

# Fundamentals of unfolding, refolding and aggregation of food proteins

Kerensa Broersen

Promotoren: **Prof. dr. R.J. Hamer**  
Hoogleraar in de technologie van graaneiwitten, Wageningen  
Universiteit

**Prof. dr. ir. A.G.J. Voragen**  
Hoogleraar in de levensmiddelenchemie, Wageningen Universiteit

Co-promotor: **Dr. H.H.J. de Jongh**  
Projectleider, Wageningen Centre for Food Sciences

Promotiecommissie: **Prof. dr. P.L. Mateo-Alarcón**  
Universidad de Granada, Spanje

**Prof. dr. I. Braakman**  
Universiteit Utrecht

**Prof. dr. ir. W. Norde**  
Rijksuniversiteit Groningen/Wageningen Universiteit

**Prof. dr. E. van der Linden**  
Wageningen Universiteit

**Dr. H. Nieuwenhuijse**  
Friesland Foods, Deventer

Dit onderzoek is uitgevoerd binnen de onderzoeksschool VLAG.

# Fundamentals of unfolding, refolding and aggregation of food proteins

Kerensa Broersen

**Proefschrift**  
ter verkrijging van de graad van doctor  
op gezag van de rector magnificus  
van Wageningen Universiteit,  
prof. dr. M.J. Kropff,  
in het openbaar te verdedigen  
op vrijdag 23 september 2005  
des namiddags om half twee in de Aula

Broersen, K.

Fundamentals of unfolding, refolding and aggregation of food proteins

Thesis Wageningen University, The Netherlands 2005 – with summary in Dutch

ISBN: 90-8504-250-x

Broersen K

Fundamentals of unfolding, refolding and aggregation of food proteins

PhD thesis

Wageningen University, The Netherlands, 2005

Key words

Food, proteins, unfolding, refolding, aggregation, chemical engineering, glycosylation, electrostatics, sulfhydryl groups, structure, stability, thermodynamics, kinetics

Protein functionality in food products and in the body strongly relies on the fact that proteins can adopt a number of different conformations. These conformations include folded and (partially) unfolded conformations and many more subtle structural variants in-between. The various conformations are separated by kinetic barriers and, as such, require an activation energy to traverse from the one thermodynamically stable state into another. Also at a molecular level these thermodynamically different conformations are structured in different ways. It was found that even very subtle differences in conformation can have dramatic consequences for the aggregation propensity of proteins. For example, thermodynamic and/or kinetic intermediately trapped protein structures have been found to be sometimes highly prone to intermolecular interactions, primarily through the partial exposure of functional groups. This led us to focus in this thesis on the process from folded to aggregation prone state as opposed to the aggregation process itself. The central aim of this thesis is to explore structural features of proteins of importance to the generation of aggregation prone protein molecules. The relevance of investigating structural in combination with thermodynamic and kinetic parameters in the context of aggregation may be clear. The approach selected involves chemical engineering in which amino groups or carboxyl groups are converted into a chemical group with different properties. The chemical modifications have been specifically selected for their previously reported effects on protein aggregation, and include glycosylation, introduction of sulfhydryl groups and charge modification. Subsequently, this led to a detailed description of the structural impact of the modifications on thermodynamic parameters related to aggregate formation. It was found that the various modifications applied interact with the aggregation process in a rather diverse manner. The accumulation of data from this work in combination with results from literature was used to significantly improve the understanding of factors relevant to aggregation and to develop a general hypothesis for aggregation propensity, the AGPRO model. This model can be used within the food and pharmaceutical industry to determine the aggregation propensity of proteins used in formulae and medication.

*In het wijde van deze wereld  
duw ik almaar het hoge gras opzij  
op zoek naar het antwoord*

# CONTENTS

## ABSTRACT

CHAPTER 1	Introduction	9
-----------	--------------	---

### **Part I Protein glycosylation**

CHAPTER 2	Glycoforms of $\beta$ -lactoglobulin with improved thermostability and preserved structural packing.	27
CHAPTER 3	Glycosylation affects protein unfolding/refolding mechanism and the formation of aggregation-prone intermediates.	47
CHAPTER 4	Glucosylation of $\beta$ -lactoglobulin lowers the heat capacity change of unfolding: a unique way to affect protein thermodynamics.	73
CHAPTER 5	The origins of the effect of glucosylation on the aggregation mechanism of $\beta$ -lactoglobulin.	91

### **Part II Electrostatics and protein stability**

CHAPTER 6	Electrostatics controls fibril formation – part I: Charge engineering of proteins affects unfolding.	117
CHAPTER 7	Electrostatics controls fibril formation – part II: Net charge can affect fibril morphology and gel properties of ovalbumin.	141

### **Part III Disulphide bonds and protein aggregation**

CHAPTER 8	Sulfhydryl groups do not affect heat-induced aggregation rate of ovalbumin but do affect aggregate morphology.	167
CHAPTER 9	Discussion: Driving forces for protein aggregation in perspective – the AGgregation PROensity (AGPRO) model.	183

SUMMARY/SAMENVATTING		219
----------------------	--	-----





# CHAPTER 1

## INTRODUCTION

### **Introduction**

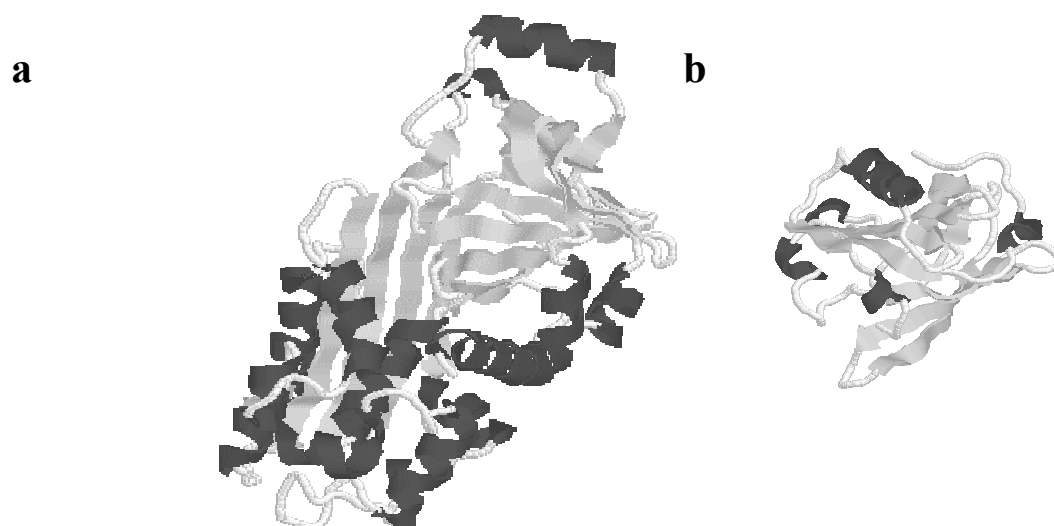
In protein-rich food products, such as cheese and yoghurt, proteins do not only serve a role as nutrients but also as texturiser. Figure 1 shows a number of naturally protein-rich food products such as nuts, meat, beans, fish, dairy products and eggs. The recognition of the versatility and value of proteins in food products led to the exploration of large-scale industrial isolation procedures of food proteins from protein-rich by-products from for example cheese production. Today these isolates are widely applied in food products to control texture. Next to a natural or additive component of food products, proteins are also an important constituent of our body and its number of appearances and functions, which have been defined today, seems beyond limits. The consumption of protein-rich food products, either originating from plants or from animals (Figure 1), provides us with the essential amino acids to build proteins required to fulfil roles in the body such as enzymes or hormones.



**Figure 1** Examples of protein-rich food products: eggs, nuts, yoghurt, meat, dairy desserts, kidney beans, cheese, tempeh and fish.

The large variety of naturally occurring proteins provides the food industry with an extensive choice for protein sources to apply to obtain the textural, sensory or nutritional properties as desired. However, the current lack of understanding of protein behaviour and the high cost to evaluate a large number of protein sources for their suitability in food products has greatly frustrated the development process of new protein-containing food products. As structure and function of the diverse proteins have been found to be strongly correlated in the past (Morley et al 2004, De Sanctis et al 2004, Coombe & Kett 2005, Deignan et al 2004, Zemser et al 1998) it seems reasonable to select protein systems based on their structural properties. However, considering only the primary sequence or molecular weight of a protein provides insufficient insight to design functionality in food products. Examples of primary structure properties are the ratio of hydrophobic to hydrophilic residues, the number of disulphide bonds and free sulfhydryl groups and positive and negative charge (Chiti et al 2002, 2003, Linding et al 2004, DuBay et al 2004, Fernandez-Escamilla 2004, Hayakawa & Nakai 1985).

Next to the primary sequence, the folding, comprised of interactions at a secondary and tertiary structural level, more importantly determines the actual extent of exposure of hydrophobic regions, disulphide bonds/sulfhydryl groups and apparent net charge (Chiti et al 2002). Below, the structures of ovalbumin, an egg white protein (Figure 2a) and  $\beta$ -lactoglobulin, a whey protein (Figure 2b) are depicted which will serve as an example to illustrate this.



**Figure 2** *Ribbon drawings of the tertiary and secondary structure of a) chicken egg ovalbumin (1OVA: Protein Data Bank) and b) bovine  $\beta$ -lactoglobulin (1BEB: Protein Data Bank) as solved by X-ray crystallography. The black structures represent helices and the dark grey structures represent sheets and the light grey structures represent random coil and turns.*

Both proteins are important food ingredients but they cannot be simply exchanged without affecting the properties of the food product. These differences originate, apart from a significant difference in size, among others, also from significantly varying proportions of the different types of secondary structure elements. As Figure 2 shows, approximately 18% of the secondary structure of  $\beta$ -lactoglobulin attains a so-called  $\alpha$ -helix conformation whereas for ovalbumin the relative proportion of  $\alpha$ -helix is 30% (Tedford et al 1998). Moreover, ovalbumin and  $\beta$ -lactoglobulin have been found to expose significantly different thermodynamic behaviour (such as heat- or denaturant-resistance). These complexities already suggest why it has been difficult in the past to predict the functional properties of proteins in food products from a single property of the polypeptide chain. Furthermore, extensive knowledge on the relative significance of structural and thermodynamic properties on protein behaviour is not always available, limiting the choice and efficient use of proteins

as food ingredients and requiring time-consuming trial-and-error studies to determine the appropriate protein system for the application. The aim of this thesis is to obtain fundamental knowledge on the driving forces for functional behaviour of proteins in food products, with specific reference to the aggregation process. This aim also results in the complementary aim of improving the understanding of specific architectural features of proteins that affect the process leading to aggregate formation.

### **Driving forces for aggregation**

Proteins in solution are dynamic ensembles of conformations. Individual protein molecules can fold and (partially) unfold many times during their lifetimes. As the unfolded state is an energetically unfavourable situation, due to the exposure of hydrophobic regions, the molecule will normally strive to refold again or to self-associate (aggregate), to avoid the exposure of hydrophobic groups as much as possible.

Both *in vivo* as well as for example in food products polypeptide chains have been found to behave under similar principles which are driven by a range of structural properties. In this thesis, a number of these determinants have been investigated for their relative importance in the aggregation process through chemical modification. These are hydrophobic exposure, electrostatic forces, glycosylation and kinetic and thermodynamic stability. These parameters have been found to significantly affect the aggregation properties of proteins (see for electrostatic forces refs e.g. Lopez de la Paz et al 2002, Zurdo et al 2001, Nakamura 1978, Ma & Holme 1982, glycosylation e.g. Fenouillet & Jones 1995, Wang et al 1996, Meldgaard & Svendsen 1994, Bouma et al 2003, Choei et al 2004, and for kinetic and thermodynamic stability e.g. Heegaard et al 2005).

### **Hydrophobic exposure**

The propensity of a polypeptide chain to aggregate has been found in the past to vary, among other factors, predominantly with the hydrophobic exposure (McSwiney et al 1994, Arntfield et al 1991, Alting et al 2004, Chiti et al 2003). In the folded polypeptide, burial of hydrophobic residues in the protein core limits the opportunity to interact in an intermolecular fashion. Upon unfolding, hydrophobic interactions in the protein core are disrupted and gradually become exposed to the solvent (Kato et al 1990). This provokes a situation in which water molecules are organised into a so-called “iceberg cage” around the exposed hydrophobic regions automatically involving a high entropic penalty (Finney et al 1993, Tanford 1997). In the proximity of other (partially) unfolded polypeptides, this entropic cost

can be redeemed by the formation of intermolecular hydrophobic interactions resulting from aggregation. It has been found before that a frequently occurring modification like glycosylation possibly results in a limited hydrophobic exposure of protein molecules having important consequences for the aggregation propensity (Marshall & Rabinowitz 1976). This issue will be further discussed in the paragraph on glycosylation.

### **Disulphide bonds**

Most secreted and membrane proteins in both bacteria and eukaryotes contain intramolecular disulphide bonds. Polypeptides containing disulphide bonds are usually first assembled in a linear form with cysteine residues in its sequence, then followed by the oxidation of the thiol groups to form disulphide bonds. These intramolecular (i.e. within a single polypeptide chain) can be important in stabilising protein structures against unfolding through a positive contribution to the enthalpy in the folded state (Betz 1993). For example, the removal of a naturally occurring disulphide bond from hen egg lysozyme has been reported to decrease the thermal transition temperature with 25°C (Cooper et al 1992).

Upon unfolding, proteins may expose internal disulphide bonds or sulfhydryl groups which have been found to undergo disulphide/sulfhydryl exchange reactions with other unfolded polypeptide chains, resulting in the formation of a covalent, intermolecular disulphide bond. In the past, kinetic models have been proposed to explain the disulphide-sulfhydryl exchange reaction resulting in covalent disulphide linked aggregates and this specific type of aggregation has been referred to as polymerisation. It was suggested that the sulfhydryl group acts as a reducing agent to the intramolecular disulphide bond. Subsequently, the disulphide bond is reduced and the free sulfhydryl group undergoes disulphide bridging with one of the liberated sulfhydryl groups. This reaction proceeds in analogy with a radical-addition polymerisation reaction (Verheul et al 1998). These observations suggest that the presence or absence of sulfhydryl groups and disulphide bonds may be a tool to control aggregation reactions. Clear evidence on the relative contribution of this mechanism to the formation of aggregates has, however, not been posed up to now and the contribution to the aggregation process is unclear (McSwiney et al 1994).

### **Electrostatic interactions**

Electrostatic interactions are of central importance for many biological processes and the functionality of proteins in food (e.g. Sims et al 2005). Electrostatic effects influence various aspects of macromolecular folding and conformational stability. Electrostatic effects also

determine the structural and functional properties of proteins, such as their three-dimensional fold, binding energies and association rates. The contribution of salt bridges and ion pairs to protein stability is a much debated issue in the literature today. Thermophilic and hyperthermophilic analogues of mesophilic proteins tend to have increased numbers of salt-bridges. However, mutational studies indicate that the relative contribution of salt bridges to stability is often small (Tanner et al 1996, Perutz & Raidt 1975, Perutz 1978, Dekker et al 1991). Apparently, in some cases charge plays a more important role to control stability issues than in others.

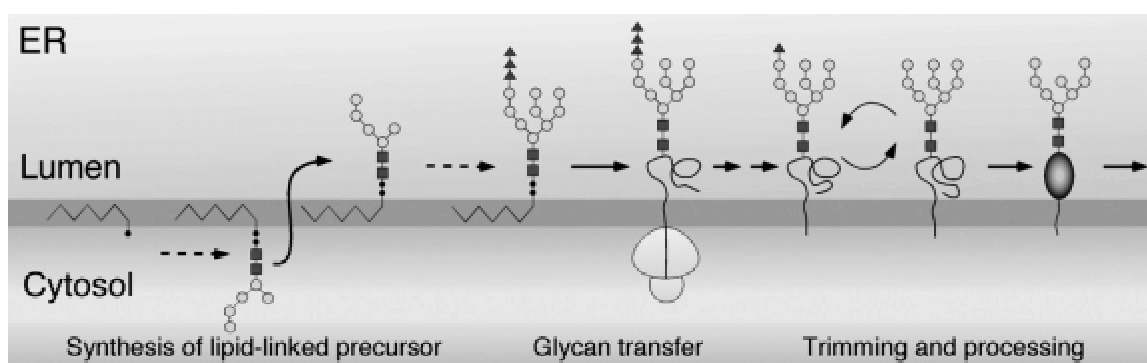
Next to the current discussion of electrostatic forces on protein stability, also the effect of electrostatics on aggregation has been an issue of discussion in literature. It has been reported before that the aggregation process is governed by a balance between attractive and repulsive forces, the latter consisting of electrostatic repulsion (Hatta et al 1986, Koseki et al 1989). At pH values far from the iso-electric point or at low ionic strength, electrostatic repulsive forces apparently hamper the formation of random aggregates and networks formed will mainly consist of fibrillar structures (Lefebvre et al 1998). Strong electrostatic repulsion appears to counteract aggregation due to the inability of (partly unfolded) protein molecules to approach close enough to enable stabilising interactions promoting the formation of aggregates. This suggests that net charge variation of a protein molecule can be used to control aggregation processes.

### **Glycosylation**

Protein glycosylation, the covalent linkage of a carbohydrate to the polypeptide chain, occurs widely in nature (*in vivo*) as well as under processing conditions and storage of food products (*in vitro*) (Lis & Sharon 1993, Dwek 1996). In food products, *in vitro* glycosylation by the Maillard reaction may affect the sensory properties of protein foods compared to non-glycosylated proteins through modifications at a molecular level such as hydrophobicity and solubility. This reaction may be desired, as in browning of meat upon baking, or undesired, as in the loss of nutritional value of proteins upon glycosylation of lysine residues. The Maillard reaction is a complex chain of reactions initiated by a condensation reaction between the  $\epsilon$ -amino group of lysine and the reducing group of a sugar to form Amadori or Heyn's rearrangement products via *N*-substituted glycosylamine. During the advanced stages the Amadori and Heyn's rearrangement products are degraded via a number of pathways (Mossine et al 1994, Röper et al 1983). The last stages of the Maillard reaction include

considerable protein cross-linking (Pellegrino et al 1999). Protein glycosylation in food products is frequently the result of the Maillard reaction and the kinetics and chemical mechanism of this reaction have been studied extensively and a review has been published by Wedzicha and Kaputo (1992).

An example of *in vivo* glycosylation includes heat shock proteins, where glycosylation infers stability (Venetianer et al 1994, Verma et al 1988). Glycosylation also plays a role in mediating a variety of protein-receptor interactions. For example, the envelope proteins of the human immunodeficiency virus type 1 are heavily glycosylated to mediate attachment of virions to the likewise glycosylated cell surface receptor molecules (Montefiori et al 1988). Protein glycosylation has been frequently reported to provide a means for protection against aggregation processes (Heal & McGivan 1998). Inhibition of self-association provides an important mode for the protection against aggregate-related diseases, such as Alzheimer's disease and Parkinson's disease. The exact origin of the aggregation inhibiting effect by glycosylation has not been identified up to now. Two means of biochemical pathways to link carbohydrates to polypeptide chains *in vivo* have been reported: First, *N*-glycosylation includes the attachment of the carbohydrate via an amide bond to the nitrogen of an asparagine side chain. Second, *O*-glycosylation involves the linkage of the carbohydrate to the hydroxyl group of a threonine or serine residue. *O*-glycosylation frequently occurs in the Golgi (e.g. Ercan & West 2005, Stornajuolo et al 2003) while *N*-glycosylation occurs co-translationally in the endoplasmic reticulum (illustrated for *N*-glycosylation in Figure 3) and have therefore been suggested to be involved in the folding process (Imperiali & Rickert 1995, Julenius et al 2005). These types of glycosylation also almost certainly have a profound effect on the structure and mobility of polypeptide chains as became evident from NMR experiments (Davis et al 1994, Andreotti & Kahne 1993).



