised form 8 June 2007; accepted 10 June 2007
Available online 21 June 2007

Abstract

The interactions of sodium dodecyl sulfate (SDS) and two glyco-variants of the enzyme phytase from \textit{Peniophora lycii} were investigated. One variant (Phy) was heavily glycosylated while the other (dgPhy) was enzymatically deglycosylated. Effects at 24 °C of titrating SDS to Phy and dgPhy were studied by Isothermal Titration Calorimetry (ITC) and Synchrotron Radiation Circular Dichroism (SRCD) spectroscopy. Comparisons of results for the two variants were used to elucidate glycan–surfactant interrelationships.

The CD spectra suggested that both the native and the SDS-denatured states of the two variants were mutually similar, and hence that the denaturation process was structurally equivalent for the two glyco-variants. The denatured state was far from fully unfolded and probably retained a substantial content of native-like structure. Furthermore, it was found that the glycans brought about only a small increase in the resistance towards SDS induced denaturation. The SDS concentration required to denature half of the protein molecules differed less than 1 mM for the two variants.

The affinity for SDS of both variants was unusually low. The amount of bound SDS (w/w) at different stages of the binding isotherm was 3–10 times lower than that reported for the most previously investigated globular proteins. Analysis of the relative affinity of the glycan and peptide moieties suggested that the carbohydrates bind much less surfactant. At saturation, glycans adsorbed about half as much SDS (in g/g) as the peptide moiety of Phy and about five times less than average proteins.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Surfactant-protein interactions; Thermodynamics; Glycosylation; Stability; Circular dichroism; Titration calorimetry

1. Introduction

Understanding the interactions of proteins and solutes affecting their physical properties is essential for improved formulation and application of industrial- and pharmaceutical proteins. Surfactants constitute an important and interesting class of such solutes. Thus, the broad range of chemical structures within this group of additives enables modulation of for example protein solubility and aggregation behavior [1]. Owing to their amphiphilic nature, surfactants tend to adsorb at interfaces or mutually associate (form micelles) to minimize the contact of their hydrophobic part with water. This tendency also leads to an unspecific binding to proteins in solution, which in turn underlies the increased solubility and conformational instability of proteins in the presence of surfactants [2]. This relationship between weak binding on the one hand, and solubility and instability on the other, is rigorously established through the so-called linkage theory by Wyman and Gill [3].

To date, the most studied surfactant is sodium dodecyl sulfate (SDS). This anionic compound, which is generally considered a potent denaturant [4], was investigated in many early works on bovine serum albumin (BSA) [5–10] and a number of other proteins [11–17]. This led to several general conclusions regarding the modes of interaction of SDS and proteins (for reviews see [18–20]). Thus, a certain pattern of binding events has been observed before the adsorption of surfactant eventually reached saturation at about 1.4 g SDS per g protein with reduced S–S bonds [15,21] and somewhat less (~1 g/g) in proteins with intact cystines [21]. Some exceptions
from this picture, however, have been reported. Glucose oxidase, for example, binds very little SDS and (in a certain pH range) retains its catalytic activity at high SDS [12,22]. Also, we have recently found that the lipolytic enzyme cutinase binds only about 0.4 g SDS/g at saturation [23,24].

The current work investigates the interactions between SDS and the enzyme phytase (Phy) from *Peniophora lycii*. This is a glycoprotein (N-glycosylated on ten glycosylation sites) consisting of a 47.5 kDa peptide and a total of ~18 kDa carbohydrate, which mainly consists of oligomeric mannose [25–27]. To address the effects of glycosylation on the SDS-induced destabilization, and to determine the relative affinity of the detergent for the peptide and glyccan moieties, we compared data for Phy with results for an enzymatically deglycosylated form (dgPhy). The importance of glycoprotein-surfactant interrelationships has been pointed out already in the seminal work of Pitt-Rivers and Impiomba [21], but none of the nearly 500 papers quoting this publication has addressed the problem. Recently, progress has been made in the understanding of surfactant interactions with soluble oligosaccharides [28–31], but protein conjugated glycans remain practically unexplored. Increased knowledge in this area may play a role in the analysis of glycoproteins in SDS gel electrophoresis [32,33]. More importantly, information on glycan-detergent interactions appears to be of particular interest in attempts to apply glycosylation as a tool to modify the properties of proteins and peptides in biotechnology. The purpose of the current work is to elucidate some thermodynamic and structural effects of the glycans on Phy during exposure to SDS. The work also revealed an unusually low SDS-affinity of this protein, which will be briefly discussed.

2. Methods

*P. lycii* phytase was expressed in *Aspergillus oryzae*, purified according to previously published procedures [27], and stored at −25 °C in buffer solution (20 mM Tris-acetate, pH 6). The molecular mass of the glycosylated phytase (~65 kDa) and the purity of this enzyme (>95%) were determined by SDS–PAGE. The enzymatic deglycosylation of Phy producing dgPhy was performed by Endo-β-N-acetylglucosaminidase F₁ (Endo F₁) E.C.3.2.1.96 purchased from Calbiochem (Darmstadt, Germany). The deglycosylation procedure is described in details elsewhere [25]. To ensure that the deglycosylation was quantitative the reaction mixture was investigated by SDS–PAGE. This showed one narrow band at ~48 kDa in accordance with a deglycosylated protein. Before the experimental trials, Phy and dgPhy were dialyzed (molecular weight cut of 12–14 kDa), at 4 °C, against milliQ water and subsequently freeze dried. All analytical (CD and ITC) was conducted with pure water as the solvent. Parallel samples of both protein variants were freeze dried in small vials and analyzed for residual water content by overnight exposure to 80 °C. In accordance with previous work on lyophilized phytase [26] no systematic weight loss could be detected in either of the protein variants. Based on this and the purity derived from the SDS–PAGE it was considered appropriate to derive the protein concentration from weighing of the lyophilized powders. SDS (>98%) was purchased from BDH laboratory (Poole, UK).

2.1. Isothermal titration calorimetry (ITC)

The freeze dried protein was dissolved in milliQ water (Spectrapore, VWR Scientific, San Francisco, CA). The protein solution was then loaded into the cell (1.43 ml) of a MicroCal VP-ITC instrument (Northampton, MA) and titrated with 110 mM SDS (dissolved in milliQ) from a 250 μl syringe. Four different protein concentrations in the 17–67 μM range were investigated for both Phy and dgPhy. Both protein and detergent solutions were degassed by stirring under vacuum prior to the experiment. The experimental temperature was 24 °C. The molar enthalpy change associated with SDS-protein interactions was calculated by integrating the raw calorimetric data (heat flow vs. time) using the Origin software, supplied with the instrument.

2.2. Synchrotron radiation circular dichroism (SRCD) spectroscopy

Stock solutions of Phy (~2 mg/ml) and dgPhy (~1.7 mg/ml) were prepared by redissolving the lyophilized powders in milliQ water. Aliquots of 500 μl stock were quantified by mass in eppendorf tubes and freeze dried in the tube. Fifteen minutes prior to the CD measurement, the proteins were hydrated in 500 μl SDS solution, giving a concentration of respectively 32 μM and 35 μM for Phy and dgPhy. Fourteen different Phy solutions and 12 dgPhy solutions covering the 0–20.4 mM SDS range were investigated.

The SRCD spectra were recorded at the UV1 beam line on the storage ring ASTRID, Institute for Storage Ring Facilities (ISA), University of Aarhus, Denmark. The instrument was carefully calibrated with respect to CD signal using camphor sulfonic acid (CSA) after each beam fill (once a day) of the storage ring. All spectra were recorded using a 0.1 mm path length quartz SUPRASIL cell (Hellma GmbH, Germany), over a wavelength range of 176–265 nm in 1 nm steps and a dwell time of 3 s per wavelength point. A pure solvent baseline collected with the same cell was subtracted all spectra which were processed and analyzed using the CDtool software package [34]. The machine unit (mdeg) was converted into the per residue molar absorption unit, delta epsilon (Δε) in M⁻¹·cm⁻¹, by normalization with respect to polypeptide concentration and path length. Secondary structure content analysis of the spectra was based on the SP29 (29 water soluble proteins) reference set from the soft ware package CDPro [35].

CD spectra for Phy were also measured in 5 °C intervals from 24 °C to 84 °C. The sample was equilibrated 5 min at each temperature prior to the measurement.

3. Results

The SDS induced structural changes of *P. lycii* phytase are illustrated by the CD spectra in Fig. 1. This figure shows Phy spectra in 14 different solvents covering the 0–20.4 mM SDS...
range. No significant changes could be detected in the seven spectra with a SDS concentration of 4.0 mM or less. Higher concentrations brought about spectral changes particularly a decreased ellipticity at about 189 nm and 206 nm and a concomitant increase in $\Delta\varepsilon$ around 222 nm. The changes around 206 nm observed at intermediate SDS concentrations resemble those reported earlier for other proteins [16,36]. The changes in the CD spectra of Phy continue until about 9–10 mM SDS. Hence, the CD data suggests that the native structure of Phy remains intact until the SDS concentration exceeds $\sim$ 4 mM, and that the structural transition is completed when [SDS] $\sim$ 10 mM.

Analogously the structural changes of the deglycosylated enzyme are studied by SRCD. Fig. 2 shows CD spectra for dgPhy in 12 solvents (0–20.4 mM SDS). The general trends in the spectral changes around 189 nm and 206 nm appear to be similar to those for Phy, however, the experimental scatter is larger in these trials and the increase in $\Delta\varepsilon$ around 222 nm is not clear for dgPhy (see below). Deviation from the native CD spectrum of dgPhy is observed at $\sim$ 3 mM SDS, thus detectable structural changes of dgPhy starts at a lower SDS concentration compared to Phy. No significant changes in the four spectra from 8.1–20.4 mM SDS could be discerned.

Absolute values of $\Delta\varepsilon$ are somewhat smaller for dgPhy at the lower SDS concentrations (Fig. 2) and this behavior is paralleled by the UV-absorption of the same samples (measured simultaneously in the SRCD instrument). Thus, the effective protein concentration is lowered in the deglycosylated samples, most likely due to some aggregation and surface adsorption of dgPhy [37, 38]. As SDS increases solubility and inhibits aggregation, the differences in the spectral intensities decrease with increasing SDS concentration, and for the highest surfactant concentrations, no differences between the two enzyme variants were detected. In pure water, $\Delta\varepsilon$ for dgPhy was consistently lower than for Phy, but application of a scaling factor provided practically superimposed spectra (Fig. 3). This suggests that the structural transition associated with SDS-induced denaturation is similar for the two variants and that the (secondary) structure of both the native and the SDS denatured state of phytase is unaffected by glycosylation. For the native conformation, this observation is expected as the catalytic activity has been shown to remain unchanged following deglycosylation [25,26]. The spectra in Figs. 1 and 3 exhibit quasi isodichroic point around 212 nm suggesting a two-state transition. Owing to the scatter in

**Fig. 1.** CD spectra of Phy in fourteen different solvents ranging from pure water to 20.4 mM SDS. The actual concentrations: 0, 0.2, 0.5, 1.0, 1.5, 3.0, 4.0, 6.0, 8.1, 9.0, 10.1, 10.2, 15.3 and 20.4 mM SDS. The $\Delta\varepsilon$ values change with increasing [SDS] in the direction of the arrow. [Phy] is 32 $\mu$M in all solutions. The inset represents an enlargement of the 179–202 nm range. The seven spectra (black) in the 0–4.0 mM SDS range are overlapping (to within the experimental precision). The red, blue and green curves represent respectively 6.0, 8.1 and 9.0 mM SDS. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Fig. 2.** CD spectra of dgPhy in twelve different solvents: 0, 0.2, 2.0, 3.0, 4.0, 5.1, 6.0, 7.1, 8.1, 10.1, 15.3 and 20.4 mM SDS. The $\Delta\varepsilon$ values change with increasing [SDS] in the direction of the arrow. [dgPhy] is 35 $\mu$M in all solutions. The inset is an enlargement of the 200–221 nm range. The spectra for 0 (black), 0.2 mM (red) and 2.0 mM (green) do not differ significantly. Structural changes induced by SDS can be detected in the spectra for higher SDS concentrations (3.0 mM (yellow), 4.0 mM (blue) and 5.1 mM (pink)). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Fig. 3.** CD spectra of Phy (circles) and dgPhy (triangles) in two different solvents: 0 mM SDS (filled) and 20.4 mM SDS (open). The recorded CD signal from dgPhy in the SDS free solution is multiplied by a factor of 1.3 (see main text). The superposition of SDS-free samples supports the conclusion that the native state of Phy is not affected by glycosylation. The scaling factor has only been applied in this figure (and only for dgPhy in pure water).
dgPhy spectra for low [SDS], a quasi isodichroic point is only found for SDS concentrations above 2 mM in Fig. 2.

To further illustrate the structural transition of Phy and dgPhy, and its cooperativity, the intensities of selected CD bands were plotted as a function of the total SDS concentration, [SDS]$_{tot}$, in Fig. 4. Panel A of this figure shows $\Delta \varepsilon$ for Phy at 189, 206 and 222 nm. This data is in accordance with a two state transition with a midpoint of about 7–8 mM SDS (i.e. half of the population of protein molecules denatured at this concentration). As discussed above, the CD data for dgPhy shows more scatter, but the results in Fig. 4B support a parallel interpretation with a midpoint value of 5–6 mM SDS.

In Fig. 5, isotherms from the ITC trials illustrate the enthalpic effects of SDS-protein interactions. The curve represented by crosses (panel A) is a reference experiment (SDS into pure water). This was used to assess the heat of dilution (and de-micellization) of the 110 mM SDS titrand and to determine the critical micelle concentration (CMC). The experimental temperature (24 °C) was chosen so that the injection of the first few aliquots of SDS was athermal ($\Delta H \sim 0$ in Fig. 5). Hence, the correction due to heats of dilution was as small as possible. Nevertheless, the heat flow associated with the dissolution of the micelles is concentration dependent and a small endothermic increase in the reference curve (crosses) builds up as the SDS concentration increases (Fig. 5A). The steeper endothermic increase in the $\sim 7.1$–$9.2$ mM SDS interval, designate the micellization process. Thus, the concentration of SDS in the calorimetric cell is raised to CMC and additional micelles will not be dissolved; c.f. [23,39]. The CMC value of SDS in the protein free solution is best described by the inflexion point of this part of the curve [39]. Four separate trials in pure water at 24 °C showed CMC $= 8.1 \pm 0.5$ mM SDS, which is in good accordance with recent data based on other experimental techniques [40].

Fig. 4. $\Delta \varepsilon$ at selected wavelengths for Phy (A) and dgPhy (B) plotted as a function of the SDS concentration. Panel A shows the intensity of the 189 nm (filled circles, left ordinate), 206 nm (open circles, right ordinate) and 222 nm (triangles, right ordinate) bands. Panel B analogously depicts the 189 nm and 206 nm bands for dgPhy.

Fig. 5. The enthalpy change, $\Delta H$ (kJ/mol SDS) associated with the titration of 110 mM SDS into different enzyme solutions at 24 °C. In panel A, The abscissa is the total (bound + free) concentration of SDS in the calorimetric cell. The upper trace (crosses) is a reference experiment (SDS into pure water). Solutions of Phy are represented by filled symbols and dgPhy solutions by open symbols. Increasing protein concentration follows the sequence: circles, squares, triangles and diamonds. The actual concentrations were 17.1, 30.3, 50.1 and 66.6 μM for Phy, and 17.1, 30.4, 50.4 and 66.9 μM for dgPhy. The inset in panel A is an enlargement of the dilute region (0–0.8 mM SDS). In panel B, the data is shown as a function of the molar ratio MR=mol SDS/mol protein in the calorimetric cell (for clarity data for $\sim 50$ μM is not included in panel B).