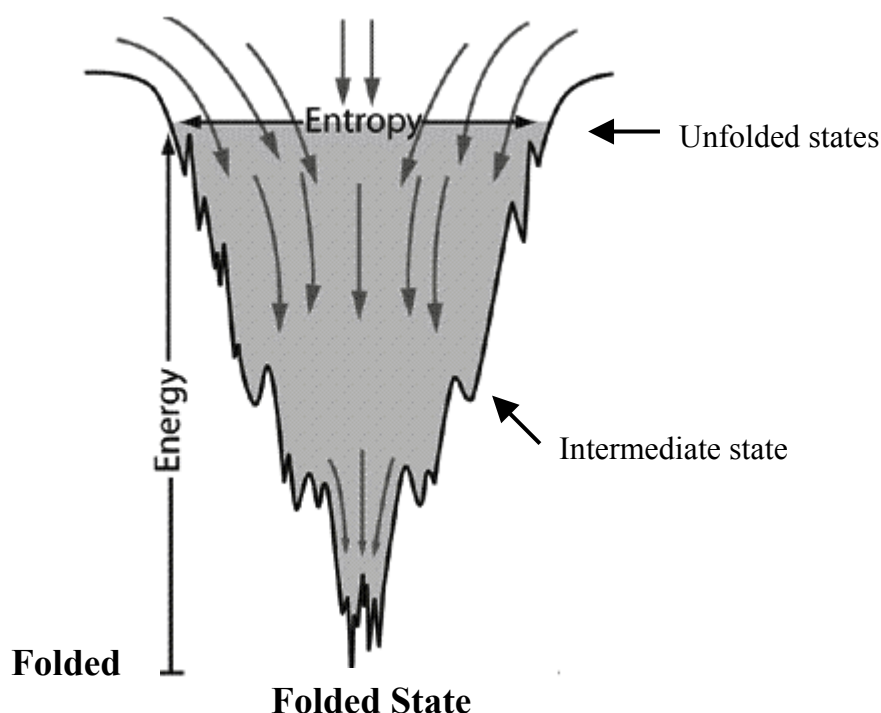
**Figure 3** Co-translational *N*-glycosylation in the endoplasmic reticulum supports correct protein folding (Helenius & Aebi 2001).

Advanced glycation end products (high molecular weight melanoidins) obtained through the Maillard reaction of readily folded proteins have been reported to result in a higher probability for protein aggregation (Kumar et al 2004). Glycosylation of proteins *in vivo*, prior to and during the process of folding, has been reported to inhibit the aggregation process (Heal & McGivan 1998). From this it seems that the various modes by which glycosylation of proteins can be accomplished, next to the various stages with which the glycosylation processes may interfere during folding, can have diverse consequences for the aggregation propensity of proteins. It has not been clear up to now which mechanisms can be accounted for explaining these observations and whether perhaps it is possible to control protein glycosylation such that the engineered proteins can be of use in the food technology industry.

### **Protein folding and stability**

The extent to which the parameters discussed, hydrophobic exposure, disulphide bonds, charge and glycosylation, are contributing to the aggregation process depends on the stability and unfolding process of a protein. For example, the high number of hydrophobic residues in the sequences of most polypeptides are largely buried in the core of the fully folded protein. Their presence renders partially folded or fully unfolded polypeptides highly prone to aggregation (Fink 1998). Since the 1970's it has been known that all the information concerning the three-dimensional fold of a protein is concealed in the primary sequence of the polypeptide chain (Anfinsen 1973). Using hydrophobicity as an example, upon folding, the polypeptide chain may encounter various thermodynamically or kinetically stable intermediate states where the hydrophobic exposure may not be substantially minimised. While residing in these states, the protein may thus be prone to aggregation processes through hydrophobic interactions (Onuchic et al 2000). The folded state, which is achieved at the end of this pathway, is the favoured entropic state (Figure 4) and the majority of the hydrophobic regions is buried in the core of the protein. This illustrates the importance of the folding process and stability of protein molecules in relation to the aggregation process.





**Figure 4** *Folding funnel of protein folding in which the folded state is achieved by a gradual decrease in entropic configurations (modified from Brown 2003).*

To further explore the association between protein folding and aggregation, the diagram in Figure 4 shows the folding funnel representation for protein folding. The roughness of the funnel reflects local free energy minima, which can transiently trap the protein in configurations that are not fully folded (Onuchic et al 2000). Late in the folding process most of the protein structure has already formed. The final structure formed is the fully folded state. The difference in energy content between the fully folded state (bottom-end of the funnel) and the fully unfolded state (top-end of the funnel) represents the thermodynamic stability. The thermodynamic stability dictates, according to Boltzmann's distribution law, the distribution of folded and (partially) unfolded molecules at a certain time. The rate of the folding process is not directly related to the thermodynamic stability. The rate at which the polypeptide chain collapses into a compact structure is reflected by the folding rate, also called kinetic stability, and is dictated by a variety of factors. Factors determining the rate of folding are investigated today and have been reported to vary from proximity of native contacts in the primary sequence or relative contact order (Kuznetsov 2004, Plaxco et al 1998, 2000, Cieplak 2004), internal friction – energetics of intrachain interactions including energy barriers to backbone

rotations and long-range interresidue interactions (Pabit et al 2004, Portman et al 2001, Kaya & Chan 2002, Qiu & Hagen 2004), rate of diffusional motion of an unfolded polypeptide chain through its solvent (Pabit et al 2004) and the existence of multiple intermediate states or smoothness of the energy landscape for folding (Baumketner 2003, Onuchic et al 1997). As aggregation is a concentration-dependent process, it may be clear that the rate of unfolding, or the rate of generation of particles prone to aggregation is very relevant to the aggregation process.

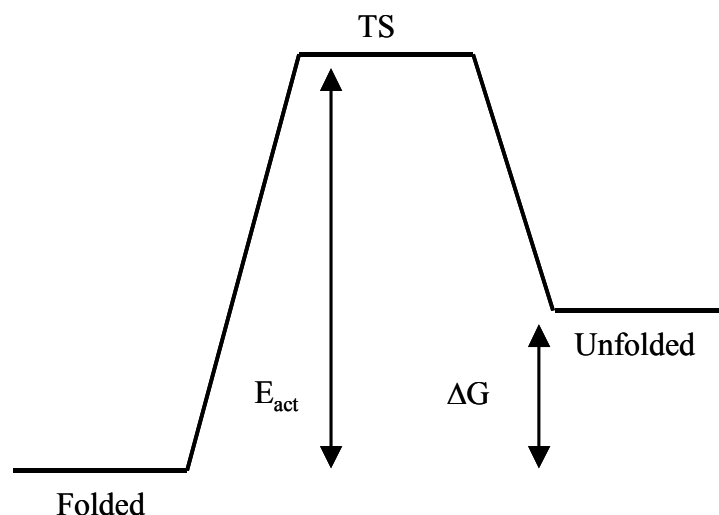
As discussed before, protein stability can also be discussed in terms of differences in Gibbs free energy,  $\Delta G$ , between the folded and unfolded state. This stability is often referred to as the thermodynamic stability and constitutes the equilibrium between folded and unfolded species at any given time under specific conditions denoted as follows.

$$\Delta G = G_u - G_f$$

where  $\Delta G$  is the difference in Gibbs free energy upon unfolding, and  $G_u$  and  $G_f$  are the Gibbs free energies of the unfolded and folded state, respectively. The two terms comprising the Gibbs free energy are the enthalpy and entropy. Forces that contribute to the enthalpy term are for example the strength of intramolecular hydrogen bonds and the interactions of the molecule with the surrounding solvent. The entropy term is mainly dictated by the hydrophobic effect of folding through an increase in disorder of water molecules upon protein folding. Enthalpy (H) and entropy (S) are related to the Gibbs free energy by the following equation.

$$G = H - T S$$

Where T is the temperature in Kelvin. The free energy difference of folding,  $\Delta G$ , is typically small, of the order of 5- 15 kcal/mol for a globular protein (compared to e.g. ~30 - 100 kcal/mol for a covalent bond). Destabilising factors include conformational entropy and the burial of peptide bonds and polar groups. Stabilising factors include disulphide bonds, burial of hydrophobic groups and the extent of hydrogen bonding. The thermodynamic and kinetic stability relate to each other as depicted in Figure 5.



**Figure 5** Kinetic ( $E_{act}$ ) and thermodynamic ( $\Delta G$ ) stability of proteins (TS=Transition State)

Both the kinetic stability and the thermodynamic stability will, in theory, have considerable implications to the aggregation propensity. For example, a protein with a high thermodynamic stability (i.e. large  $\Delta G$ ) will only expose functional groups crucial for aggregation upon extreme processing. These conditions may not be applicable to the food industry, or, otherwise never occur *in vivo*, such that aggregation by this protein molecule is unlikely to occur. The kinetic stability dictates the rate at which protein molecules unfold and is presented in Figure 5 as the Transition State (TS) which is the state with the maximum free energy in the pathway (Solís-Mendiola et al 1998).

As aggregation is a concentration dependent process, the rate at which aggregation prone-molecules are generated firmly linked with the aggregation rate. Protein stability has, for these reasons, often been appointed as a major factor in the aggregation process. Particularly in disease-related studies in which the stability of a polypeptide chain is affected by site-directed mutagenesis, increasing rates of aggregation are related to a decrease in stability. However, an explanation at a molecular level as to what extent kinetic and thermodynamic stability are genuinely associated to the rate and extent of aggregation has not been provided up to now.

### Approach

A range of approaches may offer the possibility to study forces of relevance to the aggregation propensity of proteins. In food chemistry many of the above mentioned factors

have been investigated by varying environmental conditions. This approach can offer an easily accessible way to modify the interactions involved in aggregate formation and to obtain a basic insight in the forces responsible for aggregation formation. However, it is also known that this approach may affect the aggregation mechanism in more ways and may not always allow an isolated comparison of the absence and presence of a certain effect on protein aggregation. Another approach widely applied to investigate structure/function relationships in protein science is site-directed mutagenesis. This approach involves mutation of one single or a range of specifically selected residues pre-translationally in order to investigate the specific effect or function of the affected residue in certain processes and the three-dimensional structure.

The approach selected in this work involves ‘random’ chemical engineering in which amino groups or carboxyl groups are converted into a chemical group with different properties. This approach offers the exquisite chance to obtain high quantities of material (up to hundreds of grams) and the possibility to apply the procedure to readily folded polypeptide chains such that the folding process and yield are not affected by the procedure.

The aggregation process can be probed in a number of ways. This thesis focuses on specific events during the unfolding process relevant for aggregation for a number of reasons. Preliminary literature investigations often point at the importance of the unfolding process itself to the aggregation through exposure of regions relevant to drive the aggregation. Second, thermodynamic and kinetic intermediately trapped protein structures have been found to be highly prone to intermolecular interactions, also through the partial exposure of functional groups. This knowledge led us to focus in this thesis on the process from folded to aggregation prone state as opposed to the aggregation process itself.

### **Outline of this thesis**

The central aim of this thesis is to explore structural features of proteins of importance to the generation of aggregation prone protein molecules. The approach applied throughout this thesis is that of chemical protein modification. The chemical modifications have been specifically selected for their previously reported effects on protein aggregation, and include glycosylation, introduction of sulfhydryl groups and charge modification. Accordingly, the variation of the properties of proteins that was achieved in this way, was expected to result in differences in aggregation properties. Subsequently, this led to a detailed description of the structural impact of the modifications on thermodynamic parameters related to aggregate formation.

Chapters 2 to 5 describe the effect of glycosylation on the folding/unfolding and aggregation of  $\beta$ -lactoglobulin. Chapter 2 includes a full description of the glycosylation procedure. It was found that glycosylation of proteins with glucose and fructose can be achieved in a non-destructive manner. Also, the degree of modification can be varied by varying the modification procedure and the formed glycoproteins gained thermostability. Chapter 3 describes the effect of glycosylation on the folding and unfolding process and how aggregation of a variety of proteins, e.g. lysozyme,  $\beta$ -lactoglobulin,  $\alpha$ -lactalbumin, myoglobin and cytochrome c is affected by modulating kinetic partitioning using glycosylation. As it was found that the aggregation process of proteins was largely affected by glycosylation without the elucidation of a significant parameter responsible for this observation, chapter 4 uses a thermodynamic approach to describe the effects of glycosylation on  $\beta$ -lactoglobulin A at a structural level. It was concluded that the effect of glycosylation on protein stability can be ascribed to an effect on the heat capacity change. In chapter 5 the molecular mechanism of the effect of glycosylation on aggregation is elucidated using  $\beta$ -lactoglobulin and it was found that particularly unfolding and refolding kinetics are strongly related to this effect. It was also found that hydrophobicity and net charge as such were not responsible for the observed effects.

Chapters 6 and 7 describe the effect of electrostatics on unfolding (chapter 6) and aggregation (chapter 7) of ovalbumin. It was found that, even though net charge strongly affects protein stability and also the unfolding rate, correction of the unfolding rate for the effect of electrostatics indicates no effect of electrostatics on the rate of aggregation, although the type of aggregate formed is significantly affected by the net charge.

Chapter 8 discusses the impact of sulfhydryl groups on the rate of heat-induced aggregation. The results obtained illustrated that, even though disulphide bonds play a significant role in stabilising the formed aggregates, non-covalent interactions predominantly determine the aggregation rate.

Chapter 9 analyses all obtained results from chapters 2 to 8 and summarises the obtained data into a general hypothesis for aggregation propensity. The presented model can be used as a model within the food industry and pharmaceutical industry to determine the aggregation propensity of proteins used in formulae and medication.

## References

Alting AC, Weijers M, de Hoog EH, van de Pijpekamp AM, Cohen Stuart MA, Hamer RJ, de Kruif CG, Visschers RW (2004) *J Agric Food Chem* 52: 623-631

- Andreotti AH, Kahne D (1993) *J Am Chem Soc* 115: 3352-3353
- Anfinsen CB (1973) *Science* 181: 223-230
- Arntfield SD, Murray ED, Ismond MAH (1991) *J Agric Food Chem* 39: 1378-1385
- Baumketner A, Jewett A, Shea JE (2003) *J Mol Biol* 332: 701-713
- Betz SF (1993) *Protein Sci* 2: 1551-1558
- Bouma B, Kroon-Batenburg LMJ, Wu Y-P, Brünjes B, Posthuma G, Kranenburg O, de Groot PG, Voest EE, Gebbink MFBG (2003) *J Biol Chem* 278: 41810-41819
- Brown TL (2003) In: *Making truth – the roles of metaphor in science*. University of Illinois Press pp 232
- Chiti F, Stefani M, Taddei N, Ramponi G, Dobson CM (2003) *Nature* 424: 805-808
- Chiti F, Taddei N, Baroni F, Capanni C, Stefani M, Ramponi G, Dobson CM (2002) *Nat Struct Biol* 9: 137-143
- Choei H, Sasaki N, Takeuchi M, Yoshida T, Ukai W, Yamagishi S, Kikuchi S, Saito T (2004) *Acta Neuropathol* 108: 189-193
- Cieplak M (2004) *Phys Rev E* 69: 031907
- Coombe DR, Kett WC (2005) *Cell Mol Life Sci* 62: 410-424
- Cooper A, Eyles SJ, Radford SE, Dobson CM (1992) *J Mol Biol* 225: 939-943
- Davis JT, Hirani S, Bartlett C, Reid BR (1994) *J Biol Chem* 269: 3331-3338
- Deignan ME, Prior M, Stuart LE, Comerford EJ, McMahon HEM (2004) *J Alzh Dis* 6: 283-289
- Dekker K, Yamagata H, Sakaguchi K, Udaka S (1991) *J Bacteriol* 173: 3078-3083
- De Sanctis D, Pesce A, Nardini M, Bolognesi M, Bodedi A, Ascenzi P (2004) *IUBMB Life* 56: 643-651
- DuBay KF, Pawar AP, Chiti F, Zurdo J, Dobson CM, Vendrusculo M (2004) *J Mol Biol* 341: 1317-1326
- Dwek RA (1996) *Chem Rev* 96: 683-720
- Ercan A, West CM (2005) *Glycobiology* 15: 489-500
- Fenouillet E, Jones IM (1995) *J Gen Virol* 76: 1509-1514
- Fernandez-Escamilla A-M, Rousseau F, Schymkowitz J, Serrano L (2004) *Nature Biotechnol* 22: 1302-1306
- Fink AL (1998) *Folding Des* 3: R9-R23
- Finney JL, Soper AK, Turner JZ (1993) *Pure Appl Chem* 65: 2521-2526
- Hatta H, Kitabatake N, Doi E (1986) *Agric Biol Chem* 50: 2083-2089
- Hayakawa S, Nakai S (1985) *Can Inst Food Sci Technol J* 18: 290-295
- Heal R, McGivan J (1998) *Biochem J* 329: 389-394
- Heegaard NHH, Jorgensen TJD, Rozlosnik N, Corlin DB, Pedersen JS, Tempesta AG, Roepstorff P, Bauer R, Nissen MH (2005) *Biochemistry* 44: 4397-4407
- Helenius A, Aebi M (2001) *Science* 291: 2364-2369
- Imperiali B, Rickert KW (1995) *Proc Natl Acad Sci USA* 92: 97-101
- Julenius K, Mølgaard A, Gupta R, Brunak S (2005) *Glycobiology* 15: 153-164
- Kato A, Ibrahim H, Takagi T, Kobayashi K (1990) *J Agric Food Chem* 38: 1868-1872
- Kaya H, Chan HS (2002) *J Mol Biol* 315: 899-909
- Koseki T, Kitabatake N, Doi E (1989) *Food Hydrocolloids* 3: 123-134
- Kumar MS, Reddy PY, Kumar PA, Surolia I, Reddy GB (2004) *Biochem J* 379: 273-282
- Kuznetsov IB, Rackovsky S (2004) *Proteins* 54: 333-341
- Lefebvre J, Renard D, Sanchez-Gimeno AC (1998) *Rheol Acta* 37: 346-357

- Linding R, Schymkowitz J, Rousseau F, Diella F, Serrano L (2004) *J Mol Biol* 342: 345-353
- Lis H, Sharon N (1993) *Eur J Biochem* 218: 1-27
- Lopez De La Paz M, Goldie K, Zurdo J, Lacroix E, Dobson CM, Hoenger A, Serrano L (2002) *Proc Natl Acad Sci USA* 99: 16052-16057
- Ma CY, Holme J (1982) *J Food Sci* 47: 1454-1459
- Marshall JJ, Rabinowitz ML (1976) *J Biol Chem* 251: 1081-1087
- McSwiney M, Singh H, Campanella OH (1994) *Food Hydrocolloids* 8: 441-453
- Meldgaard M, Svendsen I (1994) *Microbiology* 140: 159-166
- Montefiori DC, Robinson WE Jr, Mitchell WM (1988) *Proc Natl Acad Sci USA* 85: 9248-9250
- Morley S, Panagabko C, Stocker A, Atkinson J, Manor D (2004) *Vit E Health Ann NY Acad Sci* 1031: 332-333
- Mossine VV, Glinsky GV, Feather MS (1994) *Carbohydr Res* 262: 257-270
- Nakamura R, Sugiyama H, Sato Y (1978) *Agric Biol Chem* 42: 819-824
- Onuchic JN, Nymeyer H, Garcia AE, Chahine J, Socci ND (2000) *Adv Prot Chem* 53: 87-152
- Onuchic JN, Luthey-Schulten Z, Wolynes PG (1997) *Annu Rev Phys Chem* 48: 545-600
- Pabit SA, Roder H, Hagen SJ (2004) *Biochemistry* 43: 12532-12538
- Pellegrino L, van Boekel MAJS, Gruppen H, Resmini P, Pagani MA (1999) *Int Dairy J* 9: 255-260
- Perutz M, Raidt H (1975) *Nature* 255: 256-259
- Perutz M (1978) *Science* 201: 1185-1191
- Plaxco KW, Simons KT, Ruczinski I, Baker D. (2000) *Biochemistry* 37: 11177-11183
- Plaxco KW, Simons KT, Baker D (1998) *J Mol Biol* 277: 985-994
- Portman JJ, Takada S, Wolynes PG (2001) *J Chem Phys* 114: 5082-5096
- Qiu L, Hagen SJ (2004) *J Am Chem Soc* 126: 3398-3399
- Röper H, Röper S, Heyns K, Meyer B (1983) *Carbohydr Res* 116: 183-195
- Sims PA, Wong CF, Vuga D, McCammon JA, Sefton BM (2005) *J Comput Chem* 26: 668-681
- Solís-Mendiola S, Gutiérrez-González LH, Arroyo-Reyna A, Padilla-Zúñiga J, Rojo-Domínguez A, Hernández-Arana A (1998) *Biochim Biophys Acta* 1388: 363-372
- Stornajuolo M, Lotti LV, Borgese N, Torrisi MR, Mottola G, Martire G, Bonatti S (2003) *Mol Biol Cell* 14: 889-902
- Tanford C (1997) *Protein Sci* 6: 1358-1366
- Tanner J, Hecht R, Krause K (1996) *Biochemistry* 35: 2597-2609
- Tedford L-A, Kelly SM, Price NC, Schaschke CJ (1998) *Trans IChemE* 76: 80-86
- Venetianer A, Pirity M, Hever SA (1994) *Cell Biol Int* 18: 605-615
- Verheul M, Roefs SPFM, de Kruif CG (1998) *J Agric Food Chem* 46: 896-903
- Verma R, Iida H, Pardee AB (1988) *J Biol Chem* 263: 8569-8575
- Wang C, Eufemi M, Turano C, Giartosio A (1996) *Biochemistry* 35: 7299-7307
- Wedzicha BL, Kaputo MT (1992) *Food Chem* 43: 359-367
- Zemser M, Gorinstein S, Friedman M (1998) *Nahrung-Food* 42: 252-253
- Zurdo J, Guijarro JJ, Jiménez JL, Saibil HR, Dobson CM (2001) *J Mol Biol* 311: 325-340





## **PART I**

### **PROTEIN GLYCOSYLATION**



# CHAPTER 2

## GLYCOFORMS OF $\beta$ -LACTOGLOBULIN WITH IMPROVED THERMOSTABILITY AND PRESERVED STRUCTURAL PACKING

Broersen K, Voragen AGJ, Hamer RJ, de Jongh HHJ

Biotechnology and Bioengineering 2004, 86: 78-87

### **Abstract**

In this article we show how glycosylation can be used to control the thermal stability of proteins. The primary amines of  $\beta$ -lactoglobulin were glycosylated with glucose or fructose within a range of reaction parameters which did not affect the structure. The modified fractions were characterised, analysed for structural stability and hydrophobic exposure. The modification procedure gave rise to the production of glycoproteins with a well-defined Gaussian distribution, where glucose appeared more reactive than fructose. The integrity of the secondary, tertiary and quaternary structures remained unaffected by the modification procedure. However, upon heating the stability of the modified fractions increased up to 6°C. Here we demonstrate the effects on the thermodynamic properties of proteins by glycosylation; this work serves as a first step in understanding and controlling the process underlying aggregation of glycosylated proteins.