Titrations of SDS into solutions with different concentrations of Phy (filled symbols) and dgPhy (open symbols) are also shown in Fig. 5. The enthalpy changes, $\Delta H$, are plotted as a function of respectively the total SDS concentration (Panel A) and the protein:surfactant molar ratio, MR (Panel B). The latter representation is most commonly used in ITC work, but we have recently argued that isotherms with the total surfactant concentration on the abscissa provide some advantages for the analysis [23]. The results share a number of common traits. Thus, a conspicuous exothermic effect is observed at the lower SDS concentrations. The effect is limited to [SDS] $< 1$ mM and most likely reflects the specific binding of a few detergent molecules to the native protein[18,41–43]. To assess this further, we analyzed the data for 0–1 mM SDS by the conventional mass-action approach, using the Origin software package delivered with the ITC instrument [44]. The simple model accounted well for the measurements (below 1 mM) and
suggested that both Phy and dgPhy bound 3–4 SDS molecules with an affinity corresponding to a dissociation constant of 0.1–0.5 mM. The binding enthalpy was −5 to −10 kJ/mol SDS. Although binding in this affinity range can be measured only approximately at the current protein concentrations, we note that the derived parameters are typical for the specific binding of SDS to proteins [18, 41]. As illustrated in the inset of Fig. 5A, the enthalpy changes for the specific binding of SDS to dgPhy is almost twice that for Phy which suggests more exothermic binding to the deglycosylated enzyme.

At intermediate SDS concentrations (~2–5 mM) we observe only weak thermal effects of SDS titration (Fig. 5). At higher concentrations (~5–10 mM), corresponding to the transition range determined by SRCD (Fig. 4), the heat signal decreases (becomes increasingly exothermic) in a near-linear fashion. We ascribe this change to the cooperative break-down of the native protein structure and the concurrent binding of detergent to the additionally exposed protein surface. This part of the isotherm is displaced to the left for dgPhy relative to Phy, thus illustrating (in accord with the SRCD data) that the denaturation occurs at slightly lower SDS concentration when glycans have been removed.

At still higher SDS concentrations (~9–12 mM SDS, depending on the enzyme concentration), the isotherms curve upwards in a sigmoid manner and eventually all merge with the control experiment (Fig. 5A). This behavior has been observed for other proteins [23, 24, 41] and ascribed to the completion of the surfactant adsorption process. This saturation is governed by the appearance of micelles in the aqueous bulk (see below) and thus essentially reflects an “apparent CMC” in the presence of the protein. Hence, in accordance with the suggestions of Paula et al. [39] we use the inflexion point (D in Fig. 6) to identify saturation. The molecular origins of other characteristic points in the isotherms were inferred from the SRCD data and identified in Fig. 6 as point A, B and C (reflecting respectively onset-midpoint- and completion of the denaturation process).

More quantitative data may be derived from the ITC results by applying a simple partitioning model which considers the equilibrium between dissolved (free) and protein bound detergent [23].

\[
[\text{SDS}]_{\text{tot}} = [\text{SDS}]_{\text{free}} + [P] \cdot n
\]  

where \([\text{SDS}]_{\text{free}}\) is the concentration of SDS monomers in the aqueous bulk, \(n\) denotes the binding number (SDS/protein) and \([P]\) is the concentration of protein (Phy or dgPhy). Eq. (1) is valid when \([\text{SDS}]_{\text{free}}\) does not exceed CMC. Values of \([\text{SDS}]_{\text{tot}}\) at the saturation point (D in Fig. 6) were read off enlarged copies of Fig. 5A and plotted as a function of \([P]\) in Fig. 7. It appears that the partitioning scheme, which predicts a linear relationship (Eq. (1)), accounts well for the results. The parameters \((n, [\text{SDS}]_{\text{free}})\) were \((69 \pm 2, 8.2 \pm 0.10 \text{ mM})\) and \((57 \pm 2, 8.1 \pm 0.11 \text{ mM})\) respectively for Phy and dgPhy. Uncertainties are given as standard errors. It follows that Phy adsorbs 20% more SDS (in mol/mol) than dgPhy at saturation. Moreover, these parameters show that saturation is reached exactly when the bulk (free) concentration of SDS is equal to CMC (for \([P] = 0\), the lines in Fig. 7 extrapolate to the CMC (8.2 mM) obtained in the reference experiments). This clearly indicates that saturation in both systems is governed by the activity of SDS in solution (which becomes practically constant

![Fig. 6. Assignments of the molecular origins of certain characteristics in the enthalpograms discussed in the main text. A: onset of denaturation (a detectable fraction of the protein molecules is in the denatured form). B: denaturation midpoint (equal concentration of native and denatured protein). C: denaturation completed. D: saturation of SDS adsorption.](image)

![Fig. 7. Total (bound+free) concentration of SDS, \([\text{SDS}]_{\text{tot}}\), at the saturation point plotted as a function of the enzyme concentration \([P]\). Data for Phy and dgPhy are indicated by filled and open symbols respectively.](image)

<table>
<thead>
<tr>
<th>Phy</th>
<th>([\text{SDS}]_{\text{free}}) (mM)</th>
<th>(n)</th>
<th>([\text{SDS}]_{\text{free}}) (mM)</th>
<th>(n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D: Saturation of SDS</td>
<td>69 ± 2</td>
<td>8.2 ± 0.1</td>
<td>57 ± 2</td>
<td>8.1 ± 0.1</td>
</tr>
<tr>
<td>C: Denaturation complete</td>
<td>61 ± 5</td>
<td>7.8 ± 0.2</td>
<td>44 ± 6</td>
<td>7.4 ± 0.3</td>
</tr>
<tr>
<td>B: Denaturation midpoint</td>
<td>31 ± 2</td>
<td>6.9 ± 0.1</td>
<td>28 ± 4</td>
<td>6.2 ± 0.2</td>
</tr>
<tr>
<td>A: Initial denaturation</td>
<td>8 ± 3</td>
<td>5.4 ± 0.1</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Based on the combined interpretation of the SRCD and ITC results, four different molecular assignments were made for the characteristic features in the enthalpograms in Fig. 5. These assignments (discussed in the main text) are listed in the first column. Values of \([\text{SDS}]_{\text{tot}}\) and \([P]\) in Fig. 5A were used together with Eq. (1) to find the binding number, \(n\) (in mol SDS/mol protein), and the free concentration \([\text{SDS}]_{\text{free}}\) (in mM) as exemplified in Fig. 7. Uncertainties are standard error of the linear regressions.
above CMC) rather than other limitations such as full coverage of the protein interface (c.f. [20]).

A corresponding analysis of points A, B, and C (in Fig. 6) gave the parameters listed in Table 1. The binding number at point B (the transition midpoint) is an average of the number of SDS bound to respective native (N) and denatured (D) protein.

4. Discussion

4.1. Binding isotherms

Glycosylation appears to have a potential to modulate the physical properties of proteins without compromising enzymatic activity[25,38,45–48]. This makes controlled glycosylation an interesting avenue for the engineering and formulation of pharmaceutical peptides and industrial enzymes. Progress in this area, however, will rely on knowledge on the interactions of glycans and the surrounding solvent, and on how these interactions govern the properties of the glycosylated protein.

Interrelationship of SDS binding and deglycosylation of Phytase can be demonstrated by binding isotherms of the two glyco variants (Fig. 8). The binding isotherms are derived by combining the data in Table 1 and the results for the specific binding. The ordinate is the mass of bound SDS per mass unit of enzyme. If the affinity (per unit mass) of glycans and peptide were similar, the curves in Fig. 8 would be superimposed. It appears, however, that dgPhy adsorbs more SDS (in g/g) than Phy, i.e. that SDS interacts more favorably with the peptide. At saturation dgPhy binds 0.34 g/g whereas Phy binds 0.31 g/g. Under the assumption that the amount of SDS bound to the peptide moiety of Phy is equal to the amount adsorbed by dgPhy it appears that SDS adsorption by glycans corresponds to 0.20 g/g at saturation.

As mentioned in the Introduction, polypeptides typically adsorb 1–1.4 g SDS/g at saturation. Thus, dgPhy exhibits an unusual low affinity for this surfactant. Moreover, if the degree of SDS adsorption by glycans found here (0.2 g/g) is general, the difference between peptides and glycans might be more pronounced for other glycoproteins. This suggestion is in accord with the study by Pitt-Rivers and Impiombato[21], who found that saturation occurred at around 1.0 g/g peptide for a range of different proteins whether glycosylated or not. While the contribution from the glycans cannot be unambiguously discerned from the variability among the different proteins studied by these workers, this result indeed suggests very low SDS binding to glycans. A similar conclusion was reached by Russ and Polakova on the basis of mobility of glycoproteins in SDS–PAGE[33]. The current results are also in accord with recent data on the binding of SDS to soluble oligosaccharides. Thus, maltodextrin with 10 glucose units (i.e. a sizes comparable to the glycans on Phy) bound 1.5 mM SDS in 0.5% (w/v) maltodextrin [29,30], which translates into about 0.1 g SDS/g carbohydrate. Control ITC experiments (not shown) on solutions of pure (monomeric) mannose did not reveal any interaction with SDS, in line with more systematic work on this type of systems [30].

4.2. Protein structure

The combined analysis of CD and ITC data suggests that both forms of phytase specifically bind 3-4 SDS molecules
without detectable changes in the protein structure. This process most likely involves the interaction of the anionic surfactant and cationic groups on the protein, particularly if this binding is located around hydrophobic areas or clefts where the surfactant acyl chain can be lodged [18].

At intermediate SDS concentrations (1–5 mM) Phy interacts only weakly with SDS and the surfactants do not induce major changes in the protein structure. The binding number towards the high end of this interval corresponds to \( \sim 0.04 \) g/g which is about an order of magnitude less than typically observed values \([15,19,41]\). Above 5 mM SDS, the regime of weak interaction changes into another mode of interaction in which the heat signal becomes increasingly exothermic (Fig. 5A), the binding number increases steeply (Fig. 8) and the native structure changes (Figs. 1 & 2). This is interpreted as a typical denaturant induced transition, where the population of protein molecules is gradually brought from the N to the D state as the concentration of denaturant increases. Both methods suggest that the N–D transition is completed approximately as the free concentration of SDS reaches CMC, and this implies that formation of surfactant aggregates in the bulk promotes denaturation. This type of micelle-driven unfolding has been reported for a number of other proteins[43,49–51].

The change in secondary structure accompanying the SDS-denaturation was investigated by analysis of the CD data for secondary structure contents using a protein reference set from the soft ware package CDPro [35]. This analysis estimates the amount of \( \alpha \)-helix, \( \beta \)-sheet, turns and unordered secondary structures, and suggested about 28% of \( \alpha \)-helix and 19% of \( \beta \)-sheet in the native Phy (i.e. for [SDS]=0). These values are reasonably consistent with the crystal structure which has 33% helix and 11% sheet (Dr. L.DeMaria, Novozymes, personal communication). Denaturation by SDS only changes the secondary structure moderately (Fig. 9). Hence, the completely SDS-denatured protein (in >10 mM SDS) has a composition of \( \sim 22\% \) helix and 24% sheet. No increase in the fraction of unordered chain was identified in the structural analysis, and it appears that the SDS-denatured state of Phy is far from disordered. Hence, it is suggested that increased SDS-binding upon denaturation primarily is coupled to changes in the tertiary protein structure. Residual structure (at room temperature) in SDS denatured states of proteins is not uncommon. Serum albumin [19] and a number of other proteins show different degrees of secondary structure, when exposed to high concentrations of ionic surfactants[11,16,52–56].

To compare the structural effects of SDS-denaturation with temperature induced denaturation SRCD data from temperature scans of Phy (data not shown) were also analyzed for secondary structure content. Fig. 9 shows the estimated amount of \( \alpha \)-helix, \( \beta \)-sheet, turns and unordered secondary structures in Phy as function of respectively the SDS concentration (left panel) and temperature (right panel). It is clear that thermal denaturation induces more pronounced changes in the secondary structure. Large structural perturbations occur between 59–69 °C which is consistent with the measured denaturation temperature \( (T_d=62 \, ^\circ C) \) determined by DSC prior to the CD investigations and in previous work \([25,38]\).

5. Conclusions

The thermodynamic and spectroscopic data collectively suggests that the glycan “mantle” of phytase only confers a limited positive effect on the resistance towards SDS induced destabilization. We suggest that this relies on rather weak carbohydrate–surfactant interactions, and on peptide-surfactant interactions, which are mostly unaffected by the presence of glycans. At saturation, the adsorbed amount of SDS to the carbohydrate moieties was about 0.2 g SDS/g glycan or about 5 times less than the value found for average polypeptides. Interestingly, binding of SDS to the peptide part of phytase is also weak throughout the investigated concentration range. The binding number for dgPhy at saturation is 3–4 times less than typically reported values. Together with a few other examples \([12,22–24]\) this observation challenges the general view of an universal (mass/mass) saturation binding level for SDS. The low binding numbers are suggested to be a consequence of a strong potential of this protein to preserve secondary structure in SDS which limits the area accessible to the surfactant.

Acknowledgements

This work was supported by the Danish National Science Foundation through the establishment of the MEMPHYS centre of excellence. The financial support by the Carlsberg Foundation, Novozymes Ltd. and the Danish Research Agency (grants 26-02-0160 and 21-04-0087) is gratefully acknowledged.

References

C.A. Nelson, Binding of detergents to proteins.1. maximum amount of dodecyl sulfate bound to proteins and resistance to binding of several proteins, J. Biol. Chem. 246 (1971) 3895–3901.


Paper 4


Article in press, Biophysical Chemistry, doi:10.1016/j.bpc.2007.06.014
Solute effects on the irreversible aggregation of serum albumin

Heidi L. Bagger a,c,1, Lars H. Øgendal b, Peter Westh a,⁎

a Roskilde University, Department of Science, Models and Systems, Universitetsvej 1, DK-4000 Roskilde, Denmark
b University of Copenhagen, Department of Natural Sciences, Faculty of Life Sciences, Thorvaldsensvej 4, DK-1871 Frederiksberg C, Denmark
c Novozymes A/S, Kroghshøjvej 34, 2880 Bagsværd, Denmark

Received 15 April 2007; received in revised form 28 June 2007; accepted 28 June 2007

Abstract

Thermal stress on bovine serum albumin (BSA) promotes protein aggregation through the formation of intermolecular β-sheets. We have used light scattering and chromatography to study effects of (1 M) Na2SO4, NaSCN, sucrose, sorbitol and urea on the rate of the thermal aggregation. Both salts were strong inhibitors of BSA aggregation and they reduced both the size and number (concentration) of aggregate particles compared to non-ionic solutes (or pure buffer). Hence, the salts appear to suppress both nucleation- and growth rate. The non-electrolyte additives reduced the initial aggregation rate (compared to pure buffer), but did not significantly limit the extent of aggregation in samples quenched after 27 min. heat exposure (40–50% aggregation in all samples). The non-electrolytes did, however, modify the aggregation process as they consistently brought about smaller but more concentrated aggregates than pure buffer.