### **Keywords:**

$\beta$ -Lactoglobulin, glycosylation, monosaccharides, Maillard reaction, thermal stability

## Introduction

A common mechanism for proteins *in vivo* to cope with stress conditions is to increase the degree of glycosylation (Murakami et al 1997). Good examples for such glycosylation are found for many heat shock proteins (Venetianer et al 1994, Verma et al 1988). Another role of glycosylation *in vivo* is in mediating protein-receptor interactions (Montefiori et al 1988). Posttranslational glycosylation of, for example, lens proteins has been shown before to induce conformational changes and unfolding of proteins in the lens resulting in aggregation and tissue deposition, leading to cataract (Crabbe 1998). Also, deglycosylation of *N*-linked oligosaccharides has been reported to make recombinant human erythropoietin susceptible to aggregation on heat-treatment (Endo et al 1992). Prompt stress glycoproteins respond quickly with increased glycosylation during or within a few minutes of acute stress. For example, calreticulin is located within the endoplasmic reticulum of various cell types and is believed to protect cells by preventing protein aggregation (Heal & McGivan 1998). These results exemplify that carbohydrate chains are essential to the stability of proteins and can play an important role in the occurrence of misfolding- and aggregation-related diseases. Also, it has been reported from *in vitro* studies that glycosylation improves the conformational stability upon heating (Marshall & Rabinowitz 1976, Meldgaard & Svendsen 1994, Wang et al 1996). In view of the above-described roles, glycosylation would be an ideal tool to control protein-protein interactions both in natural environments as well as in technological applications.

In this study we focus on the first step in protein aggregation behaviour, namely the formation of unfolded proteins by heat treatment and the possibility to conjugate a defined amount of monosaccharides to a protein and to investigate the impact of such protein modification on structural integrity and stability. We chose  $\beta$ -lactoglobulin ( $\beta$ -LG) since this protein can be easily obtained in pure form and large quantities from bovine milk (Aschaffenburg & Drewry 1957, de Jongh et al 2001). The amino acid sequence of  $\beta$ -LG is reported as well as its three-dimensional structure (Braunitzer et al 1973, Godovac-Zimmermann & Braunitzer 1987, Sawyer et al 1985).  $\beta$ -LG occurs in different phenotypes and  $\beta$ -LG A differs from  $\beta$ -LG B at positions 64 and 118, where an Asp and a Val in the A variant are substituted by a Gly and an Ala in the B variant giving rise to differences in isoelectric point and molecular weight (Aschaffenburg & Drewry 1955, Braunitzer et al 1973). At ambient temperature and neutral pH in solution, the protein exists primarily as a non-covalent dimer (Creamer et al 1983).  $\beta$ -LG contains two disulphide bonds and a free

sulfhydryl group, which is inaccessible from the solvent at neutral pH and ambient temperature (Papiz et al 1986).

Upon heating,  $\beta$ -LG undergoes conformational changes, which may lead to irreversible protein-protein interactions mediated by hydrophobic and disulphide interactions. It has been recognised for years that the covalent attachment of sugars to  $\beta$ -LG increases the heat stability of the protein (Marshall & Rabinowitz et al 1976, Meldgaard & Svendsen 1994, Wang et al 1996). However, few attempts have been made to control the glycosylation process such that a more detailed description of the impact of the attachment of small sugar moieties to the protein stability could be established. Also, in the majority of the previously reported glycosylation studies the proteins had been incubated for an extended incubation time and/or high temperature such that protein polymerisation products, browning, odour compounds and advanced degradation Maillard products could be formed (e.g. Yeboah et al 2000). Moreover, often the glycosylation process conditions led to irreversible loss of the protein structural integrity, thereby already affecting the heat-induced assembly process.

The Maillard reaction has been frequently used in glycosylation studies and is a complex chain of reactions. During the early stage the  $\epsilon$ -amino group of lysine condenses with the reducing group of a sugar to form Amadori or Heyn's rearrangement products via *N*-substituted glycosylamine. Using the early stages of the Maillard reaction is a tool for chemical protein modification, revealing modified functional properties but retained structural integrity to study the structure-function relationship of proteins (Kato et al 1993, Morgan et al 1997, Nacka et al 1998). During the advanced stages the Amadori and Heyn's rearrangement products are degraded via a number of pathways (Mossine et al 1994, Röper et al 1983). The Maillard reaction between naturally occurring lactose in bovine milk and  $\beta$ -LG has been reported previously and occurs upon heat processing of milk (Morgan et al 1997). Lactose is a disaccharide which is known to react less efficiently with a protein than monosaccharide units since the Maillard reactivity of reducing sugars decreases with molecular weight of the sugar (Chevalier et al 2001). Therefore, the impact of the attachment of sugars to  $\beta$ -LG was studied using the monosaccharides fructose and glucose as model compounds. Fructose (ketohexose) and glucose (aldohexose) differ significantly in reducing activity, due to differences in mutarotation and chemical structure (Chevalier et al 2001).

Here we show that glycosylation can be controlled well *in vitro* by various reaction conditions; we provide insight as well into the glycosylation reaction itself and the impact on

the protein heat stability and hydrophobic exposure, two key factors in protein assembly reactions.

## Results

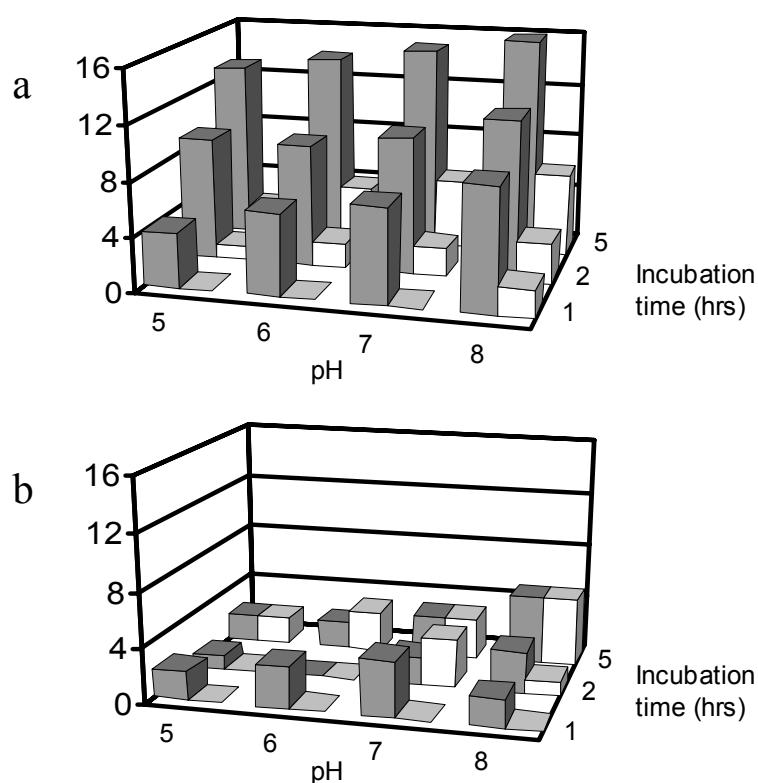
Covalently coupled sugar moieties to primary amino groups at the protein surface are expected to increase the surfacial hydrophilicity and may result in additional hydrogen bonding with the direct environment. These effects were hypothesised to improve protein stability and to affect protein-protein interactions. Our aim was to achieve such linkage under conditions that retain the folded state of the protein and do not lead to brown colouring or odour-formation as a result of advanced sugar chemical reactions. The efficiency of the conjugation of  $\beta$ -LG with fructose or glucose and at various incubation conditions can be studied by (at least) two different analysis methods. Figure 1 shows the effect of various modification conditions on the degree of glycosylation of  $\beta$ -LG. The number of reacted amino groups is determined by analysis of non-reacted amino groups (OPA-analysis) and reflects an ensemble-averaged number. As an alternative tool mass spectrometry can be used to analyse these samples, providing additional insight in the heterogeneity of the reaction (Figure 2). The degree of modification of all materials described here are comparable for these two methods. Some general results of the analysed materials can be given. Modification induces an increase in molecular mass of the protein with increments of a multiple of 162 Da. This number is equivalent to that expected upon covalent linkage of either a glucose or fructose moiety to an amino group. Secondly, modification leads to the formation of heterogeneous products with a Gaussian distribution in the degree of glycosylation (DG) with a width of 2-3 (low DG) up to 4-5 (high DG) moieties. Third, as exemplified in Figures 2b and d, the DG can exceed the number of available lysine residues on the protein. Finally, the mass spectra demonstrate that all protein used in this work was modified by processing, even under the mildest glycosylation conditions (fructose, pH 5.0, 45°C for 1 h, results not shown).

### Varying incubation conditions

*Glucose or fructose as glycosylating agent:* The average mass increase of glucosylated  $\beta$ -LG at pH 8.0 incubated for 5 h at 60°C (2590 Da equivalent to 16 glucose moieties) was significantly higher than that of fructosylated  $\beta$ -LG glycosylated under the same conditions (977 Da equivalent to six fructose moieties) (Figures 1a and 2d). Under most

of the different conditions tested, glucose appeared to react 3-4 times more efficiently to the protein than fructose.

*Effect of pH:* As illustrated in Figures 1 and 2b, increasing the pH from 5.0 to 8.0 results in an increase in molecular weight of the glucosylated and fructosylated  $\beta$ -LG. The modification of  $\beta$ -LG with sugar moieties at various pH values shows a linear increase of the degree of modification with increasing pH. While at pH 5.0 on average 13 glucose molecules were covalently attached to  $\beta$ -LG, at the most extensively modified sample, at pH 8.0, 16 glucose moieties were attached to  $\beta$ -LG (Mw = 20964). For fructosylated  $\beta$ -LG a similar trend could be observed (Figure 1).



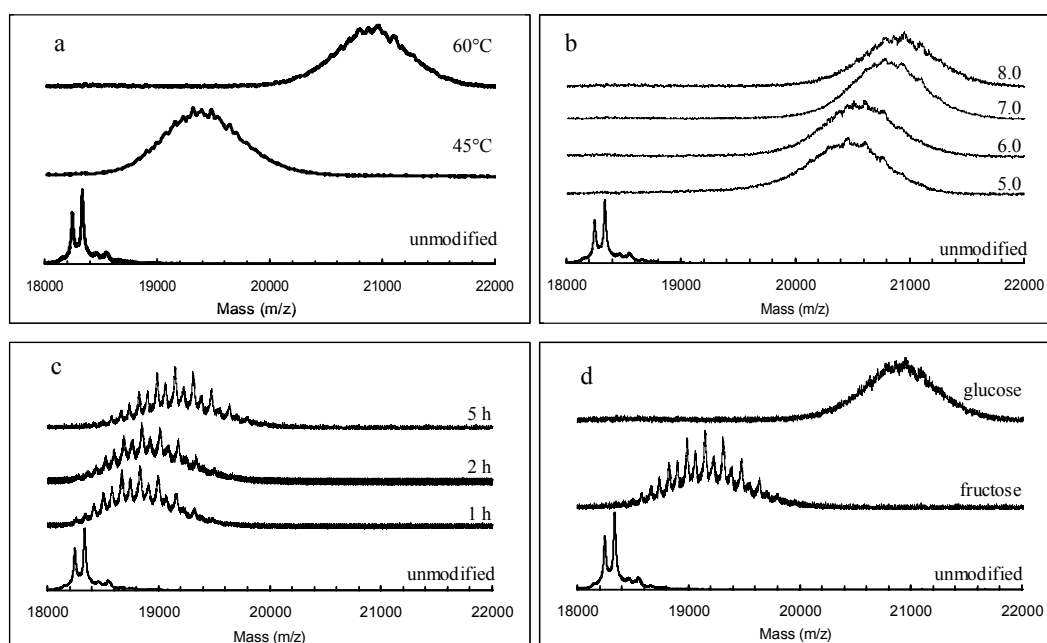
**Figure 1** Effect of various modification conditions on the degree of glycosylation (DG).  $\beta$ -LG was incubated with glucose (grey bars) or fructose (white bars) at a) 60 °C and at b) 45 °C, at pH 5.0-8.0 for 1, 2 or 5 h. The degree of modification was deduced by determining the number of free amino groups by the OPA assay.

*Effect of incubation temperature:* Figures 1 and 2a illustrate that fructosylation and glucosylation at 60°C was more effective than modification at 45°C. For example, incubation of glucosylated  $\beta$ -LG at pH 8.0 for 5 h at 45°C resulted in a mass shift of 837 Da indicating

the covalent attachment of five glucose moieties to  $\beta$ -LG. At 60°C the mass shift was 2590 Da which is equivalent to the attachment of, on average, 16 sugars to the protein (Figure 2a).

*Effect of incubation time:* In general, the reaction rate of the substitution was faster during the first 2 h of the incubation period than after 5 h (Figures 1 and 2c) in fructosylated and glucosylated  $\beta$ -LG. This reduced reaction rate can be attributed to the reduced availability of reactive groups after an initial reaction time. After 1 h, two fructose moieties were attached to  $\beta$ -LG. After 2 h three fructose moieties were bound while after 5 h the attachment of six fructose moieties had occurred.

Clearly, increasing incubation times, temperatures and pH values improve the conditions required for an efficient glycosylation reaction. Overall, it is clear that the combination of experimental conditions during the modification procedure provide a direct control on the final degree of glycosylation (DG) of the (glyco)protein. Also, analysis of browning products and Amadori compounds (results not shown) indicated that under all conditions used here, no advanced reaction products could be detected.



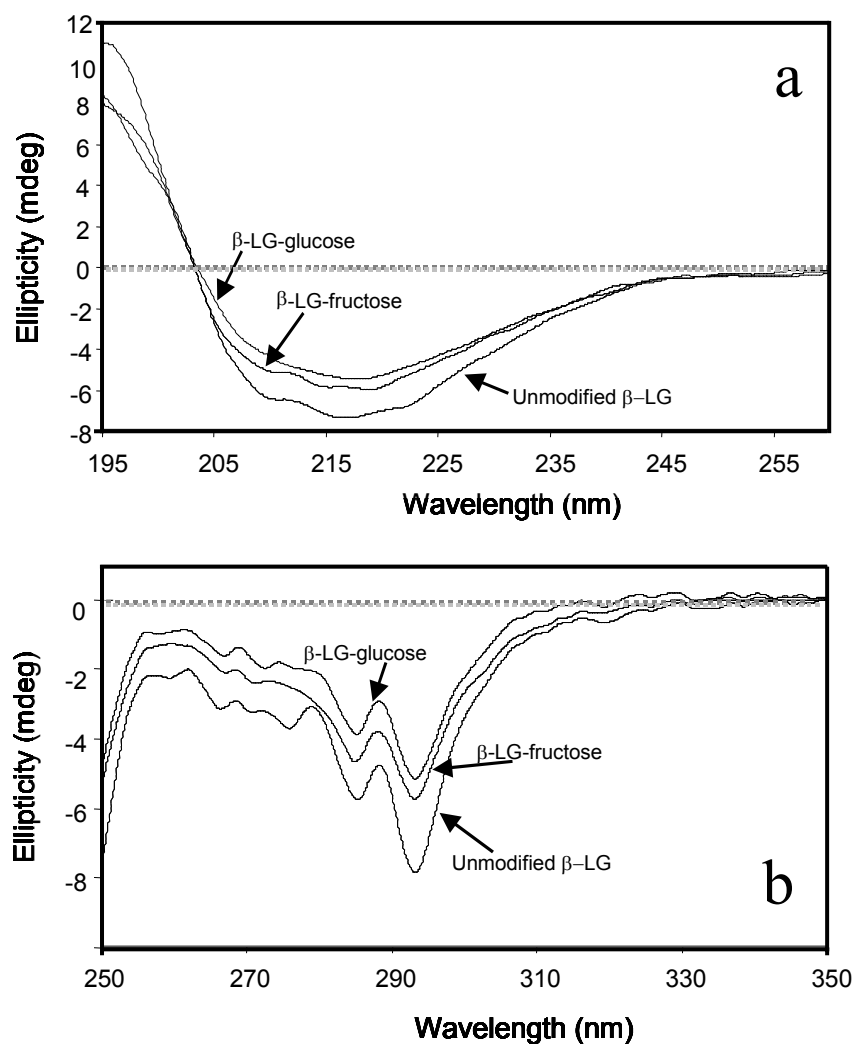
**Figure 2** Typical examples of mass ( $m/z$ )-spectra of unmodified and modified  $\beta$ -LG. The unmodified  $\beta$ -LG spectrum consists of two peaks (18281 and 18372 Da) reflecting the two phenotypes B and A, respectively, in a molar ratio of 40:60. a) effect of incubation temperature at pH 8.0, incubation with glucose for 5 h, b) effect of pH at 60 °C for incubation with glucose for 5 h, c) effect of incubation time at pH 8.0 and 60 °C for incubation with fructose, d) effect of sugar type at pH 8.0 and incubated at 60 °C for 5 h.



### Impact on structural integrity

*Secondary structure:* While modifying the primary amino groups of  $\beta$ -LG, it was aimed to retain the secondary, tertiary and quaternary structure of the unmodified protein. The effect of glycosylation on the secondary structure of  $\beta$ -LG was studied with far-UV CD. The far-UV CD spectrum (Figure 3a) of unmodified  $\beta$ -LG displays extremes around 195, 210 and 218 nm comparable to previously reported spectra (Hirota-Nakaoka & Goto 1999). Spectra were analysed using the method by de Jongh et al (1994) providing estimates of 11%  $\alpha$ -helix, 50%  $\beta$ -sheet, 10%  $\beta$ -turn and 29% random coil (Figure 3a) for the unmodified protein. These estimates are in agreement with literature values (de Jongh et al 2001). The spectra of the glycosylated samples with the highest DG of fructose and glucose are also shown in Figure 3a. The shapes of the spectra show only minor changes upon glycosylation. The spectra have the same zero-crossing at 203.5 nm and differ mainly in intensity. Spectral analysis demonstrated that in fructosylated  $\beta$ -LG approximately 5%  $\beta$ -sheet had been transformed into a random coil compared to unmodified  $\beta$ -LG. The estimated secondary structure of glucosylated  $\beta$ -LG did not differ significantly from the unmodified protein. For samples with a lower degree of glycosylation these differences were even smaller. The effect of the treatment itself on the secondary structure was negligible (results not shown). From these data it is evident that neutralisation of all positively charged lysine groups by conjugation with monosaccharide moieties does apparently not lead to loss of secondary structure.

*Tertiary structure:* Intrinsic fluorescence studies demonstrated that glycosylation did not affect the tertiary structure (results not shown). Modification of  $\beta$ -LG with glucose or fructose did not affect the shape of the emission spectra (i.e. peak position and intensities) for all tested samples compared to the unmodified protein. Hence, the local environment of the tryptophan residues of the glycosylated  $\beta$ -LG was considered to be the same as that in unmodified  $\beta$ -LG. Comparison of the fluorescence spectra intensities upon excitation at 274 nm (tyrosine) or 295 nm (tryptophan) also did not show any significant changes (results not shown). This illustrates that the energy transfer from tyrosine to tryptophan residues was not affected by the glycosylation, indicating that the intramolecular distances between tyrosine and tryptophan or spatial orientations were not affected. That indeed the tertiary structure was not significantly affected by the modification was also found by recording near-UV CD-spectra (Figure 3b). The shape of the spectra is comparable for the modified and unmodified protein, with for the latter a slightly larger intensity.

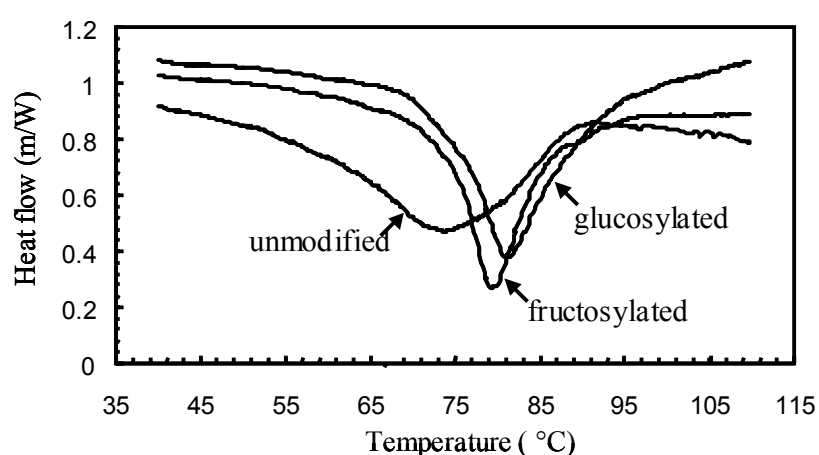


**Figure 3** a) Far-UV CD spectra of unmodified and modified  $\beta$ -LG, b) Near-UV CD spectra of unmodified and modified  $\beta$ -LG. Modified  $\beta$ -LG was incubated with glucose or fructose at 60 °C at pH 8.0 for 5 h.

**Quaternary structure:** Gel permeation chromatography (GPC) studies showed that the quaternary structure of  $\beta$ -LG remained dimeric upon glycosylation even for the highest degrees of glycosylation achieved in this study (results not shown). A minor contribution (< 5 %) was found at elution volumes indicative for aggregated protein fractions, but this was found to be similar to  $\beta$ -LG that was treated in the absence of sugars.

**Heat stability:** DSC measurements show that upon increasing temperature, the heat flow of the unmodified  $\beta$ -LG starts to decrease rapidly above approximately 47.0°C and reaches a minimum at 74.9°C (Figure 4). The thermal transition temperature of the

glucosylated  $\beta$ -LG conjugate (incubated at 60°C for 5 h at pH 8.0) as determined by DSC was 81.2°C ( $\pm 0.1^\circ\text{C}$ ) and that of fructosylated  $\beta$ -LG (incubated at 60°C for 5 h at pH 8.0) 79.4°C ( $\pm 0.1^\circ\text{C}$ ) which was higher than that of unmodified  $\beta$ -LG (74.9°C  $\pm 0.4^\circ\text{C}$ ). Also, the endothermic peaks obtained from the glycosylated proteins appear sharper with an onset at approximately 67°C compared to unmodified  $\beta$ -LG. The treatment alone also increased the thermal transition temperature of the protein slightly (results not shown). However, the attachment of sugar moieties to  $\beta$ -LG increased the thermal transition temperature significantly higher than the treatment alone.



**Figure 4** Differential scanning calorimetric spectra of (un)modified  $\beta$ -LG. Modified  $\beta$ -LG was incubated with glucose or fructose at 60°C, pH 8.0 for 5 h.

While DSC provides insight in the enthalpy changes occurring in the protein during unfolding, near-UV CD spectra give an indication about the tertiary structural integrity of the protein. Temperature traces of near-UV CD (Figure 5a) as a function of temperature show changes in intensity of the aromatic residues while heating the samples from 25°C to 90°C. Increasing the temperature results in a decreased intensity of the ellipticity at 293 nm. The relative ellipticity change is plotted in Figure 5a. The midpoint unfolding temperature of glucosylated and fructosylated  $\beta$ -LG obtained by near-UV CD was not significantly different from the unmodified protein. Additional temperature traces in the far-UV CD region (203.5 nm) provides information of the structural integrity at a secondary folding level (Figure 5b). Glycosylation of  $\beta$ -LG results in a shift of about 7°C toward higher temperatures of the transition from folded secondary structure to more unfolded structures. No difference in the

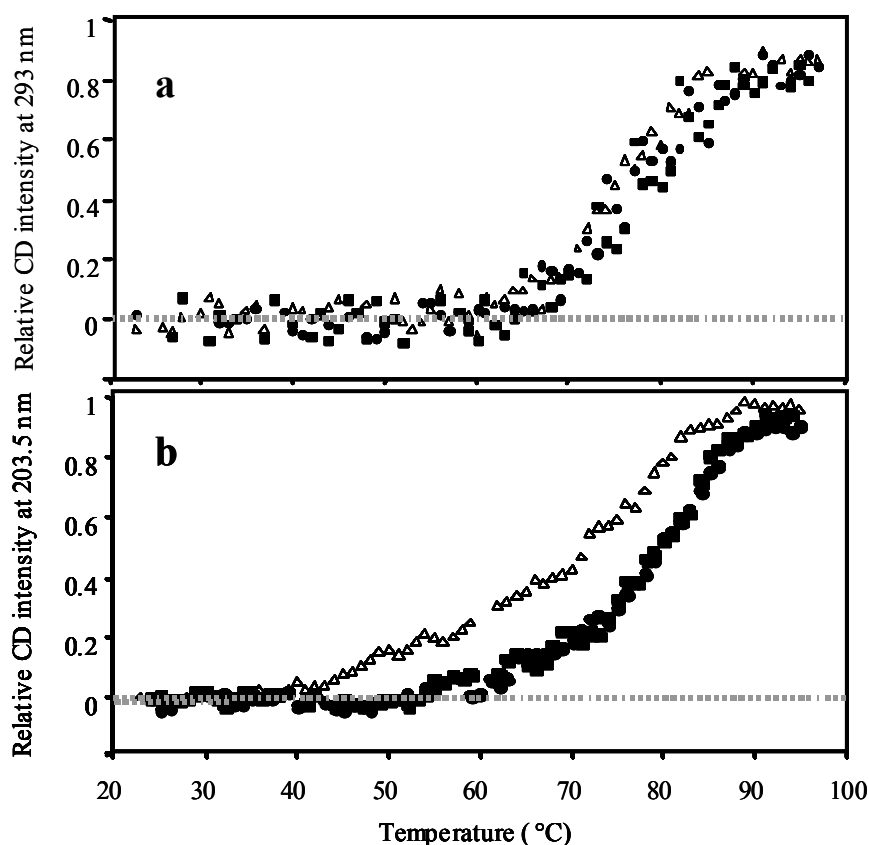
midpoint of the transition between glucosylated and fructosylated  $\beta$ -LG was found. The values obtained for the thermal transition temperatures from DSC and CD temperature measurements are summarised in Table I.

**Table I** *Chemical and thermodynamic properties of unmodified and modified  $\beta$ -LG. Modified  $\beta$ -LG was incubated with glucose or fructose at 60 °C at pH 8.0 for 5 h. The thermal transition temperature ( $T_t$ ) of DSC is the maximum heat flow. The thermal transition temperature from CD is the midpoint of the transition from folded to unfolded state.*

	DG <sup>a</sup>	ANS binding Mol ANS/Mol protein	T <sub>t</sub> (°C)		
			DSC	near-UV CD	far-UV CD
Unmodified	0	1.21	75	74	68
Glucosylated	16	0.66	81	74	77
Fructosylated	7	0.32	79	78	77

<sup>a</sup>DG = degree of glycosylation

*ANS fluorescence:* Upon non-covalent binding of the fluorescent probe ANS to hydrophobic patches on protein surfaces, its fluorescence quantum yield increases and the wavelength of maximal absorbance shifts. Monitoring ANS fluorescence upon titrating a protein with this probe can therefore be used to scale the protein's exposed hydrophobicity. The number of ANS molecules that could be bound to the surface of the modified proteins (fructosylated 0.66 Mol/Mol protein, glucosylated 0.32 Mol/Mol protein) was significantly lower than for unmodified  $\beta$ -LG (1.21 Mol/Mol protein) (Chapters 4 and 5 of this thesis describe a more detailed analysis of the impact of glucosylation on the hydrophobicity of  $\beta$ -LG).



**Figure 5** Temperature-dependence of near-UV (a) and far-UV (b) circular dichroism intensities of unmodified and modified  $\beta$ -LG. a) Near UV CD at 293 nm of unmodified and modified  $\beta$ -LG and b) Far UV CD at 203.5 nm of unmodified and modified  $\beta$ -LG. Modified  $\beta$ -LG was incubated with glucose or fructose at pH 8.0 for 5 h at 60 °C ( $\Delta$ ) unmodified  $\beta$ -LG, ( $\bullet$ )  $\beta$ -LG-glucose, ( $\blacksquare$ )  $\beta$ -LG-fructose). The ellipticity is plotted relative to the differences of the intensities at 20 °C and 90 °C.

## Discussion

As reported in literature, glycosylation of proteins may have an effect on protein aggregation properties (e.g. Marquardt & Helenius 1992). The Maillard-products of monosaccharides that reacted with multiple primary amino groups on  $\beta$ -LG, as described in this study, display a clear, slower aggregation kinetics compared to unmodified protein (e.g. Endo et al 1992, Heal & McGivan 1998). It can be questioned, however, whether differences in colloidal stability, due to altered protein-protein interactions and related to a different degree of neutralisation of positive charges, are the only cause. Also, the conformational stability might be affected. The aim of this work was to demonstrate that, by using a generic process, it is possible to obtain relatively defined materials, that bear potential grips to fine-tune protein-protein interactions, both in terms of exposed hydrophobicity and net surface

charge. Moreover, upon covalent attachment of monosaccharides to the primary amino groups of  $\beta$ -LG, the heat stability was improved while, as aimed, the structure of the unmodified protein was retained.

#### *Preparation of glycoforms of $\beta$ -LG*

Unmodified  $\beta$ -LG contains 16 potential sites (15 lysine residues and one *N*-terminal group) for coupling glucose or fructose molecules via Maillard reactions. In some cases (pH 8.0, 5 h, glucose) it was, however, found that more glucose moieties were attached than lysine groups were available (Figures 1 and 2). The attachment of more than 16 sugar moieties to  $\beta$ -LG implies that, next to the lysine residues, most likely arginine participates in the Maillard reaction as well. Since mass spectrometry does not make any distinction between different components with a similar molecular weight, it is uncertain whether the products obtained are true Amadori and Heyns products or whether these are actually degradation products obtained during advanced stages of the Maillard reaction. The degradation products can be formed by the 3-deoxyaldoketose route or the 1-deoxy-2,3-diketose route (Brands & van Boekel 2001). The intermediate products of these routes differ only by one H-atom and therefore they cannot be distinguished using mass spectrometry analysis. That the stability of formed products is limited was demonstrated by mass spectrometry of glycosylated  $\beta$ -LG stored for 6 months at -20°C in lyophilised form. This material displayed a clear lower degree of glycosylation compared to the freshly prepared product (unpublished results). It has been described that the presence of sulphite species inhibits the browning reaction as the reaction of 3,4-dideoxyhexulose-3-ene with sulphite ion to form 3,4-dideoxy-4-sulphohexosulose leads to irreversible binding of the sulphite species (Leong & Wedzicha 2000).

Our results revealed that glucose was 3 to 4 times more reactive than fructose moieties in the reaction with primary amino groups. Previously reported results on the reactivity of the different chemical sugar structures (Suárez et al 1995) indicated that the different steps of the Maillard reaction involving different sugar types occur at different reaction rates. For example, a publication by Aoki et al (1999) showed that the disappearance of amino groups is faster for ovalbumin reacting with glucose than for ovalbumin with glucuronic acid. However, polymerisation and melanoid formation appeared faster for glucuronic acid than for glucose. In literature, some discussion exists concerning Maillard reactivity related to the nature of sugar. These discrepancies can be explained twofold. First, various conditions have been applied to prepare Amadori products e.g. modification in solution (Brands & van Boekel

2001) or in lyophilised form (Morgan et al 1999). Second, various methods have been applied to determine the degree of glycosylation, like detection of browning at  $A_{340}$  nm (Suárez et al 1995), of acid hydrolysis followed by measurement of sugar released (Marsh et al 1977) or of the reacted or blocked lysine residues by furosine (Moreaux & Birlouez-Aragon 1997). Mass spectrometry has been used to study the lactosylation of  $\beta$ -LG in milk during processing and to locate glycosylation sites (Morgan et al 1997), to study the association behaviour of glycosylated proteins (Morgan et al 1999) and to identify glycosylated species of lysozyme formed during glycosylation with D-glucose and D-fructose (Yeboah et al 2000). From these previous studies, mass spectrometry appeared a valid method to obtain data on the degree of glycosylation of proteins.

There is an apparent linear relationship between the DG and the pH of the original incubation solution. Although it can be argued what the effective pH in the lyophilised material is, the result of incubation is highly reproducible. Under more alkaline conditions, the reactive acyclic conformation of glucose and fructose tends to dominate over the  $\alpha$ - and  $\beta$ -conformation (Burton & McWeeney 1963). Our findings are in general agreement with studies on bovine serum albumin conjugated with glucose by Lapolla et al (1993) who indicated that increasing pH resulted in lower degrees of glycosylation. An increasing pH results in decreased activation energy of the Maillard reaction causing a faster reaction rate (Lee et al 1984). The pKa of the lysine  $\epsilon$ -amino group is 10.5 (Lee et al 1984). With increasing pH, the reaction rate of the Maillard reaction increases because lysine residues must be unprotonated to function as satisfactory nucleophiles.

The time of incubation significantly affected the number of sugar groups covalently attached to the protein. During the first stages of the glycosylation reaction, after 1 and 2 h, more reactive groups were labelled with sugar moieties. However, with increasing incubation time (after 5 h) and therefore decreasing availability of reactive groups, the rate of glycosylation decreased. A positive correlation between time of incubation and shift of glycoforms toward higher glycoconjugates has been reported previously (Morgan et al 1997, Nacka et al 1998). The Maillard reaction itself also tends to slowly decrease the pH of the reaction medium due to the progressive formation of organic acids during advanced stages of the Maillard reaction (Lingnert et al 1979), lowering the reaction rate of the glycosylation.

Incubation temperature also appeared an important tool to control the degree of glycosylation. An incubation temperature of 60°C resulted in much higher numbers of reacted sugar moieties compared to 45°C in glucosylated as well as fructosylated  $\beta$ -LG. A clear

drawback in using higher incubation temperatures is that proteins will become progressively irreversibly unfolded, implying a loss of functionality.

The CD, fluorescence and GPC studies indicated that the mild glycosylation procedures as performed in this study had no effect on the structural packing at ambient temperature (Figure 3). Attached sugar molecules did not shield the area around the tryptophan residues as determined by fluorescence analysis. Although in some studies extensive glycosylation of proteins resulted in protein cross-linking and loss of protein solubility (Kochinsky et al 1997), our modification procedure did not imply polymerisation or protein cross-linking reactions. It is evident from these data that it is possible to obtain Maillard products with various degrees of glycosylation and with preserved structural integrity.

#### *Colloidal stability*

For the unfolded particle to participate in the aggregation process a number of properties have to be evaluated. Next to the colloidal stability reflecting attractive and repulsive interactions of the particles, the concentration and irreversibility of follow-up reactions have to be taken into account with respect to the colloidal stability (Chi et al 2003). The decreased exposed hydrophobicity by glycosylation of  $\beta$ -LG (Table I) in the unfolded state will certainly reduce non-covalent protein-protein interactions, slowing down aggregation kinetics. Another aspect determining the colloidal stability is the apparent protein net charge. In view of the DG of the materials it could be expected that the isoelectric point of the materials would lower significantly upon attachment of sugar moieties. However, determination of the isoelectric point showed that it was only affected marginally (unpublished results). We expect that this contribution is therefore a minor one in the aggregation process.

#### *Conformational stability*

According to differential scanning calorimetry glucosylated  $\beta$ -LG is more heat-stable than the unmodified  $\beta$ -LG by 4°C for the fructosylated and by 6°C for glucosylated  $\beta$ -LG (Figure 4, Table I). Several other studies showed that L-asparaginase (Marsh et al 1977), ovalbumin (Aoki et al 1999) and  $\beta$ -LG (Moreaux & Birlouez-Aragon 1997) were stabilised by glycosylation and that the heat stability of a glycoprotein, bacterial (1,3-1,4)- $\beta$ -glucanase secreted from yeast, was decreased by deglycation (Meldgaard & Svendsen 1994). In view of the broad and complex thermogram of unmodified protein (Figure 4) it is, however, not



unambiguous to establish a protein thermal transition temperature, since it coincides with aggregation phenomena. The far- and near-UV CD data provide, in our opinion, a much clearer picture of the unfolding process (Figure 5). Here it can be seen that the midpoint of unfolding of the secondary structure occurs in line with tertiary unfolding for fructosylated  $\beta$ -LG, while it occurs at slightly higher temperatures for the glucosylated example. This in contrast to unmodified protein, where the secondary structure starts to unfold at already much lower temperatures. Clearly, glycosylation affects the stability of the secondary rather than tertiary structure of the protein. The origin of the thermodynamic effect may be found in an altered intramolecular hydrogen-bonding network, but also in differences in bound water molecules.

Here we demonstrated that, by using a generic process, it is possible to obtain relatively defined materials, that bear potential grips to fine-tune protein-protein interactions, both in terms of colloidal and conformational stability. We showed that under the conditions applied the degree of glycosylation can be very well controlled by pH, incubation time and temperature. Also the nature of the sugar is of importance. Second, glycosylation itself does not affect the structural integrity of  $\beta$ -LG, but does have a significant impact on the thermostability of the structure. Non-site specific glycosylation is a perfect tool for proteins to adapt to externally altered conditions without the use of complex enzyme systems and a complex sugar moiety synthesis apparatus. The current line of investigation focuses on the impact of glycosylation on the conformational stability aspects, related to the first step in generating reactive particles that may associate according to colloidal interactions.

## Materials and Methods

### *Materials*

Ortho-phthaldialdehyde (OPA), 3,5-dimethoxy-4-hydroxycinnamic acid, D-(+)-glucose, NaCl, NaOH, L-leucine, NaNO<sub>2</sub>, acetonitrile, trifluoroacetic acid, acrylamide, 8-anilino-1-naphtalenesulfonic acid, Na<sub>2</sub>HPO<sub>4</sub>, NaH<sub>2</sub>PO<sub>4</sub>, D(-)-fructose, and 2-(dimethyl amino) ethanethiol hydrochloride were all of analytical grade and purchased from Sigma-Aldrich and Coomassie Brilliant Blue R-250 from Pharmacia. *N,N*-dimethyl-2-mercaptoethylammoniumchloride (DMA) and di-sodiumtetraboratodeca-hydrate (borax buffer) were obtained from Merck. Sodium dodecylsulphate (SDS) was from Serva. All chemicals were used without further purification.  $\beta$ -LG was isolated and purified (>98% purity) from

fresh cow milk (A:B ratio 60:40) using the protocol as described by de Jongh et al (2001). The material was freeze-dried and stored at  $-40^{\circ}\text{C}$  until further usage.

#### *Modification of $\beta$ -lactoglobulin*

Freeze-dried  $\beta$ -LG was solubilised in demineralised water (0.3 mM) and mixed with fructose or glucose solutions in demineralised water (0.03 M) (molar ratio of primary amino groups:monosaccharide = 1:6.5). The pH of the mixtures was adjusted to 5.0, 6.0, 7.0 and 8.0 using either NaOH or HCl. The solutions were subsequently frozen at  $-20^{\circ}\text{C}$  and lyophilised at 50 mbar at  $-50^{\circ}\text{C}$ . Next, the samples were incubated at  $45^{\circ}\text{C}$  or  $60^{\circ}\text{C}$  at a relative humidity of 65% (exposed to an equilibrated atmosphere containing a saturated  $\text{NaNO}_2$  solution) for 1, 2 or 5 h. Finally, the modified samples were extensively dialysed (molecular weight cut-off 12-14000 Da) at  $4^{\circ}\text{C}$  against distilled water, lyophilised and stored at  $-18^{\circ}\text{C}$  until use.