The results are discussed along the lines of linkage- and transition state theories. In this framework, the rate of the aggregation process is governed by the equilibrium between a thermally denatured state (D) and the transition state D≠. Thus, the effect of a solute relies on its preferential interactions with respectively D and D≠. The current results do not show any correlation between the solutes' preferential interactions with native BSA and their effect on the rate of aggregation. This suggests that non-specific, “Hofmeister-type” interactions, which scale with the solvent accessible surface area, are of minor importance. Rather, salt induced suppression of aggregation is suggested to depend on the modulation of specific electrostatic forces in the D≠ state.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Multi angle light scattering; Size exclusion chromatography; Linkage theory; Transition state theory; Hofmeister series; Preferential interactions

1. Introduction

The physical instability of proteins in solution poses a major challenge for their technological and pharmaceutical utilization. One way of addressing this problem is to stabilize the protein’s secondary and higher order structures through the addition of solutes such as certain sugars, polyols, salts and amino acids. For equilibrium processes, the way these changes in solvent composition modulates protein stability can be rationalized through the so-called Wyman relationships for linked equilibria [1]. One key aspect of this theory stipulates – along the lines of Le Chatelier’s principle – that addition of a solute will favor the state with which it has the strongest (most favorable) interactions. More specifically the solute will change the equilibrium constant K for a denaturation equilibrium (N ⇆ D) according to Eq. (1)

$$\left( \frac{\partial \ln K}{\partial \ln a_3} \right)_{T, p, a_2} = \nu_D - \nu_N = \Delta \nu$$

where $a_3$ is the thermodynamic activity of the solute. The function $\nu$ is the preferential binding parameter, $\nu = (\partial m_3 / \partial m_2)_{T, p, a_2}$, where $m_2$ and $m_3$ denotes the molal concentration of respectively protein and solute. The preferential binding parameter can be experimentally assessed e.g. through dialysis equilibrium measurements, and this approach has been successful in attempts to understand solute effects on various protein equilibria [2–6]. In many cases, however, the loss of biological activity in protein
solutions does not progress reversibly, but relates to changes under kinetic control far from equilibrium. In these cases, it is challenging to establish a framework for the understanding of solute effects. However, one fundamental starting point, which has proven useful (for review see e.g. [7,8]), is the two-step molecular picture originally introduced by Lumry and Eyring [9].

\[ N \leftrightarrow D \] (2a)

\[ D + A_n \rightarrow A_{n+1} \] (2b)

The first step (Eq. (2a)) is an equilibrium transition (\( \Delta G_{2a} < 0 \)) from the native (N) state to a non-native, aggregation competent conformation, D. The second (Eq. (2b)) is an irreversible growth step (\( \Delta G_{2b} > 0 \)) in which the D state merges with a pre-established aggregate. Combining Eqs. (1), (2a) and (2b) suggests that solutes for which \( \Delta \nu > 0 \) are likely to slow down the rate of aggregate formation since they displace the equilibrium (2a) towards the N state and thus lower the concentration of one “reactant” in (2b). Analogously, solutes which displace (2a) towards the right (i.e. \( \Delta \nu > 0 \) so that \( K \) is increased) may be expected to accelerate the observed rate of aggregation. This view suggests a direct correlation between the stabilization of N in (2a) and the kinetically controlled life-time of the active protein. Many reports have been in line with this interpretation both for stabilizing solutes such as sucrose (\( \Delta \nu > 0 \)) and destabilizing solutes such as urea (\( \Delta \nu < 0 \)) [7,10–15]. However, this strict coupling between equilibrium and kinetic stability is not universal and may be valid only at conditions where \( K \) is below or close to unity (i.e. when the N state is significantly populated) [16]. Moreover, the effects of solutes may be particularly complex as pointed out by Eronia et al. [11], who reported that additives which stabilized the N state of glucagon phosphorylase also promoted the rate of the protein’s aggregation. Dual effects of solutes have also been reported for denaturants which promote aggregation and/or precipitation at low concentrations but inhibits in it somewhat more concentrated solutions [10,12,15]. One plausible interpretation of these dual effects is that the solute modulates the two steps in Eqs. (2a) and (2b) oppositely.

In the light of this, studies on possible interrelationships between preferential interactions and aggregation kinetics appear to be of interest. To this end we have investigated the effects of selected solutes (for which information on \( \Delta \nu \) and/or \( \nu_N \) is available) on the thermal aggregation of bovine serum albumin (BSA). To emphasize solute effects on the irreversible step (Eq. (2b)), we have studied the aggregation kinetics at the temperature \( T_d \) (determined separately by DSC), where the extent of reaction (2a) is 0.5 and hence \( K = 1 \) for all solutes (i.e. at slightly different temperatures for different types and concentration of solute). Hence, the initial D-concentration was expected to be equal in all trials. At neutral pH, the irreversible step for the thermal denaturation of BSA controlled by the formation of soluble “β-aggregates” in which a moderate part of the native helices are converted into intermolecular β-sheets which connect the protein monomers [17]. Under these conditions the aggregates are small, without pronounced polydisperse and their growth is governed by a reaction-limited mechanism [17–19]. These properties make BSA an adequate model system for studies of solute effects on irreversible transitions.

2. Methods

Albumin from bovine serum (BSA, >98%) was purchased from Sigma (St. Louis, USA). Prior to the experiments, BSA was dialyzed (MWCO 12–14 kDa), at 4 °C, against milliQ water and subsequently freeze dried. Sodium sulphate, Na2SO4, (>99.0%); sucrose (>99.5%) and sodium thiocyanate, NaSCN (>98%) were purchased from Fluka (Steinheim, Germany). Urea (>98%) and sorbitol (>98%) and succinic acid (>99%) was from Sigma. All solutions for calorimetry and light scattering measurements were prepared in 20 mM succinate buffer, pH 7.0. The freeze dried protein was quantified gravimetrically and dissolved in the succinic acid buffer to a stock solution with a protein concentration of 5.2 mg/ml. The investigated BSA samples were prepared by mixing this stock and solute–buffer mixtures to a final BSA concentration of 1.04 mg/ml.

Differential Scanning Calorimetry (DSC) was used to determine the denaturation temperature of BSA in the different solvents. The instrument was a SCAL-1 (Puschino, Russia). The protein solution was loaded into the sample cell (331 μl) and pure solvent into the reference cell. Both protein- and reference solution were degassed by stirring under vacuum prior to the experiment. The temperature was raised from 6 °C to 90 °C at a rate of 2 °C/minute under an excess pressure of 2 atm. The denaturation temperature, \( T_d \), was defined at the maximum value of the heat capacity in the transition peak.

Static light scattering. The bulk aggregation kinetics of BSA was monitored in real time at the temperature \( T_d \). We used a home built instrument for multi angle simultaneous static and dynamic light scattering [20]. The instrument uses a Nd-YAG laser (ADLAS325) with a wavelength of 532 nm and a power of 150 mW. The sample temperature is controlled to within 0.1 °C. For the current measurements, only the photon count rates recorded at 90° scattering angle were used.

Size exclusion chromatography with multi angle static light scattering and refractive index detection (SEC-MALLS) was used to characterize quenched samples. Aliquots of 500 μl were injected onto a 300 mm Superdex 200 column (Amersham Pharmacia) and eluted at 0.5 ml/min. The elution buffer was 50 mM NaH2PO4/Na2HPO4, pH 7.0, 15 mM NaCl. The species that eluted from the column were first passed through a light scattering instrument (DAWN EOS, Wyatt Technology, Santa Barbara, CA) equipped with a K2 flow cell, and subsequently through a differential refractometer (RID-10A, Shimadzu, Japan) for concentration determination. The light scattering instrument has 18 detectors that measure the intensity of the scattered light at different angles. In principle all but the detector corresponding to the smallest scattering angle are accessible with the K2 flow cell. In practice, however, a varying number of detectors both at the smallest and at the largest scattering angles have to be disregarded because the signal to noise ratio at the
The distribution of monomers, dimers and aggregated proteins in the samples, which were quenched after 27 min at $T_d$, was investigated by SEC-MALLS. This method is particularly useful since the combined analysis of the RI and LS data provides information on the size and concentration of the aggregates even if they are too large to be resolved chromatographically. Fig. 2 shows examples of the SLS and RI signals from trials on both thermally denatured samples and samples which were kept refrigerated prior to the chromatography (referred to as reference samples). Panel A and B illustrate the effect of heat treatment of BSA in either pure buffer or 0.58 M sucrose. The RI signal (Fig. 2B) shows that in addition to the monomeric protein, the reference sample also contains a population (∼14% w/w) of a species with twice the molecular weight (determined from the ASTRA software introduced below). This is most likely the Cys34 disulphide dimer, which is commonly found in BSA preparations [25]. The SLS profiles of reference samples, represented by bold lines in Fig. 2A, show four peaks corresponding to respectively aggregates (retention time 13–17 min), some smaller oligomers probably tri-, tetra and pentamers (20–22 min), dimers (22–25 min) and monomers (25–29 min). There was no significant difference between chromatograms (SLS and RI) of the eleven reference samples. The thin lines in Fig. 2A and B represent heat-treated samples of BSA in pure buffer (solid line) and in 0.58 M sucrose (dotted line). These results show a clear heat-induced increase in the