#### *Analytical methods for product characterisation*

*Chromogenic OPA assay:* The degree of glycosylation (DG) is defined as the mole sugar groups reacted to a mole protein. The DG was determined indirectly by a chromogenic assay described by Church et al (1983). This method is based on the specific reaction between ortho-phthaldialdehyde (OPA) and free primary amino groups in proteins in the presence of 2-(dimethyl amino) ethanethiol hydrochloride (DMA), resulting in alkyl-iso-indole derivatives that absorb at 340 nm. The OPA-reagent was prepared by dissolving 40 mg OPA in 1 mL methanol, followed by the addition of 25 mL 0.1 M borax buffer, 200 mg DMA and 5 mL 10% SDS. The volume was adjusted to 50 mL with demineralised water. A quartz cuvette containing 3 mL reagents and 15  $\mu\text{L}$  0.3 mM protein solution was measured for absorbance at 340 nm. The absorbance of the reagent itself was subtracted from the protein-containing sample. A calibration curve was obtained by adding aliquots of a 2 mM L-leucine (Sigma) solution in water to 3 mL of OPA-reagent, yielding final concentrations in the range from 6.6 to 9.5  $\mu\text{M}$ . All measurements were performed in triplicate.

*Matrix Assisted Laser Desorption/Ionisation Time-of-Flight Mass Spectrometry (MALDI-ToF-MS):* MALDI-TOF mass spectra were acquired in a linear mode on a Voyager-DE<sup>TM</sup> RP (PerSeptive Biosystems Inc., USA), controlled by Voyager RP Software. One  $\mu\text{L}$  of a 0.3 mM  $\beta$ -LG solution in demineralised water was mixed with 9  $\mu\text{L}$  of a 3,5-dimethoxy-4-hydroxycinnamic acid matrix (10 mg/ml of 3,5-dimethoxy—hydroxy cinnamic acid in acetonitrile/0.3% trifluoroacetic acid/demineralised water (3:1:6, v/v/v). Next, 2  $\mu\text{L}$  of the

final mixture was deposited on a 100-well golden plate and air-dried at room temperature to crystallise. Subsequently, the samples were analysed at an accelerating voltage of 25 kV and an initial velocity of 350 m/sec by taking 250 shots per spectrum. The mass range ( $m/z$ ) of peaks were determined from 2 to 25 kDa. Prior to each experiment the apparatus was calibrated using a reference protein solution. All samples were prepared and measured in duplicate.

*SDS-PAGE:* Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to a protocol described by Leammli (1970). A 15% (w/v) acrylamide separating gel and a 4% (w/v) acrylamide stacking gel containing 0.1% SDS (w/v) were run using a Mini-PROTEAN II Electrophoresis Cell (Biorad). Samples of 0.1 mM  $\beta$ -LG were prepared in sample buffer containing 10% SDS and 1.25%  $\beta$ -mercaptoethanol. Gels were stained with Coomassie Brilliant Blue R-250 and destained with 30% methanol/10% acetic acid in water.

*Isoelectric focusing:* The apparent isoelectric points (IEPs) of unmodified and modified  $\beta$ -LG fractions were determined using the Phast System (Pharmacia). Protein samples of 0.3 mM were prepared in demineralised water. Ready-to-use PhastGel IEF gels (4.5-6.5) were used. Gels were stained with Coomassie Brilliant Blue R-250 and destained in a 30% methanol/10% acetic acid in water. A calibration kit containing proteins with isoelectric points ranging from 2.5 to 6.5 (Pharmacia) was used.

*Advanced reaction products:* Amadori compounds (Brands & van Boekel 2001) and browning products (Brands et al 2002) were analysed using existing standard reference methods published previously.

### *Conformational Analysis*

*Gel Permeation Chromatography (GPC):* The quaternary structure of unmodified and modified  $\beta$ -LG in a 10 mM phosphate buffer (pH 7.0) containing 50 mM NaCl at 20°C was determined by GPC using a 24 mL Superdex 75 column HR 10/30 (Amersham) at a flow rate of 0.4 mL/min and detection of protein in the eluents at 280 nm.

*Circular Dichroism (CD) spectroscopy:* Far-UV CD spectra of 5.5  $\mu$ M of unmodified and modified  $\beta$ -LG in a 10 mM phosphate buffer (pH 7.0) were recorded at 20°C in the spectral range from 260 to 190 nm with a spectral resolution of 0.5 nm on a Jasco J-715 spectropolarimeter (Jasco Corp Japan). Spectra were recorded as averages of eight spectra. The scanning speed was 100 nm/min and the response time was 0.125 s with a bandwidth of

1.0 nm. An instrument sensitivity of 20 mdeg was used. Quartz cuvettes with an optical path of 1 mm were used. The spectra were corrected for the corresponding protein-free sample. The resulting spectra were analysed for the secondary structure estimates using a non-linear Least Squares fitting procedure using reference spectra for proteins as described by de Jongh et al (1994).

*Tryptophan fluorescence:* Emission and excitation fluorescence spectra of unmodified and modified  $\beta$ -LG solutions of 2.7  $\mu$ M in a phosphate buffer (10 mM, pH 7.0), were analysed using a Perkin Elmer (LS 50 B) luminescence spectrometer at 20°C. Quartz cuvettes with an optical path of 1 cm were used. The emission spectra from 300 to 400 nm were determined using excitation at 295 nm. Excitation and emission slit widths were 5 nm and a scan speed of 120 nm/min was used. The excitation spectra were measured from 250-335 nm at an emission wavelength of 338 nm which corresponds to the peak emission of tryptophan fluorescence. All spectra were corrected for the corresponding protein-free sample. The measurements were performed in duplicate.

#### *Conformational stability*

*Circular dichroism spectroscopy:* Near-UV and far-UV CD spectra of the ellipticity of unmodified and modified  $\beta$ -LG in a 10 mM phosphate buffer (pH 7.0) (0.05 mM and 5.5  $\mu$ M, resp.) were recorded at a temperature range of 25 to 90°C with a heating rate of 12 °C/h while recording CD spectra at intervals of 0.1°C at 293 nm and 203.5 nm respectively. The response time was 4 s with a bandwidth of 1 nm. Quartz cuvettes with an optical path of 10 mm and 1 mm, respectively, were used. All spectra and temperature traces were corrected for the corresponding protein-free sample.

*Differential scanning calorimetry (DSC):* Unmodified and modified  $\beta$ -LG were dissolved in a 10 mM phosphate buffer (pH 7.0) at a protein concentration of 2.73 mM. An 800  $\mu$ L aliquot was sealed in a DSC cell (Setaram MicroDSC III calorimeter equipped with a Setaram controller CS32 and a Julabo F32 thermostat). A 10 mM phosphate buffer (pH 7.0) was used as reference sample. The heat flow was recorded in separate duplicates from 25 to 110°C at a heating rate of 1°C /min and a cooling rate of 3°C /min. A heating rate of 1°C /min was found to be sufficient to match the equilibrium condition of the denaturing protein (de Jongh et al 2001). The thermal transition temperature ( $T_i$ ) was defined as the temperature of maximum heat flow.

### *Hydrophobic Exposure*

**ANS Fluorescence:** Fluorescence of ANS (8-anilino-1-naphtalensulfonic acid; Sigma) was measured at 293 K using a Shimadzu (UV-Vis 1240) spectrophotometer. 0.722 mg/mL ANS was dissolved in 10 mM phosphate buffer (pH 7.0). A protein solution of 0.2 mg/mL in 10 mM phosphate buffer (pH 7.0) was prepared. The fluorescence of 0.5 mL of this solution was measured in a 1 mL quartz cuvette.  $\lambda_{\text{ex}}=385$  nm;  $\lambda_{\text{em}}=400\text{-}650$  nm (5-nm slit). After the spectrum was measured, 10  $\mu\text{L}$  ANS solution was added to the protein solution, equilibrated for 2 minutes and measured for fluorescence. An equilibrium time of 2 minutes was sufficient to allow maximum binding of ANS to the protein as was indicated by Pots et al (1998). The addition of 10  $\mu\text{L}$  ANS solution was repeated until the fluorescence intensity of the bound ANS to the protein became a constant value. The concentration ANS bound to  $\beta$ -LG was determined using the extinction coefficient of ANS in methanol,  $8.0 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$  (Demarest et al 1998).

### **Acknowledgements**

The authors thank Jolan de Groot for technical assistance with the purification of the  $\beta$ -LG, Hans Kusters for assistance with analytical techniques and Carline Brands for support with the analysis of advanced reaction products. We also thank Henk Schols and Tiny van Boekel for useful discussions.

### **References**

- Aoki T, Hiidome Y, Kitahata K, Sugimoto Y, Ibrahim HR, Kato Y (1999) Food Res Int 32: 129-133
- Aschaffenburg R, Drewry J (1957) Biochem J 65: 273-277
- Aschaffenburg R, Drewry J (1955) Nature 176: 218-219
- Brands CMJ, Wedzicha BL, van Boekel MAJS (2002) J Agric Food Chem 50: 1178-1183
- Brands CMJ, van Boekel MAJS (2001) J Agric Food Chem 49: 4667-4675
- Braunitzer G, Chen R, Schrank B, Stangl A (1973) Hoppe-Seyler's Z Physiol Chem 354: 867-878
- Burton HS, McWeeney DJ (1963) Nature 197: 266-268
- Chevalier F, Chobert J-M, Popineau Y, Nicolas MG, Heartlé T (2001) Int Dairy J 11: 145-152
- Chi EY, Krishnan S, Kendrick BS, Chang B, Carpenter JF, Randolph TW (2003) Protein Sci 12: 903-913
- Church FC, Swaisgood HE, Porter DH, Catignani GL (1983) J Dairy Sci 66: 1219-1227
- Crabbe MJ (1998) Cell Mol Biol (Noisy-le-grand) 44: 1047-1050
- Creamer LK, Parry DAD, Malcolm GN (1983) Arch Biochem Biophys 227: 98-105
- de Jongh HHJ, Gröneveld T, de Groot J (2001) J Dairy Sci 84: 562-571
- de Jongh HHJ, Goormaghtigh E, Killian JA (1994) Biochemistry 33: 14521-14528
- Demarest SJ, Fairman R, Raleigh DP (1998) J Mol Biol 283: 279-291

- Endo Y, Nagai H, Watanabe Y, Ochi K, Takagi T (1992) *J Biochem (Tokyo)* 112: 700-706
- Godovac-Zimmermann J, Braunitzer G (1987) *Milchwissenschaft* 42: 294-297
- Heal R, McGivan J (1998) *Biochem J* 329: 389-394
- Hirota-Nakaoka N, Goto Y (1999) *Bioorg Med Chem* 7: 67-73
- Kato Y, Matsuda T, Nakamura R (1993) *Biosci Biotech Bioch* 57: 1-5
- Koschinsky T, He CJ, Mitsuhashi T, Bucala R, Liu C, Buenting C, Heitmann K, Vlassara H (1997) *Proc Natl Acad Sci USA* 94: 6474-6479
- Lapolla A, Gerhardinger C, Baldo L, Fedele D, Keane A, Seraglia R, Catinella S, Traldi P (1993) *Biochim Biophys Acta* 1225: 33-38
- Leammli UK (1970) *Nature* 227: 680-685
- Lee CM, Sherr B, Koh Y-N (1984) *J Agric Food Chem* 32: 379-282
- Leong LP, Wedzicha BL (2000) *Food Chem* 68:21-28
- Lingnert H (1979) SIK Rapport no. 458
- Marquardt T, Helenius A (1992) *J Cell Biol* 117: 505-513
- Marsh JW, Denis J, Jr Wriston JC (1977) *J Biol Chem* 252: 7678-7684
- Marshall JJ, Rabinowitz ML (1976) *J Biol Chem* 251: 1081-1087
- Meldgaard M, Svendsen I (1994) *Microbiology* 140: 159-166
- Montefiori DC, Robinson WE jr, Mitchell WM (1988) *Proc Natl Acad Sci USA* 85: 9248-9250
- Moreaux V, Birlouez-Aragon I (1997) *J Agric Food Chem* 45: 1905-1910
- Morgan F, Leonil J, Molle D, Bouhallab S (1999) *J Agric Food Chem* 47: 83-91
- Morgan F, Léonil J, Mollé D, Bouhallab S (1997) *Biochem Biophys Res Commun* 236: 413-417
- Mossine VV, Glinsky GV, Feather MS (1994) *Carbohydr Res* 262: 257-270
- Murakami MK, Nishikawa K, Hirakawa E, Murofushi H (1997) *J Biol Chem* 272: 486-489
- Nacka F, Chobert J-M, Burova T, Léonil J, Haertlé T (1998) *J Protein Chem* 17: 495-503
- Papiz MZ, Sawyer L, Eliopoulos EE, North ACT, Findlay JBC, Sivaprasadorao R, Jones TA, Newcomer ME, Kraulis PJ (1986) *Nature* 324: 383-385
- Pots AM, de Jongh HHJ, Gruppen H, Hessing M, Voragen AGJ (1998) *J Agric Food Chem* 46: 2546-2553
- Röper H, Röper S, Heyns K, Meyer B (1983) *Carbohydr Res* 116: 183-195
- Sawyer L, Papiz MZ, North ACT, Eliopoulos EE (1985) *Biochem Soc Trans* 13: 265-266
- Suárez G, Etlinger DJ, Maturana J, Weitman D (1995) *Arch Biochem Biophys* 321: 209-213
- Venetianer A, Purity M, Hever SA (1994) *Cell Biol Int* 18: 605-615
- Verma R, Iida H, Pardee AB (1988) *J Biol Chem* 263: 8569-8575
- Wang C, Eufemi M, Turano C, Giartosio A (1996) *Biochemistry* 35: 7299-7307
- Yeboah FK, Alli I, Yaylayan VA, Konishi Y, Stefanowicz P (2000) *J Agric Food Chem* 48: 2766-2774

# CHAPTER 3

## GLYCOSYLATION AFFECTS PROTEIN UNFOLDING/REFOLDING MECHANISM AND THE FORMATION OF AGGREGATION-PRONE INTERMEDIATES

Broersen K, Meinders MBJ, Hamer RJ, Voragen AGJ, de Jongh HHJ

Manuscript in preparation

### **Abstract**

This study shows that glycosylation of proteins significantly influences kinetic partitioning between folding and aggregation. The propensity to self-associate dramatically decreases upon glycosylation and the refolding efficiency is increased. The mechanism behind this effect of glycosylation has been investigated by comparing non-glucosylated and glucosylated  $\beta$ -lactoglobulin obtained by coupling glucose moieties to lysine residues. It was found that glucosylation changes the energy landscape for folding as was shown by a decrease in free energy ( $\Delta G$ ) upon unfolding and a lower activation energy both for unfolding and refolding. The glucosylation-induced kinetic and thermodynamic destabilisation is suggested to be a coupled effect of the interruption of the ion pair network on the surface of the protein as a result of glucosylation and altered binding of water-molecules around the glucosylated protein, as has been suggested from the altered heat capacity change upon unfolding found before (van Teeffelen et al 2005). The results indicate that decreased protein stability is not automatically related to a higher tendency to self-associate, as often presumed. The explanation behind the reduced self-association upon glucosylation can be twofold. First, the improved refolding efficiency indicates that glycosylation may direct the folding route by specific tagging of exposed regions of the folded protein consistent with the finding that glucosylation of proteins increases the refolding rate. Second, it was found that  $\beta$ -lactoglobulin requires only very limited unfolding to enable protein aggregation while upon glucosylation, protein molecules need to be unfolded to a higher degree. Models illustrating these hypotheses are presented.

### **Keywords:**

Globular proteins, glycosylation, folding/unfolding mechanism, protein stability, aggregation, kinetic partitioning

## Introduction

Correct protein folding is required to ensure a functional protein system *in vivo*. Considerable attention is currently being paid to factors that influence the unfolding/refolding mechanism of polypeptide chains (Dobson 2003, Otzen & Oliveberg 2004). The high number of hydrophobic residues in the sequences of most polypeptides are largely buried in the protein core in the fully folded form. However, their presence renders partially folded or unfolded polypeptides to molecules that are highly prone to aggregation (Fink 1998). Unwanted aggregation of incompletely folded, newly synthesised proteins is typically prevented through the interaction of molecular chaperones with glycans attached to the peptides (Land & Braakman 2001). Many examples in literature confirm that glycosylation may aid the (re)folding process. Fenouillet and Jones (1995) showed for example that the envelope glycoprotein gp160 requires its N-linked glycans for correct folding. Glycosylation occurs as soon as the growing polypeptide chain enters the endoplasmic reticulum and guides the protein through specific interactions with the chaperone systems to the folded state (Land & Braakman 2001, Helenius et al 1997, Hebert et al 1996). However, this pathway for *in vivo* folding has been proven to be insufficient to fully explain the effect of attached sugars on folding (Hurtley et al 1989). It has been suggested by Helenius (1994) that the attached sugar units must have functions in folding beyond their established role in the chaperone path. This suggestion was further supported by the finding that also *in vitro* (re)folding of glycoproteins is improved and misfolding and subsequent aggregation avoided by the presence of sugar units (f.e. Wang et al 1996). The aggregate-preventing effect of glycosylation has been related to protein stability parameters before: stabilisation of the globular fold generally disfavors the aggregation pathway by not allowing unfolded or partially folded states to be significantly populated (Hurle et al 1994). It has indeed been shown that covalently linked sugar moieties exert a large effect on the thermostability and resistance of proteins against unfolding. Consistent with this, Meldgaard and Svendsen (1994) showed that removal of N-glycans by glucanase resulted in a decrease in thermostability. However, that the preventive effect of glycosylation on protein aggregation is not unambiguous has been shown by recent studies which proposed a causal relationship between the occurrence of Alzheimer's disease or cataract and the accumulation of protein-advanced Maillard reaction end-product adducts, which involve heavily glycosylated proteins (Bouma et al 2003, Choei et al 2004). For example, cataract has been associated with conformational changes and unfolding of proteins in the eye lens, as a direct result of a reaction of glucose with eye lens proteins (Crabbe 1998). These examples show that the effect exerted by protein glycosylation on aggregation is more



complex than has been previously thought and may involve other aspects than those related to stability. Furthermore, Chi et al (2003) found that, although the intrinsic conformational stability of the folded protein plays an important role in its propensity to aggregate, the conformational stability and the resistance to aggregation are not necessarily linked. Clearly, the effects of glycosylation on the mechanism of protein aggregation are not clear today.

In this work we aim to elucidate the mechanism of the aggregation suppressing effect and the improved refolding efficiency upon protein glucosylation. To investigate the contribution of sugar chains in the (re)folding process of proteins by a random type of (post-translational) glycosylation, the Maillard reaction can be used. The Maillard reaction is brought about by a non-enzymatic reaction between reducing sugars and protein amino groups (Brownlee et al 1984, Monnier & Cerami 1981). This reaction can proceed from reversible Schiff bases to covalently bound Amadori rearrangement products (Brownlee et al 1984, Monnier & Cerami 1981).

It has been assumed that, in general, a relatively small protein with a molecular weight below 20 kDa can reversibly refold from the unfolded state (Ptitsyn 1987). However, for  $\beta$ -lactoglobulin, a globular protein with a molecular weight of 18.4 kDa as a monomer, refolding has been reported to be only poorly reversible (Subramaniam et al 1996, Yagi et al 2003). It can be hypothesised that glycosylation perhaps affects the reversibility of the unfolding process by driving the kinetic partitioning toward refolding. Within the context of  $\beta$ -lactoglobulin, also the reversible folding behaviour of other small globular proteins upon glycosylation such as myoglobin, lysozyme and  $\alpha$ -lactalbumin will be considered in order to evaluate the more general character of our findings. We will also evaluate the suitability of glucosylation of lysines as an approach to investigate the thermodynamic fundamentals of the preference of glycosylated proteins for refolding from a (partially) unfolded state over aggregation.