---

**Table 1**

<table>
<thead>
<tr>
<th>Concentration (M)</th>
<th>Sucrose</th>
<th>Urea</th>
<th>Sorbitol</th>
<th>Na$_2$SO$_4$</th>
<th>NaSCN</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.29</td>
<td>0.58</td>
<td>0.37</td>
<td>0.74</td>
<td>0.31</td>
<td>0.62</td>
<td>0.22</td>
</tr>
<tr>
<td>0.58</td>
<td>0.37</td>
<td>0.74</td>
<td>0.31</td>
<td>0.62</td>
<td>0.22</td>
<td>0.44</td>
</tr>
<tr>
<td>0.37</td>
<td>0.74</td>
<td>0.31</td>
<td>0.62</td>
<td>0.22</td>
<td>0.44</td>
<td>0.36</td>
</tr>
<tr>
<td>0.22</td>
<td>0.44</td>
<td>0.36</td>
<td>0.74</td>
<td>0.72</td>
<td>0.36</td>
<td>0.73</td>
</tr>
<tr>
<td>0.36</td>
<td>0.74</td>
<td>0.36</td>
<td>0.72</td>
<td>0.36</td>
<td>0.73</td>
<td>0.73</td>
</tr>
</tbody>
</table>

The slope is in s$^{-1}$. The experimental temperatures ($T_d$ from the DSC trials) and slopes immediately after the lag period for the light scattering data in Fig. 1 are given in Table 1.

### 3. Results

#### 3.1. Calorimetry

Variations in $T_d$ were small. Thus, it increased 2 °C for sorbitol and sucrose over the concentration interval investigated here (0–0.6 M). A similar decrease was found for urea while no significant changes in $T_d$ were detected for Na$_2$SO$_4$ (0–0.4 M). The only solute that produced distinctive effects was NaSCN which increased $T_d$ by about 5 °C. This is highly unusual for this solute which is a strong chaotrope and accordingly commonly found to destabilize the native conformation [21]. However, an identical effect of thiocyanate on BSA has been reported before [22] and probably relates to a specific binding of the anion to native BSA. The values of $T_d$ in all solvents, which were used as the experimental temperatures for the light scattering measurements, are listed in Table 1.

#### 3.2. Static light scattering

The increase in the intensity of scattered light reflects the gradual aggregation of BSA in the sample. In the appendix we propose an approach for the analysis of the time course of the scattering. Fig. 1 shows that all trials share the common feature of a 2–3 min lag time in which no increase in NSI is observed. This behavior is typical to this type of measurements [8,17,23,24] and probably reflects the time required for thermal equilibration, equilibration of Eq. (2a) and nucleation of the aggregation process. Subsequent to the lag time, all solutes appear to inhibit thermal BSA aggregation as suggested directly from NSI values. This effect is particularly strong for the ionic solutes at the higher concentrations (>0.4 M). To get an estimate of the relative rates of aggregation, we determined the slope $d$NSI$/dt$ in the near-linear region (Fig. 1) just after the lag time. These values are listed in Table 1.

---

**Fig. 1.** Normalized scattering intensity, NSI, as a function of time for BSA in pure buffer (solid line) and in 0.58 M sucrose (dotted line). These results show a clear heat-induced increase in the

---

The distribution of monomers, dimers and aggregated proteins in the samples, which were quenched after 27 min at $T_d$, was investigated by SEC-MALLS. This method is particularly useful since the combined analysis of the RI and LS data provides information on the size and concentration of the aggregates even if they are too large to be resolved chromatographically. Fig. 2 shows examples of the SLS and RI signals from trials on both thermally denatured samples and samples which were kept refrigerated prior to the chromatography (referred to as reference samples). Panel A and B illustrate the effect of heat treatment of BSA in either pure buffer or 0.58 M sucrose. The RI signal (Fig. 2B) shows that in addition to the monomeric protein, the reference sample also contains a population (∼14% w/w) of a species with twice the molecular weight (determined from the ASTRA software introduced below). This is most likely the Cys34 disulphide dimer, which is commonly found in BSA preparations [25]. The SLS profiles of reference samples, represented by bold lines in Fig. 2A, show four peaks corresponding to respectively aggregates (retention time 13–17 min), some smaller oligomers probably tri-, tetra and pentamers (20–22 min), dimers (22–25 min) and monomers (25–29 min). There was no significant difference between chromatograms (SLS and RI) of the eleven reference samples. The thin lines in Fig. 2A and B represent heat-treated samples of BSA in pure buffer (solid line) and in 0.58 M sucrose (dotted line). These results show a clear heat-induced increase in the
amount of large aggregates and a concomitant decrease in the concentration of monomers and dimers. The chromatograms of all aggregated BSA solutions are shown in Fig. 2C and D. While present in the native samples, tetramers cannot be detected in the aggregated samples by either RI or SLS, and except the 0.73 M NaSCN solution, which shows a relative strong dimer peak (Fig. 2D), aggregation also seems to preferentially diminish the amount of dimers in all solutions. This enhanced tendency of oligomers to aggregate is in accord with the observation that BSA with a fluorophor covalently linked to cysteine 34 (preventing dimerization) exhibits reduced aggregation [19].

In general the aggregates are too large to be fully separated by this column, but their average molecular weight is reflected by the SLS data (Fig. 2C). Using the ASTRA software supplied with the SEC-MALLS instrument the average molecular weights of aggregates in the eleven different solvents were determined. The procedure for the evaluation of molecular weights deviated slightly from the recommend standard procedure for the DAWN light scattering instrument. First, baselines representing no light scattering or no RI signal are determined for all detectors. All subsequent calculations are then based on baseline corrected signals. The instrument has 18 LS detectors measuring the intensity of the scattered light at 18 different scattering angles, $\theta_1, \theta_2, ..., \theta_{18}$, but only detectors #9–14 (representing scattering angles 69.3°–121.2°) were used in the current analysis. The detectors do not have the same intrinsic sensitivity, so for each sample the central portion of the peak for the monomer (known to scatter isotropically) was used to normalize the detectors. This means that they should all yield the same signal as detector # 11 (90°) for the monomer peak after multiplication with individual correction factors $k_j$, $j=9,10,...,14$. Finally, an additional calibration factor, $K_{cal}$, common to all the detectors, is introduced. The value of this constant is determined by the requirement that the apparent molecular weight $M_{app}(\theta_j)$, i.e. one for each detector angle, calculated by the ASTRA software (Eq. (4)) all yield the correct molecular weight of the BSA monomer, 66.4 kD.

$$M_{app}(\theta_j) = K_{cal} \cdot \frac{\sum_i LS_i(\theta_j)}{\sum_i RI_i} \text{ where } j = 9,10,...,14 \quad (4)$$

In this equation $LS_i(\theta_j)$ is the baseline corrected signal from the light scattering detector at scattering angle $\theta_j$, and $RI_i$ is the
signal from the refractometer. The index, \( i \), in the sums has a range corresponding to the points constituting the monomer peak. In order to determine the molecular weight corresponding to the other peaks in the chromatograms Eq. (4) is again used to calculate the apparent molecular weights, where now the index \( i \) in the sums has a range corresponding to the peak whose molecular weight is being determined. The calibration constant \( K_{cal} \) and the normalization constants \( k_i \) have the values established by the procedure described above. The ASTRA program then determines the average molecular weight, \( M_{w} \), by extrapolating the value of \( M_{app}(\theta) \) to zero scattering angle using a Debye plot, i.e. assuming a linear relationship between \( M_{app}(\theta) \) and \( \sin^2(\theta/2) \).