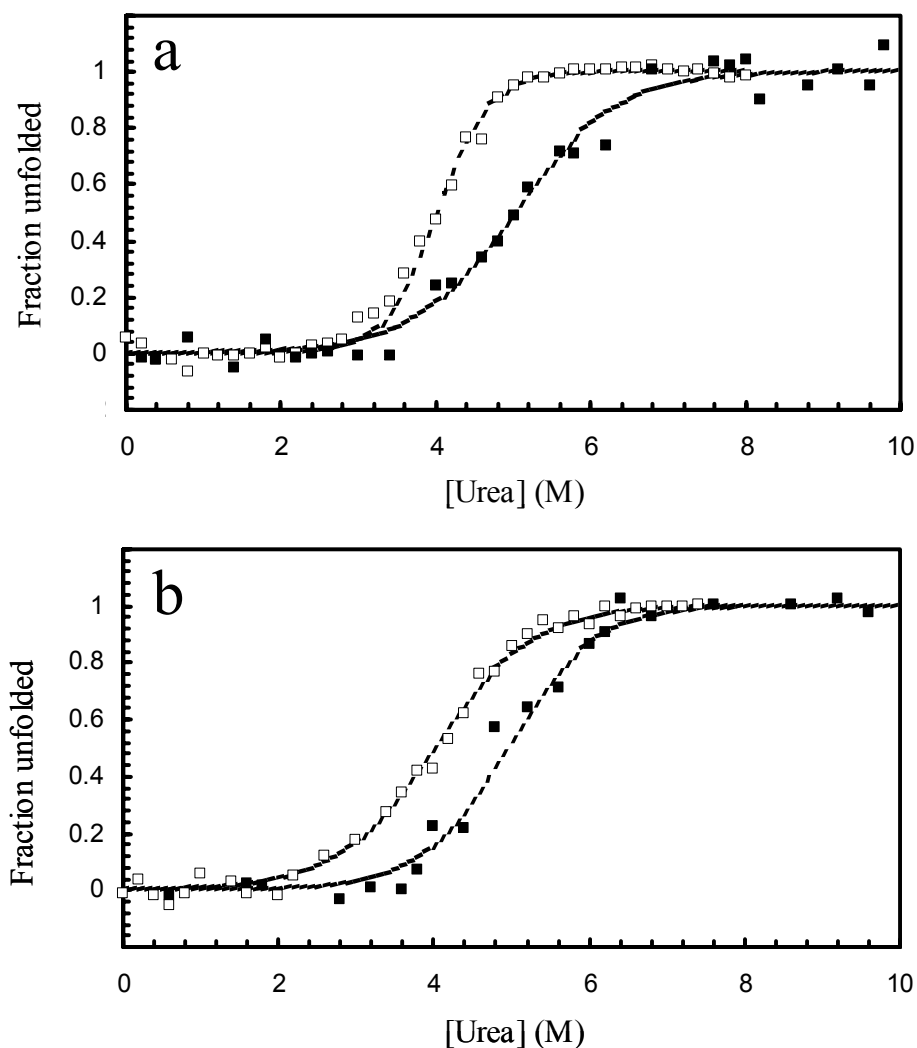
## Results

This paper describes the effect of protein glycosylation on kinetic partitioning between refolding and aggregation. Our strategy is based on determining the rates of aggregation and refolding upon the conjugation of glucose moieties to the small, thermodynamically and structurally well-characterised proteins  $\beta$ -lactoglobulin ( $\beta$ -LG), myoglobin (MYO), lysozyme (LYS) and  $\alpha$ -lactalbumin ( $\alpha$ -LAC) by the Maillard reaction. We first analysed the consequences of glucosylation in detail for  $\beta$ -LG investigating its conformational stability and

aggregation properties. The other proteins investigated will be considered at the last part of this work to verify the more generic applicability of the effect of glucosylation by the Maillard reaction on kinetic partitioning.

**Table I** *Physicochemical characteristics of non-glucosylated and glucosylated proteins: degree of modification and molecular weight*

	Number of attached glucose groups (deduced from MALDI tof-MS-data)		Mean molecular mass (Da) of main peak ( $\pm$ heterogeneity)	
	Non-glucosylated	Glucosylated	Non- glucosylated	Glucosylated
$\beta$ -LG	0	$16 \pm 2$	$18321 \pm 0$	$20884 \pm 2$
LYS	0	$5 \pm 1$	$14275 \pm 0$	$15087 \pm 2$
$\alpha$ -LAC	0	$11 \pm 2$	$14150 \pm 0$	$15932 \pm 2$
MYO	0	$15 \pm 2$	$16928 \pm 0$	$19344 \pm 2$



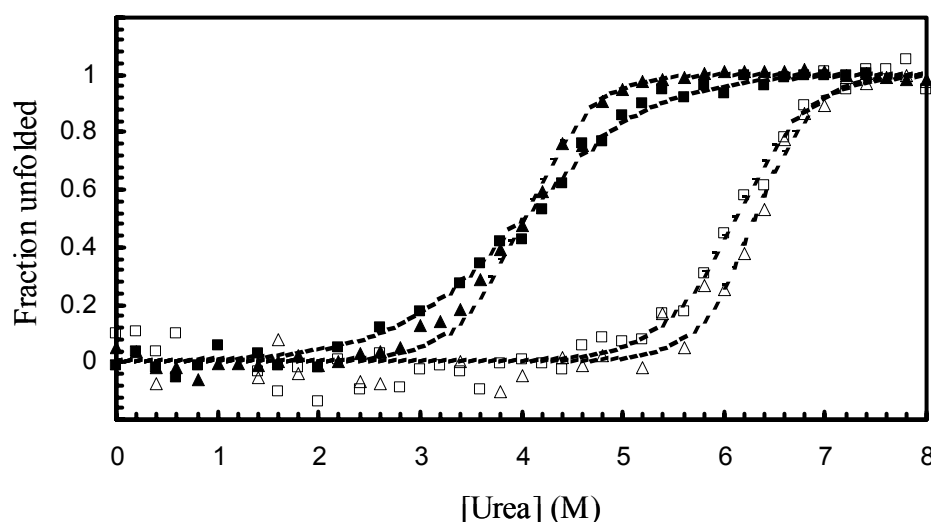
**Figure 1** Urea-induced unfolding of non-glucosylated a) and glucosylated b)  $\beta$ -LG using fluorescence and circular dichroism. (□) fluorescence signal at 320 nm, (■) CD-signal at 203.5 nm. The dashed line represents the fit based on the two state transition theory using the thermodynamic parameters in Table II and equations (1) and (2).

*The effect of sugars on the protein structure* – The degree of modification of the glucosylated variants of  $\beta$ -LG is presented in Table I. As can be observed in this table, 16 glucose moieties were attached to  $\beta$ -LG as determined by mass spectrometry (results not shown), thereby occupying all available locations for lysine-modification in the primary sequence of  $\beta$ -LG. Mass spectrometry also indicated that this random glucosylation procedure resulted in the production of a heterogeneous range of glycoproteins within a limited distribution of  $\pm 2$  (results not shown, see also Broersen et al 2004). As established with near- and far UV-CD and intrinsic tryptophan-fluorescence, the covalent linkage of sugars as well

as the modification procedure itself did not have any detectable impact on the secondary and tertiary structure of the selected proteins (results not shown). This result is in agreement with previous results (Broersen et al 2004). A conserved structural conformation allows a proper comparison of glucosylated and non-glucosylated proteins for the unfolding and refolding rates in addition to thermodynamic stability.

*The effect of sugars on  $\beta$ -LG unfolding thermodynamics* –An increasing fraction of the molecules becomes unfolded upon titration of the folded  $\beta$ -LG at pH 7.0 with urea as determined by fluorescence with a midpoint at 4.0 M urea (Figure 1a, open squares). For far-UV CD also a sigmoidal curve is obtained but the midpoint of unfolding is shifted to 5.2 M urea (Figure 1a, closed squares). It was also observed that the two unfolding curves found using far-UV CD and fluorescence do not coincide but that both curves individually can be well-described by a two-state unfolding transition from equations (1) and (2). Preliminary results indicated that under the conditions (pH 7.0, 10 mM phosphate buffer) and protein concentration (2.7  $\mu$ M) selected for the urea titration experiment, no aggregates were formed upon refolding and that the unfolding transition was fully reversible (unpublished results), allowing calculation of thermodynamic parameters from the results. Apparently, no thermodynamically stable intermediate was populated under the selected conditions, apart from a so-called ‘molten globule’ state between 3 and 7 M urea. This latter state is a well-known phenomenon that has been first described for  $\alpha$ -lactalbumin and is characterised by an affected tertiary structure but an intact secondary structure (Kuwajima et al 1975). Also the effect of the heterogeneous character of the glucose introduction has been verified. By comparing the urea-induced unfolding of  $\beta$ -LG with various degrees of glucosylation with a similar heterogeneity (similar distribution of masses, see Broersen et al 2004), we found that the  $m$ -value significantly shifted with decreasing degree of glucosylation (unpublished results). The effect of heterogeneity on the obtained results was therefore assumed to be negligible. The  $\Delta G$  calculated from the fluorescence results yielded a value of 30.4 kJ/mol and from far-UV CD a value of 19.2 kJ/mol. Figure 1b (open squares) shows that the midpoint of the unfolding transition is not affected by glucosylation when measured by fluorescence (Table II). Similarly, the unfolding curve determined from fluorescence could also be well-described by a two-state unfolding process and the process was found to be reversible (no aggregates were formed upon refolding). The calculated  $\Delta G$  however was significantly decreased from 30.4 kJ/mol for non-glucosylated  $\beta$ -LG down to 16.2 kJ/mol

upon glucosylation resulting from a shift of the  $m$ -value of the unfolding transition. However, determination of the unfolding transition of  $\beta$ -LG, at a secondary folding level using far-UV CD (closed squares), yielded only very small differences in  $\Delta G$  between glucosylated (22.8 kJ/mol) and non-glucosylated (19.2 kJ/mol)  $\beta$ -LG (Figures 1a and b). Similar to the observations of non-glucosylated  $\beta$ -LG, the unfolding transitions observed using fluorescence and far-UV CD do not coincide for glucosylated  $\beta$ -LG also suggesting the existence of an intermediate state. For glucosylated  $\beta$ -LG, shown in Figure 1b, this state populates between 2 and 6M urea. Based on these data we suggest that glucosylation affects the thermodynamic stability of  $\beta$ -LG, in particular at a tertiary level. A state with intact secondary but affected tertiary structure is still abundant at intermediate urea concentrations. On the other hand, the individual far-UV CD and fluorescence titration curves could be well described using a two-state unfolding model suggesting that at a secondary or tertiary level in isolation, the unfolding process does not implicate the accumulation of thermodynamically stable intermediates.



**Figure 2** Urea-induced titration fluorescence: the effect of pH on stability. Unfolding of non-glucosylated and glucosylated  $\beta$ -LG at pH 7.0 and at the isoelectric point (non-glucosylated pH 5.2, glucosylated pH 4.5). ( $\Delta$ ) non-glucosylated  $\beta$ -LG at pH 5.2, ( $\blacktriangle$ ) non-glucosylated  $\beta$ -LG at pH 7.0, ( $\square$ ) glucosylated  $\beta$ -LG at pH 4.5, ( $\blacksquare$ ) glucosylated  $\beta$ -LG at pH 7.0. The dashed line represents a fit using equations (1) and (2) based on a two-state theory.

The glucosylation procedure resulted in the removal of positive charge by covalently attaching glucose to primary amino groups. Therefore, differences in intramolecular

electrostatic interactions introduced by the glucosylation could be responsible for the observed differences in the thermodynamic stability of the tertiary structure. The unfolding transition was therefore studied using intrinsic tryptophan fluorescence at the isoelectric point of the glucosylated (pH 4.5) and non-glucosylated (pH 5.2)  $\beta$ -LG. Using these conditions in combination with a low protein concentration (2.7  $\mu$ M), the unfolding curves were fully reversible and no aggregates accumulated upon unfolding (results not shown). The unfolding transition was well described by a two-state unfolding model (Figure 2). For both  $\beta$ -LG variants, the unfolding transition shifted significantly to a higher midpoint of 6.1 M urea for glucosylated  $\beta$ -LG and 6.3 M urea for non-glucosylated  $\beta$ -LG (Figure 2, Table II). Also the transition appeared more co-operative compared to the transition at pH 7.0 illustrated by a shift in the  $m$ -value for the unfolding transition of both variants (Table II). Despite the unfolding processes of glucosylated and non-glucosylated  $\beta$ -LG appearing more similar at their isoelectric points compared to at pH 7.0, the  $\Delta G$  and the  $m$ -values were significantly different when the proteins were having a similar net charge (Table II). This suggests that the modified net charge by random glucosylation alone does presumably not explain the observed differences in the stability of the tertiary structure as a consequence of glucosylation (Figures 1 and 2).

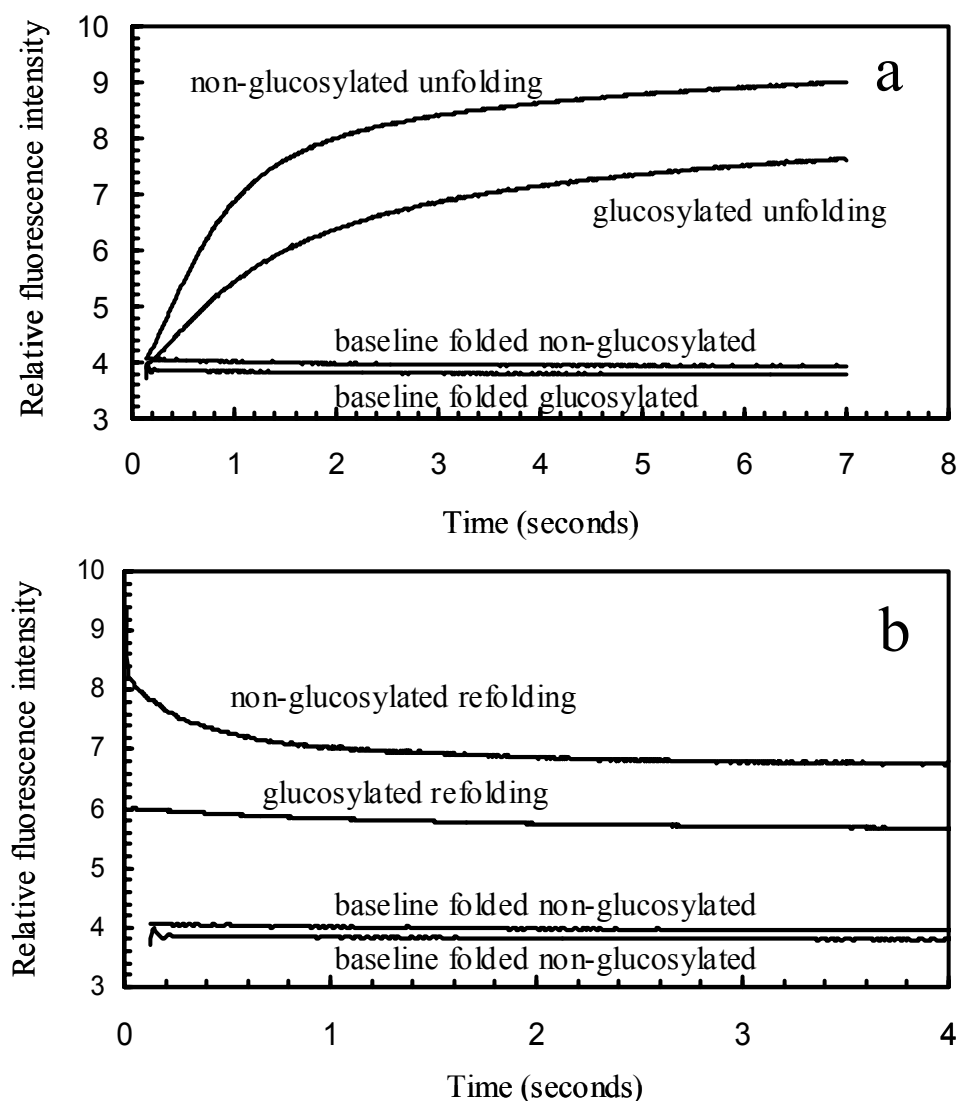
**Table II** *Thermodynamic parameters of the tertiary structure of non-glucosylated and glucosylated proteins determined by intrinsic tryptophan fluorescence*

Protein	$\Delta G^b$ (kJ/mol)	$m$ -value <sup>b</sup> (kJ/mol M <sup>-1</sup> )	$C_M^c$ (M urea/GdnHCl)
$\beta$ -LG (pH 7.0)			
non-gluc	30 ( $\pm$ 1)	-7.6 ( $\pm$ 0.2)	4.0 ( $\pm$ 0.1)
gluc	16 ( $\pm$ 1)	-4.0 ( $\pm$ 0.1)	4.0 ( $\pm$ 0.1)
$\beta$ -LG (IEP <sup>a</sup> )			
non-gluc	55 ( $\pm$ 2)	-8.8 ( $\pm$ 0.4)	6.3 ( $\pm$ 0.3)
gluc	42 ( $\pm$ 2)	-6.9 ( $\pm$ 0.3)	6.1 ( $\pm$ 0.3)
$\alpha$ -LAC			
non-gluc	41 ( $\pm$ 3)	-9.1 ( $\pm$ 0.6)	4.5 ( $\pm$ 0.3)
gluc	15 ( $\pm$ 2)	-2.8 ( $\pm$ 0.3)	5.3 ( $\pm$ 0.6)
MYO			
non-gluc	31 ( $\pm$ 2)	-22.9 ( $\pm$ 1.2)	1.4 ( $\pm$ 0.1)
gluc	24 ( $\pm$ 1)	-16.4 ( $\pm$ 0.5)	1.4 ( $\pm$ 0.1)
LYS			
non-gluc	50 ( $\pm$ 1)	-12.2 ( $\pm$ 0.3)	4.1 ( $\pm$ 0.1)
gluc	39 ( $\pm$ 2)	-9.8 ( $\pm$ 0.4)	4.0 ( $\pm$ 0.2)

<sup>a</sup>IEP = iso electric point

<sup>b</sup> $\Delta G$  and  $m$ -value derived from fluorescence data.

<sup>c</sup> $C_M$  = denaturant-concentration midpoint of unfolding, from fluorescence data

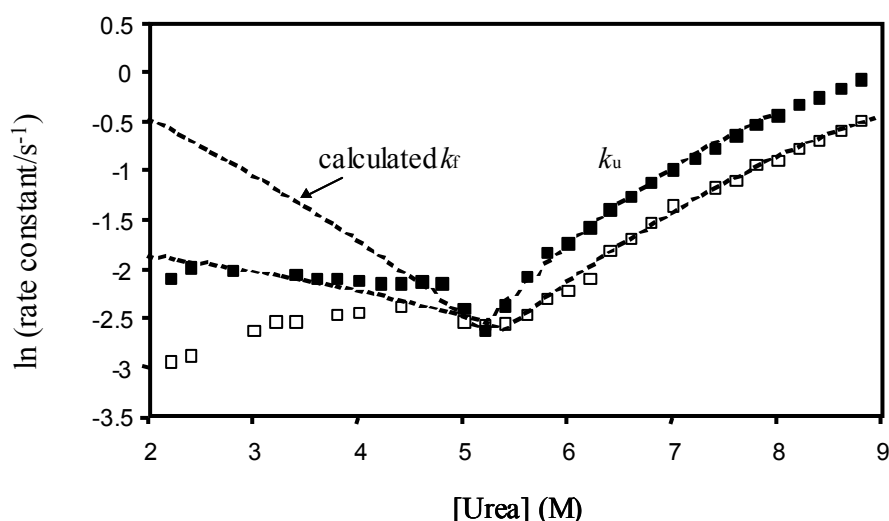


**Figure 3** Unfolding (a) and refolding (b) kinetics of glucosylated and non-glucosylated  $\beta$ -LG. Unfolding rate determined at 8M urea, pH 7.0 and 25 °C using stopped-flow fluorescence,  $\lambda_{ex}$  295 nm,  $\lambda_{em}$  320 nm. The final protein concentration is 0.55 mM. Refolding kinetics obtained from a 10x dilution of protein in 8M urea at similar conditions as unfolding. The baselines of the folded proteins were obtained from dilution in a 10 mM sodium phosphate buffer pH 7.0 to a protein concentration of 0.55 mM.