For each peak in the chromatogram, specified by the range of the index variable \( i \), the corresponding weight averaged molecular weight is determined using Eq. (3) followed by the extrapolation as described above. Results for the aggregate peak (retention time \( \sim 15 \) min) given as average aggregation numbers, \( n = M_w/66.4 \, \text{kDa} \), are plotted as a function of the solute concentration in Fig. 3A. It appears that all investigated solutes bring about a decrease in the average aggregate size and that the ionic solutes, \( \text{Na}_2\text{SO}_4 \) and \( \text{Na SCN} \), exert the strongest effect. The non-ionic polyols sucrose and sorbitol cause a linear decrease in \( n \), while urea shows a minimum in Fig. 3A. In Fig. 3B, the solute systems are ranked according to their initial rate \( \text{dNSI}/\text{dt} \) (Table 1). This plot shows a gradual increase, which reveals that the aggregate size after 27 min. scales with the initial rate irrespectively of the chemical structure of the solute.

From a technological point of view, the most important parameter is not the aggregate size but the fraction of non-aggregated monomers (which may become biologically active if favorable conditions are reestablished). To elucidate this, we integrated the peaks in the RI chromatograms of respectively aggregated (Fig. 2D) and reference samples (bold lines in Fig. 2B). Before heat treatment the samples consisted of 82% (w/w) monomeric BSA, \( \sim 14\% \) dimers, and \( \sim 4\% \) higher order aggregates. The monomer peak decreased to different extends in different solvents upon heat treatment (Fig. 2) while the oligomers generally disappeared entirely. To illustrate this we calculated the mass fraction of BSA, \( f_A \), which had become part of larger aggregates during the 27 min exposure to \( T_d \).

The results in Fig. 4A and B show that \( \text{Na}_2\text{SO}_4 \) and \( \text{Na SCN} \) most effectively inhibit the thermally induced aggregation of BSA. For these solutes there is a clear correlation between the initial rates derived from real time light scattering and the fraction of BSA in aggregates in the quenched samples (Fig. 4B). For sorbitol, sucrose and urea, we find aggregation of 35–45% of the BSA, which is not significantly different from \( f_A \) in pure buffer. This suggests that in spite of the effects of these solutes on protein equilibria (sorbitol and sucrose as stabilizers, and urea as a denaturant) they exert little or no effect on the extent of BSA aggregation after 27 min.
Having established the amount and average size of the aggregates in the quenched samples (Figs. 3 and 4), we estimated the concentration of aggregate particles, [A], according to Eq. (5),

$$[A] = \frac{f_A \cdot [\text{BSA}]_{\text{tot}}}{M_w}$$  \hspace{1cm} (5)

where $[\text{BSA}]_{\text{tot}}$ is the total amount of BSA (1.04 mg/ml or 16 μM). Fig. 5 shows that [A] is ∼0.1–0.3 μM in the quenched samples, and that the ionic additives consistently decrease the particle concentration while the nonionic solutes have the opposite effect (Fig. 5A).

Taken together, Figs. 3–5 show that the ionic solutes are the strongest inhibitors of the investigated aggregation process—they induce both smaller and fewer BSA aggregates. In contrast, the non-electrolytes promote a higher concentration (Fig. 5) of smaller sized aggregates (Fig. 3), and thus exert limited effects of the overall degree of aggregation (Fig. 4).

### 3.3. Modeling of the real-time SLS

Above we have assessed the initial aggregation rates directly from the time course of the curves in Fig. 1 (subsequent to the lag time) by assuming proportionality of the scattered light intensity and the extent of aggregation. This approach has previously provided important insight into aggregation processes for other systems [8,11,12,23,24,26,27] but it is not unambiguous, for example because it neglects the fact that the intensity of the scattered light scales with the square of the molecular weight. To examine the consistency of the solvent-sequence used in Figs. 3B–5B, we derived an expression that correlated kinetic parameters and the intensity of the scattered light (see Appendix). The derivation was based on a number of approximations including the assumption of first order kinetics with respect to [D], constant particle concentration, [A], in the growth phase and rate limitation by the growth step, Eq. (2b). In general, non-linear regression by Eq. (A12) accounted well for the experimental data and Fig. 6 shows two representative examples. The kinetic parameter, $k_2$ (Eq. (A14)) derived from the regression analysis of the 11 solvents scaled with $d\text{NSI}/dt$, thus corroborating the sequence used in Figs. 3–5. We conclude that the initial slope in Fig. 1 may be used as a measure of the relative rate of aggregation in different solvents. The ratio $k_2/d\text{NSI}/dt$, however, was not constant and this calls for caution in the use of $d\text{NSI}/dt$ as an absolute measure of the (first order) rate constant for thermal aggregation of BSA. A more detailed analysis of the relationship of slopes in plots like Fig. 1 and rate constants has been given elsewhere [8].

### 4. Discussion

#### 4.1. Aggregation mechanism and nucleation rate

Previous studies on thermal denaturation of BSA in pure buffer have suggested that the protein forms “soluble” aggregates in which a moderate fraction of the native α-helix is transformed into intermolecular β-sheet [17,19,28,29]. A similar trend has been found for numerous other proteins [30–32]. At pH values close to neutral, the aggregation of BSA is
reaction controlled whereas it becomes faster and diffusion controlled if pH is lowered toward pl which is ~ 5 [17,18].

An initial lag time in light scattering data (Figs. 1 and 6) is typical for protein aggregation and reflects at least in part the nucleation process [23,33]. Accordingly, it has been suggested that the subsequent increase in scattering intensity is governed by the growth of embryo-aggregates formed during this lag period [8]. This simplification implies that the growth phase involves the adsorption of unfolded protein to a constant number of pre-established particles. The current data sheds some light on this, since the combined RI and MALLS analysis quantifies both the size (\(M_a\)) and concentration ([\(A\)]) of aggregate particles in the quenched samples. It appears that the salts decrease [\(A\)], while the non-ionic additives have the opposite effect (Fig. 5A). As all solutes decreased both the size of the aggregates after 27 min. (Fig. 3) and the initial aggregation rate (Table 1) one plausible interpretation is that the non-ionic additives promote the rate of nucleation in the lag phase, but retard the growth of particles at later stages. Hence, these solutions have a comparable large number of small aggregates. Addition of salts, on the other hand, inhibits both nucleation- and growth rates. This may suggest that although the BSA molecules carry the same net charge, nucleation and growth are both favored by electrostatic interactions, which are shielded by the ionic solutes.

4.2. Rate of aggregation

The pronounced effect of solutes on the rate of protein aggregation is illustrated in Table 1, which suggests that moderate concentrations (0.2–0.7 M) may change the initial rate of aggregation of BSA by more than an order of magnitude. To highlight the solute induced effect on the irreversible growth step (Eq. (2b)) all experiments were conducted at \(T_d\), where the initial concentration of the D-form is (approximately) half of the total BSA concentration. One inevitable limitation of this experimental approach is the variation in experimental temperature. However, corrections using the Arrhenius equation and an activation energy of 300 kJ/mol [34] revealed only minor modifications. The largest changes were observed for 0.74 M urea, which increased to approximately the level of the pure buffer and thiocyanate which fell to around the values for sulfate. Comparisons of the initial rates (corrected or not) do not reveal any clear correlations to literature values of \(\Delta v\) (Eq. (1)). For example, we did not find any significant difference between the stabilizer sorbitol (\(\Delta v < 0\)) and the denaturant urea (\(\Delta v > 0\)) [35–38] or between the two salt which are located in opposite ends of the Hofmeister series [39]. To examine this further we apply traditional transition state theory [40,41] for reaction controlled processes. If we assume that the irreversible step (Eq. (2B)) is rate limiting for the overall aggregation process we may expand Eq. (2a) to read.

\[
N \rightleftharpoons D = D^* \rightarrow A
\]  