*The effect of sugars on the kinetics of unfolding* – Kinetic unfolding experiments by time-resolved fluorescence spectroscopy were used to reveal differences in the unfolding kinetics upon glucosylation of  $\beta$ -LG. Figure 3a shows that the rate of unfolding in 8M urea is decreased upon glycosylation. No significant events at a tertiary structure level were apparent during the dead time of the observation as the baseline of the folded proteins starts at the same fluorescence intensity as the unfolding transition signal (Figure 3a). A single exponential can



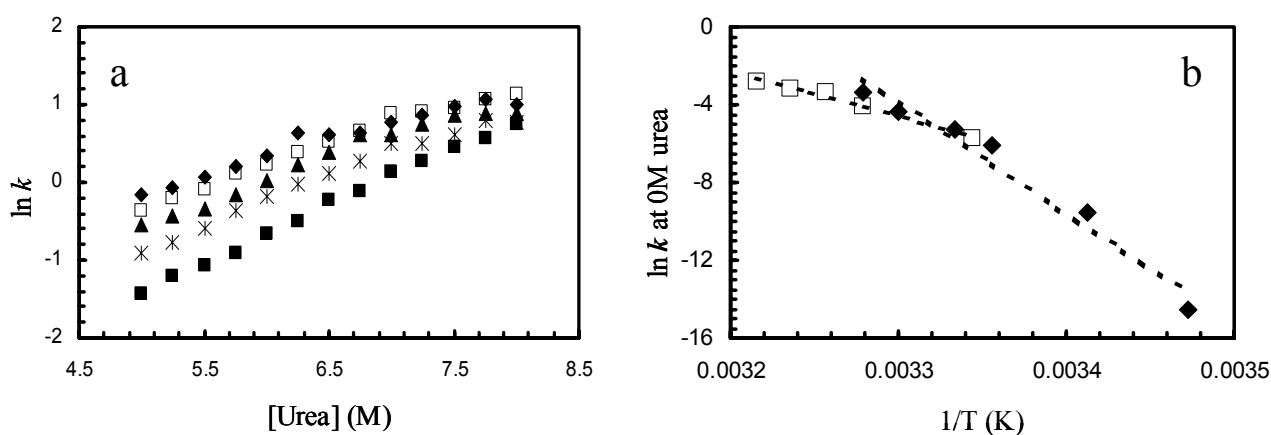
describe glucosylated and non-glucosylated  $\beta$ -LG unfolding at low unfolding rates. The unfolding rate in 8M urea at 25°C was  $0.78 \pm 0.02 \text{ sec}^{-1}$  for non-glucosylated  $\beta$ -LG and  $1.10 \pm 0.02 \text{ sec}^{-1}$  for glucosylated  $\beta$ -LG. The observed rate constants at 25°C for glucosylated and non-glucosylated  $\beta$ -LG folding and unfolding are shown in Figure 4 as a Chevron plot. A downward curvature was observed upon unfolding. According to Matouschek et al (1990) and Camacho and Thirumalai (1996) this indicates the population of an intermediate retarding the reaction rate. For glucosylated  $\beta$ -LG this transition occurred in a similar fashion. From steady-state unfolding experiments, presented in Figures 1 and 2, no thermodynamically stable intermediate was observed using fluorescence upon urea-induced unfolding. The reason why a kinetically rate-limiting unfolding intermediate is present in the unfolding process from kinetic experiments suggests that an intermediate state is involved in the process but that this state is however not a thermodynamically stable state.



**Figure 4** Urea-dependence of the rate of unfolding of non-glucosylated and glucosylated  $\beta$ -LG. Natural logarithm of the observed rate constants as a function of urea concentration (Chevron plot) of non-glucosylated (■) and glucosylated (□)  $\beta$ -LG unfolding and refolding. The calculated  $k_f$  and  $k_u$  were calculated using equations (3) and (4) in the materials & methods section.

Determination of the rate of unfolding at various denaturant concentrations and temperatures (below the thermal transition temperature of the protein), allows the evaluation of the apparent activation energy of unfolding of proteins using equation (5). Figure 5a shows the dependence of the unfolding rate of non-glucosylated  $\beta$ -LG determined at various temperatures and concentrations of urea. At unfolding rates (using higher unfolding

temperatures or higher denaturant concentrations) above  $1.65 \text{ sec}^{-1}$  for non-glucosylated  $\beta$ -LG (Figure 5a) and above  $1.8 \text{ sec}^{-1}$  for glucosylated  $\beta$ -LG (results not shown), the rate of unfolding appears to deviate from linearity indicating that a kinetic intermediate accumulates upon unfolding. Apparently, the pathway of unfolding for both proteins involves a fast unfolding step preceded by a rate-limiting transition. The activation energy calculated from these data returned a value of 400 kJ/mol for non-glucosylated  $\beta$ -LG. Glucosylation caused the activation energy to decrease significantly to 175 kJ/mol (Figure 5b).



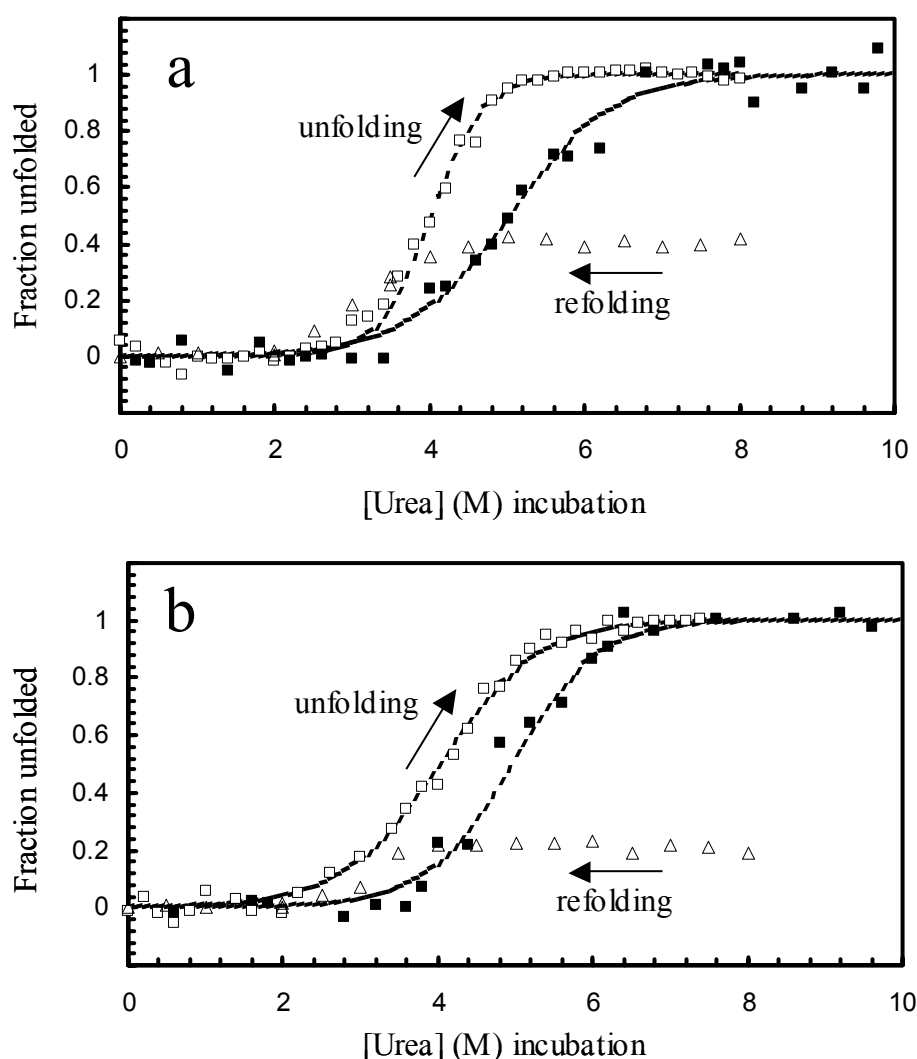
**Figure 5** Arrhenius plot and activation energies of non-glucosylated and glucosylated  $\beta$ -LG using stopped-flow fluorescence. The dependence of  $\ln k$  of unfolding on the urea-concentration (a) at various temperatures from 26 °C to 38 °C (■ 26 °C, \* 32 °C, ▲ 34 °C, □ 36 °C, ◆ 38 °C). The  $\ln k$  values depicted in Figure 5a were extrapolated to 0M urea for the various temperatures tested disclosing the activation energy of unfolding in Figure 5b. The dashed lines represent the fit based on a two-state transition and equation (4).

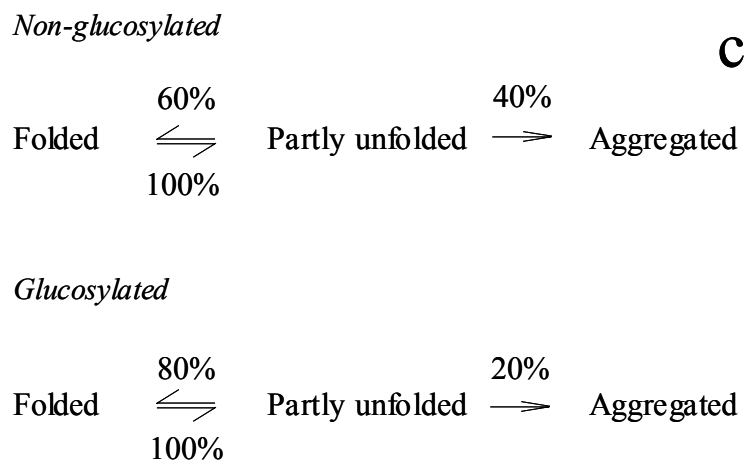
*The effect of sugars on aggregation and refolding* – To determine kinetic partitioning between refolding and aggregation, both  $\beta$ -LG variants were first unfolded in 8M urea and subsequently exposed to an environment, by dilution of the buffer, in which either refolding or aggregation is energetically favourable over the unfolded state. Figure 3b already indicates that non-glucosylated  $\beta$ -LG does not completely refold as the fluorescence trace does not return to the baseline of folded protein. Upon glucosylation, the refolding rate was found to be much increased compared to that of non-glucosylated  $\beta$ -LG. But even though the fluorescence intensity upon refolding approaches the baseline for the folded protein more than that for non-glucosylated protein, still a significant discrepancy between the fluorescence intensities for folded and refolded  $\beta$ -LG is observed (Figure 3b). The fraction refolded or aggregated protein

upon dilution was determined by gel permeation chromatography (GPC). Upon refolding up to an incubation concentration of 2M urea no aggregates were observed (see Figures 6a and b, open triangles). Incubation above this urea-concentration resulted in a gradual increase in the concentration of aggregates formed (as deduced from a decrease in the fraction of monomeric proteins). Apparently, the refolding yield is denaturant-concentration dependent. Above 4M urea no effect of urea-concentration is observed anymore on the refolding yield. For non-glucosylated  $\beta$ -LG it was found that approximately 40% of the unfolded molecules preferred a self-associated state upon dilution (Figure 6a). 20% of the glucosylated  $\beta$ -LG formed aggregates (Figure 6b). Similar percentages were found upon refolding using dialysis against buffer (results not shown). The remaining 60 and 80% of the non-glucosylated and glucosylated proteins respectively, was monomeric and had refolded to their 'native-like' state as was found by far- and near-UV CD (results not shown) and fluorescence studies (Figure 6c). Also, urea-induced unfolding transition curves of refolded glucosylated and non-glucosylated  $\beta$ -LG by fluorescence were yielding comparable results ( $\Delta G$ ,  $C_M$  and  $m$ -value) as the unfolding curves of  $\beta$ -LG that had not been unfolded before (results not shown). Figure 7 summarises the results obtained from GPC, steady-state unfolding by fluorescence and far-UV CD. This figure shows that the extent of unfolding required to induce aggregation varies upon glucosylation. For non-glucosylated  $\beta$ -LG it was found that the aggregation process followed by GPC is detected from a urea-concentration of 2M urea while the unfolding process only initiates at a urea-concentration of 3M. This result suggests that small locally unfolded regions are sufficient to drive the aggregation process. For glucosylated  $\beta$ -LG the detection of aggregates is coincident with the population of intermediates detected by fluorescence and far-UV CD. Upon glucosylation, the tertiary structure is required to unfold to a significant extent to allow the aggregation of proteins.

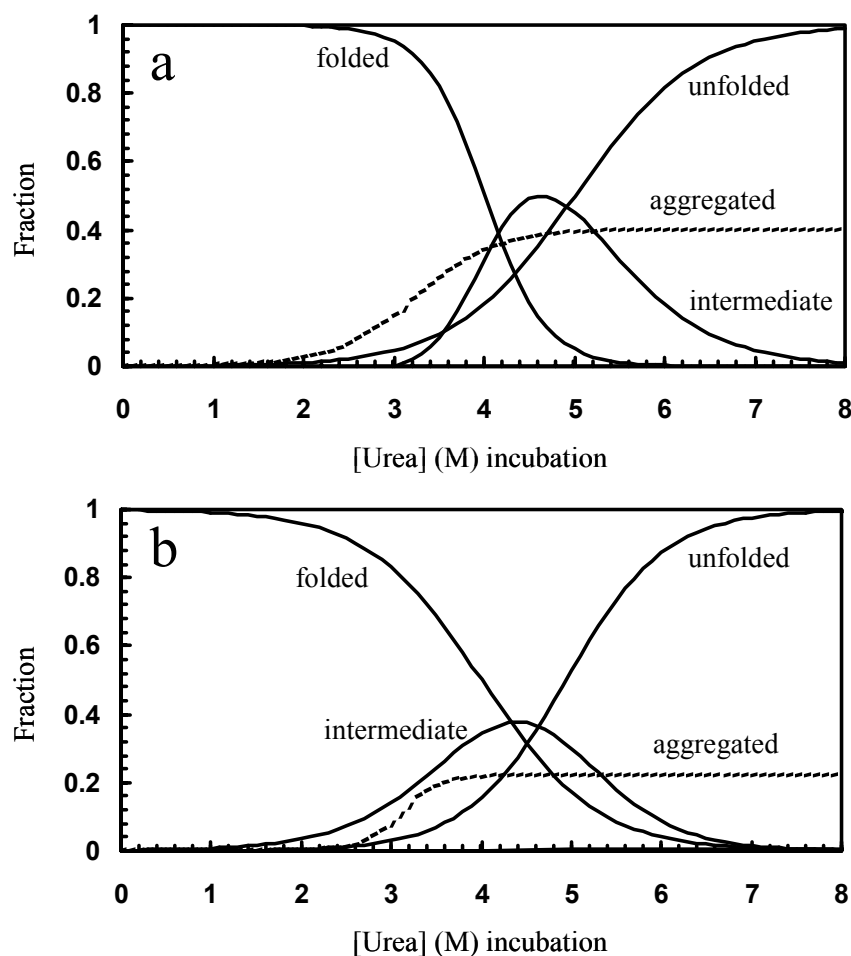
*Is the effect of glucosylation on kinetic partitioning generic?* – To investigate whether the observed changes in  $\beta$ -LG stability and kinetic partitioning upon glucosylation can be ascribed to a general mechanism, other small proteins were glucosylated in a similar way as  $\beta$ -LG. Table I shows that high degrees of glucosylation could be obtained by the glucosylation of LYS,  $\alpha$ -LAC and MYO observed by mass spectrometry. Similar to  $\beta$ -LG, the tertiary and secondary structure of LYS,  $\alpha$ -LAC and MYO, tested by intrinsic tryptophan fluorescence, far-UV CD and near-UV CD, were not affected by the glucosylation procedure (results not shown). It was found that, at a low protein concentration of 2.7  $\mu$ M, all selected

proteins folded reversibly, whether they were glucosylated or not. Table II summarises the results of the thermodynamic characterisation of the selected proteins. It was found that, in general, the  $\Delta G$  upon glucosylation decreased to about half the value of the non-glucosylated variant. We did not find the extent of this decrease to be related to the number of glucose groups attached to the protein. As was found for  $\beta$ -LG, the midpoint of the urea-induced unfolding transition was not affected by glucosylation while the  $m$ -values show a significant change upon glucosylation. Thermodynamically, glucosylation thus appears to decrease protein stability in a similar fashion for a range of proteins.





**Figure 6** Kinetic partitioning upon unfolding using fluorescence, circular dichroism and gel permeation chromatography of non-glucosylated (a) and glucosylated (b)  $\beta$ -LG. ( $\square$ ) fluorescence at 320 nm, ( $\blacksquare$ ) circular dichroism at 203.5 nm, ( $\Delta$ ) gel permeation chromatography peak of aggregated material. The arrows indicate whether the observations are unfolding or refolding experiments. c) Refolding and aggregation yield upon refolding from a urea-denatured state. Data deduced from GPC, CD and fluorescence results.



**Figure 7** Thermodynamic model for the equilibrium between the folded, intermediate, unfolded and aggregated fractions at various urea-concentrations for non-glucosylated (a) and glucosylated (b)  $\beta$ -LG. The model is based on data obtained from Figures 1 and 6.

To investigate whether the change in activation energy for  $\beta$ -LG upon glucosylation is a generic response to protein glucosylation, we performed similar urea-induced kinetic unfolding experiments with glucosylated and non-glucosylated  $\alpha$ -LAC (results not shown). In contrast to  $\beta$ -LG, where glucosylation was found to decrease the activation energy of unfolding, the unfolding activation energy of  $\alpha$ -LAC was determined 200 kJ/mol upon glucosylation, while the activation energy of the non-glucosylated  $\alpha$ -LAC was estimated 101 kJ/mol. The unfolding rates of LYS and MYO could not be determined due to the high rate of unfolding which mainly took place during the dead-time of the experiment. Clearly, glucosylation affects the activation energy of unfolding in a way that is as yet not understood.

As kinetic partitioning between refolding and aggregation was previously found to predict the probability for self-association of  $\beta$ -LG, similar refolding experiments were

performed for glucosylated and non-glucosylated variants of the selected proteins. MYO glucosylation resulted in a higher preference to refold compared to non-glucosylated protein, similar to  $\beta$ -LG. The aggregation propensity was thus decreased upon attaching glucose residues to the protein. Refolding of non-glucosylated LYS and  $\alpha$ -LAC gave no aggregates, even at very high protein concentrations such as 17.5 mM. Glucosylation of LYS and  $\alpha$ -LAC also did not result in self-association despite the fact that these proteins displayed a significantly reduced stability.

## Discussion

*In vivo* glycosylation of proteins has been associated with a comprehensive spectrum of functions, for example to interference with the folding process of many proteins synthesised in eukaryotes (Land & Braakman 2001, Helenius & Aebi 2001). In this paper it is shown that non-specific glucosylation of a number of small globular proteins also directs protein folding and limits the possibilities for aggregation *in vitro*. The precise impact of glucosylation on thermodynamic parameters related to kinetic partitioning between aggregation and refolding have been explored and the outcome can be of significance in the interpretation of the relevance of protein glycosylation.