In Eq. (6), \(D^*\) is the transition state or “activated complex”. The observed rate of the aggregation process is proportional to the concentration of \(D^*\), and the standard activation free energy of the process is \(\Delta G^\ddagger = -RT \ln K^\ddagger\), where \(R\) and \(T\) have their usual meanings and \(K^\ddagger\) is the equilibrium constant \([D^*]/[D]\). The advantage of this approach is that the effect of a solute can be analyzed by equilibrium theory. Thus, solute effects on the second (rate determining) equilibrium in Eq. (6) may be expressed as \(\partial \ln K^\ddagger / \partial \ln [A] = \Delta \nu^\ddagger\), in complete analogy to Eq. (1). This approach thus stipulates that a solute’s effect on the aggregation rate is governed by the difference of its preferential interaction with respectively the D and \(D^*\) state (\(\Delta \nu^\ddagger\)). The pronounced inhibition of aggregation elicited by the two investigated salts thus implies negative (and rather similar) values of \(\Delta \nu^\ddagger\). This consistency is interesting in the light of the disparity in their preferential interactions with typical proteins. Sulfate is a strong kosmotrope (\(\nu_N < 0\)) while thiocyanate is a strong chaotrope (\(\nu_N > 0\)) [21,42–44]. In many cases, the preferential binding parameter has been shown to be “non-specific” in the sense that it approximately scales with the solvent accessible surface area [4,38]. In such cases a preferentially excluded solute (\(\nu_N < 0\)) will displace any protein equilibrium towards the state with the lowest accessible area and vice versa. Since the two salts with different sign of \(\nu_N\) act rather equally, we conclude that their effect on the thermal aggregation of BSA relies on specific interactions with the transition state (i.e. the preferential binding parameter is not proportional to the solvent accessible area). One likely interpretation is that the transition state is stabilized by electrostatic interactions, which are weakened by the ionic additive so that \(\Delta G^\ddagger(\nu_N < 0)\) increases with the salt concentration. Also, the results for the non-electrolytes suggest that the D- and \(D^*\)-states have similar solvent accessible areas, and consequently that \(\Delta \nu^\ddagger\) will be close to zero for solutes interacting nonspecifically. This latter suggestion is in line with recent work by Rodriguez-Larrea et al. [45], who reported modest changes in the solvent accessibility of the transition state during the irreversible denaturation of a lipase. The limited importance of non-specific interactions is also emphasized by the rather constant extend of aggregation (Fig. 4) in samples with respectively sorbitol, sucrose and urea.

In closing we suggest that the combined use of linkage- and transition state theories offers a promising approach for systematic analysis and comparison of solute effects on protein aggregation rates — an area of considerable interest in biotechnology. Related suggestions has previously been put forward [7,46]. The current results show that for BSA, the lyotropic properties of a solute (reflected in its position in the Hofmeister series [39]) are of minor importance both for the rate of the irreversible step and for the size and concentration of aggregate particles in heat-treated samples.

Acknowledgements

This work was supported by Novozymes A/S through the establishment of the “Protein Stability Network”. Financial support by the Carlsberg Foundation, Danish Research Agency (grants 26-02-0160 and 21-04-0087 to PW) and the Ministry of Science Technology and Innovation is gratefully acknowledged.
Appendix A

In order to interpret the bulk SLS data a model for the aggregation and the resulting intensity of the scattered light is needed. We take as a starting point, the Lumry and Eyring model mentioned in the introduction

\[ N \equiv \frac{D}{K} = \frac{[D]}{[N]} \quad (A1) \]

\[ D + A_n \rightarrow A_{n+1} \quad (A2) \]

We will assume that the formation of aggregates proceeds through the formation of nucleation species, seeds, with a concentration \([A]\). All subsequent aggregation proceeds by addition of denatured protein monomer to already existing aggregates. Thus, the number of aggregates remains constant and equal to \([A]\). The aggregation process is assumed to be first order in the sense that

\[ \frac{dn}{dt} = k_2 \cdot [D] \quad (A3) \]

where \(k_2\) is a rate constant (unit \(M^{-1}s^{-1}\)). Throughout the aggregation process mass is conserved which can be expressed as

\[ [N] + [D] + n \cdot [A] = c_0 \quad (A4) \]

where \(c_0\) is the molar concentration of added protein when the solution is prepared. We now make a differential equation for \([D]\) by first differentiating and next rearranging Eq. (A4) giving

\[ \frac{dn}{dt} = -\frac{1}{[A]} \cdot \frac{d([N] + [D])}{dt} \quad (A5) \]

If the Equilibrium (A1) is fast enough that the native and denatured protein molecules remain in equilibrium while aggregation proceeds then Eq. (A5) can be combined with Eq. (A3) to give the differential equation

\[ \frac{d[D]}{dt} = -\frac{k_2[A]}{1 + 1/K} \cdot [D] \quad (A6) \]

having the general solution

\[ [D] = D_0 \cdot e^{-\lambda t} \quad (A7) \]

where

\[ \lambda = k_2[A]/(1 + 1/K) \quad (A8) \]

At infinite times no denatured or native protein is left and the all monomers are bound in aggregates with the aggregation number

\[ n_{\text{max}} = \frac{c_0}{[A]} \quad (A9) \]

Particles in solution scatter light in proportion to their molar concentration and the square of their molecular weight. Therefore the intensity of scattered light of the aggregating solution can be written as

\[ I_{\text{scat}} = k_{\text{scat}} \cdot (\left[N\right] + [D] + n^2[A]) = k_{\text{scat}} \cdot ((1 + 1/K)[D] + n^2[A]) \quad (A10) \]

where \(k_{\text{scat}}\) is a constant depending on instrumental details as well as properties of the investigated system but not on the aggregation state of the protein. Solving Eq. (A4) for \(n\) and using Eq. (A7) the scattered intensity (A10) can be written

\[ I_{\text{scat}} = k_{\text{scat}} \cdot \left((1 + 1/K)D_0 \cdot e^{-\lambda t} + \left(c_0 - (1 + 1/K)D_0 \cdot e^{-\lambda t}\right)^2/S_0\right) \quad (A11) \]

If the parameter \(\lambda_0 = k_2D_0\) is introduced Eq. (A11) can be rewritten

\[ I_{\text{scat}} = I_0 \cdot \left(\frac{\lambda_0}{\lambda_{\text{max}}} \cdot n_{\text{max}} \cdot e^{-\lambda t} + n_{\text{max}} \cdot \left(1 - \frac{\lambda_0}{\lambda_{\text{max}}} \cdot n_{\text{max}} \cdot e^{-\lambda t}\right)^2\right) \quad (A12) \]

This expression can be fitted to the bulk scattering data through adjustment of the free parameters \(I_0, \lambda_0, \lambda\) and \(n_{\text{max}}\). Notice that the expression works equally well with normalized and non-normalized scattering data.

Although not evident from the mathematical structure of the fitting expression (A12) it turns out that the parameter \(n_{\text{max}}\) correlates strongly with the other parameters. This problem can be circumvented because both the weight fraction \(f\) in the aggregated state and the value of the aggregation number \(n(t_0)\) has been determined in the SEC-MALLS experiments at time \(t_0 = 6120\) s = 27 min. Still assuming that all aggregation after \(t_0\) only increases the aggregation number (and not molar concentration of the aggregates) it is seen that

\[ n_{\text{max}} = \frac{n(t_0)}{f} \quad (A13) \]

which can be added as a constraint on Eq. (A12) while fitting.

When the fitting parameters have been determined, the physical parameters can be calculated using \(K = 1\). Since the measurements were done at \(T = T_{\text{me}}\) \([A] = \frac{c_0}{n_{\text{max}}}\) and

\[ k_2 = 2\frac{\lambda_{\text{max}}}{c_0} \quad (A14) \]

References


Please cite this article as: H.L. Bagger et al., Solute effects on the irreversible aggregation of serum albumin, Biophys. Chemist. (2007), doi:10.1016/j.bpc.2007.06.014.