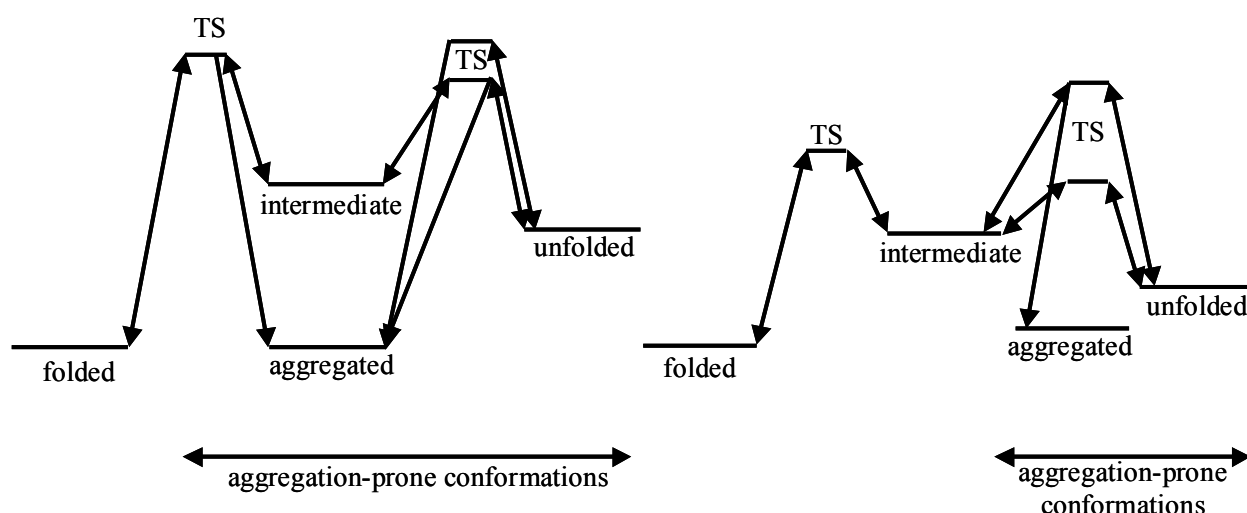
*How do covalently conjugated sugar moieties affect the in vitro folding of small globular proteins?* – It has been reported that glycosylation can have a direct effect on the folding process of proteins in the endoplasmic reticulum (Helenius & Aebi 2001). When glycosylation is inhibited, the most commonly observed effect is the generation of misfolded or aggregated proteins that fail to reach a functional state (Paulson 1989, Helenius 1994, Varki 1993). These results imply that the kinetic partitioning between folding and aggregation can be altered by cotranslationally introduced sugar chains. Our results on the effect of glucosylation on *in vitro*  $\beta$ -LG folding support these results in a number of ways. The first effect observed of glucosylation of  $\beta$ -LG is that the aggregation propensity upon dilution of urea-induced unfolded  $\beta$ -LG is reduced, as was found by kinetic partitioning experiments (Figure 6). This effect can be explained in a number of ways. One likely mechanism for the improved efficiency of refolding reported before in literature (Imperiali & O'Connor 1999, Kern et al 1993) is that the presence of sugar units may promote the refolding process of proteins by orienting the polypeptide segment toward the surface of protein domains. *In vitro* folding studies using glycosylated and non-glycosylated counterparts of the same protein have

also confirmed that oligosaccharide moieties guide the folding process by such a mechanism and thus simulate the chaperone effect *in vivo*. Figure 3b indeed suggests that the refolding rate upon glucosylation is increased upon dilution from a denaturant containing environment. As the protein concentration used in time-resolved fluorescence experiments is 0.55 mM, which was found to induce self-association upon unfolding (Figures 6 and 7), the refolding side of the Chevron plot (Figure 4) is a combined result of refolding and aggregation and the rate of refolding can therefore not be deduced. A second indication lies in the finding that non-specific glucosylation reduced the activation energy of unfolding of  $\beta$ -LG (Figure 5b). This reduction is possibly related to an effect of the interruption of existing ion pairs on the surface of the protein. The contribution of ion pairs in the kinetic control of unfolding has been investigated before (Oliveberg & Fersht 1996, Waldburger et al 1996). Values for the hydration enthalpies for carboxylic and amino groups have been reported in the order of 38 kJ/ mol (Makhatadze & Privalov 1992, Gruia et al 2004, Masunov & Lazaridis 2003). Glucosylation by the reaction of glucose with the lysine residue potentially results in a reduced number of potentially ion-paired charges of the protein. Based on the findings of the contribution of ion pairs to the kinetic barrier for unfolding, it can be assumed that the glucosylation reaction and resulting rupture of ion pairs implies a reduction of the energy penalty for dehydration and rehydration upon unfolding causing a decrease in the activation energy of unfolding upon glucosylation. Glucosylation reduced the activation energy for unfolding of  $\beta$ -LG from 400 kJ/mol down to 175 kJ/mol. This difference can be theoretically accounted for by the disruption of six ion pairs. This suggests that not all lysine residues may be involved in ion pairs but, for the larger part, may be already hydrated when exposed at the surface of the protein. However, literature clearly states the phenomenon of a strong impact of disruption of ion pairs on the activation energy of unfolding and, based on the approach we took for glycosylation, the results obtained seem to fit reasonably well this assumption.

*Direct consequences of glucosylation for (un)folding* – Glucosylation lowers the stability of proteins at ambient temperatures (Table II) as reflected by a strong reduction of the  $\Delta G$  upon denaturant-induced unfolding. This can be primarily attributed to a reduction of the  $m$ -value of the unfolding reaction upon glucosylation whereas the midpoint of unfolding is not affected by glycosylation. In contrast, the denaturant-induced unfolding transition of  $\beta$ -LG determined using far-UV CD indicates only a small change in  $\Delta G$ . These results suggest that glucosylation reduces the difference in stability of the folded and unfolded state quite specifically of the tertiary structure of  $\beta$ -LG and it apparently does not lead to a global



destabilisation of the protein structure, but only affects the stability in specific regions (Figures 1a and b). It was found that electrostatics only partially explained the differences observed in thermodynamic stability (Figure 2). Previously, it has been found that, at elevated temperatures, glycosylation leads to a stabilisation of the protein structure, as reflected by an increased thermal transition temperature (Broersen et al 2004). These, at first sight contrasting results, can be explained by the recently shown significant lowering of the heat capacity change upon unfolding by glucosylation, leading to a different temperature dependence of thermodynamic stability (van Teeffelen et al 2005). Van Teeffelen et al (2005) assigns this effect to the ability of the introduced sugar molecules to undergo complementary hydrogen bonding with the polypeptide backbone. This explanation reflects the possibility that an entropic penalty of the restricted backbone flexibility reduces the entropic stability of the protein as has been previously established (Pieper et al 1996, Jentoft 1990, Kato 2000).



**Figure 8** Thermodynamic and kinetic model describing the effect of protein glucosylation on urea-induced unfolding and aggregation. The model describes the kinetic and thermodynamic barriers for unfolding and aggregation processes. TS is the state with the maximum free energy in the pathway (Solís-Mendiola et al 1998).

The kinetic implications of glucosylation cannot be explained by a generic mechanism. It was found that the activation energy of unfolding does not react in a predictable manner upon glucosylation. It was found that the activation energy of unfolding is significantly reduced upon glucosylation of  $\beta$ -LG (Figure 5) while a significant increase was found upon glucosylation of  $\alpha$ -LAC (results not shown). Upon glucosylation, the potential positive charge of the primary amines is removed. Breaking ion pairs and salt bridges in the

protein structure should be particularly reflected in a reduction of the activation energy of unfolding. As has been outlined before, unfolding requires a large energy penalty resulting from dehydration of the intact ion pairs and subsequent rehydration of the individual charged groups. As the effect of glucosylation was not found to be consistent, it is more likely that specific effects of glucosylation associated to the location of the modifications in the primary sequences, the properties of proximate residues and fold play a key role in determining the activation energy of unfolding.

*Does glucosylation improve the refolding efficiency of proteins?* – The Chevron plot showed that the refolding curve deviates from the theoretically calculated fitted curve by a significant rollover (Figure 4). It was found that this rollover originated from irreversible aggregation as confirmed by gel permeation chromatography experiments (Figure 6). Low reversibility upon refolding after urea-incubation of  $\beta$ -LG at pH 7.0 has been found before by Yagi et al (2003). Using competition experiments it was investigated whether glucosylation affects kinetic partitioning between refolding and aggregation. Upon dilution of the denaturant in buffer, a fraction of the molecules were found to self-associate instead of to refold to their fully folded state. It was found that glucosylation improves the refolding efficiency of MYO (results not shown) and  $\beta$ -LG (Figure 6). The refolded fractions for both glucosylated and non-glucosylated proteins consisted of fully refolded species showing a similar sigmoidal unfolding curve upon urea-denaturation (results not shown) as the folded protein. Figure 8 depicts a model in which the aggregated system can be defined as an energetically stable end-point occurring only upon partial unfolding of non-glucosylated  $\beta$ -LG (Figure 8a) and more significant unfolding required for aggregation upon glucosylation (Figure 8b). The reduced aggregation upon glucosylation can be explained twofold on a molecular level. Faster refolding through glycosylation can be a consequence of hydrophilic tagging of the outside of the folded polypeptide chain guiding the unfolded molecule more efficiently to its folded state. It has been reported before that sugar chains may interact with hydrophobic regions of the protein by reorientation of the sugar backbone (Imperiali & O'Connor 1999, Wormald & Dwek 1999). The limited flexibility of the sugar-amide bond limit this possibility in the folded state, as was found by PRODAN-binding measurements of the folded state (results not shown). Upon unfolding, the polypeptide chain gains entropic flexibility. In addition, the covalently attached sugar chains interact with the hydrophobic regions providing an energetically more favourable situation in the unfolded state (depicted in

Figure 8 by an affected energy state of the unfolded protein upon glucosylation). These results are supported by the finding that the heat capacity change significantly decreases upon glucosylation (van Teeffelen et al 2005). In this way, hydrophobic regions of the protein are less accessible for self-association. This finding also places into context why, upon glucosylation, partial unfolding is not sufficient to induce aggregation (Figure 7), while non-glucosylated  $\beta$ -LG allows aggregation by minimal unfolding (Figure 7). This, combined with a very fast transition to the folded state can reduce the possibilities of glycosylated proteins to adopt an aggregated state even though the protein is destabilised upon glucosylation. Next to this, the appearance of the intermediate structures are an important factor in aggregation propensity. Of course the generalisation of these results obtained from glucose attached to small globular proteins to larger glycan residues to far more complex protein systems can be questioned and further studies can be used to confirm this. However, the fact that varies types of glycosylation to a wide range of proteins have shown consistent effects with respect to protein stability and aggregation propensity, should not be neglected. From the observations, it can be established that protein stability alone does not predict the propensity of proteins to aggregate. The refolding rate and efficiency seem to dictate the balance of molecules that are prone to aggregation under defined conditions. This balance can no doubt be strongly affected by glycosylation.

## Materials and methods

*Preparation of glucosylated proteins* – Hen lysozyme (LYS), horse heart myoglobin (MYO) and bovine  $\alpha$ -lactalbumin ( $\alpha$ -LAC) were obtained from Sigma and used without further purification. Bovine  $\beta$ -lactoglobulin ( $\beta$ -LG) was purified by a protocol by de Jongh et al (2001). Typically, 0.273 mM protein was dissolved in demineralised water. Per mL protein solution, 27.8 mM D-glucose (Sigma) was added. The pH was adjusted to 8.0 using 1M NaOH and the mixture was lyophilised. Next, the mixed powders were incubated at 60°C for 5 h at a relative humidity of 65% (atmosphere equilibrated with supersaturated NaNO<sub>2</sub> solution). The incubated samples were next dissolved in demineralised water and dialysed against demineralised water to remove all the non-reacted sugar from the solution (see also reference Broersen et al (2004) for details and validation of this procedure). Subsequently, the solutions were lyophilised and stored at -20°C until use.

*Matrix-Assisted-Laser-Desorption/Adsorption Ionization-time-of-flight spectroscopy (MALDI-tof MS)* – MALDI-TOF mass spectra were acquired to determine the degree of

modification as reported previously (Broersen et al 2004) on a Voyager-DE™ RP (PerSeptive Biosystems Inc., USA) using a 3,5-dimethoxy-4-hydroxycinnamic acid matrix. Repeat measurements resulted in a standard error of  $\pm 4$  Da.

*Circular Dichroism (CD) spectroscopy* - Far-UV CD spectra of 5.5  $\mu$ M non-glucosylated and glucosylated proteins in a 10 mM sodium phosphate buffer (pH 7.0) were recorded on a Jasco J-715 spectropolarimeter (Jasco Corporation Japan) at 25°C in the spectral range from 260 to 190 nm with a spectral resolution of 0.5 nm. Spectra were recorded as averages of eight scans. The scanning speed was 100 nm/min and the response time was 0.50 s with a bandwidth of 1.0 nm. Quartz cuvettes with an optical path of 1 mm were used. The spectra were corrected for the corresponding protein-free sample.

*Steady-State fluorescence spectroscopy* - Emission fluorescence spectra of non-glucosylated and glucosylated protein solutions of 2.7  $\mu$ M in a phosphate buffer (10 mM, pH 7.0) were obtained using a Cary Eclipse (Varian) fluorimeter at 20°C. The emission spectra from 300 to 400 nm were determined using excitation at 295 nm. Excitation and emission slit widths were 5 nm and a scan speed of 100 nm/min was used. All spectra were corrected for the corresponding protein-free sample. The measurements were performed in duplicate.

*Change in free energy by urea/guanidine hydrochloride titration* – Emission fluorescence spectra were recorded using a Cary Eclipse (Varian) fluorimeter at a protein concentration of 2.7  $\mu$ M in a phosphate buffer (10 mM, pH 7.0).  $\beta$ -LG and  $\alpha$ -LAC were equilibrated overnight at a range of 0 – 8 M urea and MYO and LYS were equilibrated overnight at a range of GuHCl concentrations from 0 - 6 M at 25°C in a 10 mM phosphate buffer (pH 7.0). The fluorescence emission was recorded at a scan speed of 100 nm/min between 300 and 400 nm with an excitation at 295 nm and an excitation and emission slit width of 5 nm. Denaturant titration curves were obtained by plotting the band intensity (selected at the maximum intensity of the folded protein emission spectrum) for a fixed emission wavelength as a function of denaturant concentration.

Alternatively, the change in free energy was determined using the same samples described above using far-UV CD on a Jasco J-715 spectropolarimeter (Jasco Corporation Japan) in the range from 260 to 190 nm with a spectral resolution of 0.5 nm. Spectra were recorded as averages of 16 spectra. The scanning speed was 100 nm/min and the response time was 0.50 s with a bandwidth of 1.0 nm. Quartz cuvettes with an optical path of 1 mm were used. The spectra were corrected for the corresponding protein-free sample. The titration

curves were plotted by the CD intensity at 203.5 nm (zero-crossing in the folded state) as a function of denaturant concentration.

*Stopped-flow fluorimetry* – The kinetics of urea-induced unfolding of glucosylated and non-glucosylated proteins was monitored at 26°C, 32°C, 34°C, 36°C and 38°C. In stopped-flow experiments, folded  $\beta$ -LG at a concentration of 0.55 mM in 10 mM phosphate buffer (pH 7.0) was unfolded in various urea concentrations from 5 to 8 M. Two buffers, one containing ~ 10 M urea and another with no denaturant, both at pH 7.0, were degassed, transferred to gastight syringes, loaded in the stopped-flow set-up and equilibrated for 15 min to the appropriate temperature. Unfolding to different final concentrations of urea was initiated by mixing 40  $\mu$ L of the folded protein solution with 360  $\mu$ L of the two buffers mixed to yield the desired denaturant concentration. These measurements were performed using a SFM4 stopped-flow fluorescence instrument (Biologic). An 0.8 mm square flow cell was used. The dead time of the instrument, determined by using the procedure prescribed by the instrument manufacturer, was ~2 ms. The excitation and emission slit widths were 1 nm. The reproducibility of ten individually recorded unfolding rates was high with an average unfolding rate of  $1.105408 \pm 0.02213 \text{ sec}^{-1}$  which indicates a standard error of 2%.

*Aggregation* – Proteins were incubated at a range of urea concentrations (0 – 8M) at pH 7.0 for 24 h at 25°C. Then the protein/denaturant mixtures were diluted tenfold in 10 mM phosphate buffer (pH 7.0) and 50 mM NaCl. To quantify the refolding efficiency, the diluted samples were immediately injected onto a 24 mL Superdex 75 column HR 10/30 (Amersham) at 25°C at a flow rate of 0.4 mL/min for analysis with gel permeation chromatography (GPC). The monomeric protein concentration in the eluent was detected at 280 nm. The monomeric and aggregated fractions were collected separately. The refolded fraction of the proteins was obtained and tested for native-like structural elements using far-UV CD and urea titrations with fluorescence as described above.

*Data analysis* – The change in free energy from the data obtained from denaturant-titration using CD and fluorescence are fitted using software (MATLAB) based on the following equations describing a two-state unfolding transition assuming the thermodynamic existence of only a folded and an unfolded fraction (Matouschek et al 1990):

$$\Delta G = \Delta G_{H_2O} - m [\text{denaturant}] \quad (1)$$

$$\Delta G = - R T \ln K \quad (2)$$

Where  $\Delta G$  represents the change in free energy,  $\Delta G_{H_2O}$ , change in free energy in water,  $m$  is a constant proportional to the increase in degree of exposure to solvent of the protein on denaturation,  $R$  is the gas constant ( $8.135 \times 10^{-3} \text{ kJ / mol} \cdot \text{K}$ ),  $T$ , temperature (K), and  $K$  is the equilibrium constant (fraction folded/fraction unfolded).

The rate constant of unfolding  $k_u$ , increases with increasing denaturant concentration according to equation (3) where  $k_u^{H_2O}$  is the rate constant in  $H_2O$  (Kuwait et al 1989, Matouschek et al 1989, Bycroft et al 1990).

$$\log k_u = \log k_u^{H_2O} + m_{ku} [\text{denaturant}] \quad (3)$$

The rate constant of folding,  $k_f$ , of a protein that involves a reversible transition between the folded and the unfolded state must follow the rate law:

$$\log k_f^{H_2O} = \log (k_u^{H_2O} / K_u^{H_2O}) \quad (4)$$

$k_f$  can be calculated from the unfolding rate constants and equilibrium data (Matouschek et al 1990) and will be compared to the measured data to obtain information on the reversibility of the refolding reaction.

The activation energy of unfolding was calculated by the Arrhenius equation:

$$\ln k = - (E_{act} / R) \cdot (1 / T) + \ln A \quad (5)$$

Where  $k$  represents the apparent reaction rate,  $E_{act}$  is the activation energy (kJ/mol),  $R$  is the gas constant,  $T$ , the temperature in Kelvin and  $A$  is the frequency factor.

## References

- Bouma B, Kroon-Batenburg LMJ, Wu Y-P, Brünjes B, Posthuma G, Kranenburg O, de Groot PG, Voest EE, Gebbink MFBG (2003) J Biol Chem 278: 41810-41819
- Broersen K, Voragen AGJ, Hamer RJ, de Jongh HHJ (2004) Biotech Bioeng 86: 78-87
- Brownlee M, Vlassara H, Cerami A (1984) Ann Intern Med 101: 527-537
- Bycroft M, Matouschek A, Kellis JT, Serrano L, Fersht AR (1990) Nature 340: 488-490
- Camacho CJ, Thirumalai D (1996) Protein Sci 5: 1826-1832
- Chi EY, Krishnan S, Kendrick BS, Chang BS, Carpenter JF, Randolph TW (2003). Protein Sci 12: 903-913
- Choei H, Sasaki N, Takeuchi M, Yoshida T, Ukai W, Yamagishi S, Kikuchi S, Saito T (2004) Acta Neuropathol 108: 189-193

- Crabbe MJ (1998) *Cell Mol Biol* (Noisy-le-Grand) 44: 1047-1050
- De Jongh HHJ, Gröneveld T, de Groot J (2001) *J Dairy Sci* 84: 562-571
- Dobson CM (2003) *Nature* 426: 884-890
- Fenouillet E, Jones IM (1995) *J Gen Virol* 76: 1509-1514
- Fink AL (1998) *Fold Des* 3: R9-R23
- Gruia AD, Fischer S, Smith JC (2004) *Chem Phys Let* 385: 337-340
- Hebert DN, Foellmer B, Helenius A (1996) *EMBO J* 15: 2961-2968
- Helenius A, Aeby M (2001) *Science* 291: 2364-2369
- Helenius A (1994) *Mol Biol Cell* 5: 253-265
- Helenius A, Trombetta ES, Hebert DN, Simons JF (1997) *Trends Cell Biol* 7: 193-200
- Hurle MR, Helms LR, Li L, Chan WN, Wetzel R (1994) *Proc Natl Acad Sci USA* 91: 5446-5450
- Hurtley SM, Bole DG, Hoover-Litty H, Helenius A, Copeland CS (1989) *J Cell Biol* 108: 2117-2126
- Imperiali B, O'Connor SE (1999) *Curr Opin Chem Biol* 3: 643-649
- Jentoft N (1990) *Trends Biochem Sci* 15: 291-294
- Kato A, Nakamura S, Ban M, Azakami H, Yutani K (2000) *Biochim Biophys Acta* 1481: 88-96
- Kern G, Kern D, Jaenicke R, Seckler R (1993) *Protein Sci* 2: 1862-1868
- Kuwajima K, Mitani M, Sugai S (1989) *J Mol Biol* 206: 547-561
- Kuwajima K, Nitta K, Sugai S (1975) *J Biochem (Tokyo)* 78: 205-211
- Land A, Braakman I (2001) *Biochimie* 83: 783-790
- Makhatadze GI, Privalov PL (1992) *J Mol Biol* 226: 491-505
- Masunov A, Lazaridis T (2003) *J Am Chem Soc* 125: 1722-1730
- Matouschek A, Kellis JT Jr, Serrano L, Bycroft M, Fersht AR (1990) *Nature* 346: 440-445
- Matouschek A, Kellis JT, Serrano L, Fersht AR (1989) *Nature* 340: 122-126
- Meldgaard M, Svendsen I (1994) *Microbiology* 140: 159-166
- Monnier V, Cerami A (1981) *Science* 211: 491-493
- Oliveberg M, Fersht AR (1996) *Biochemistry* 35: 6795-6805
- Otzen DE, Oliveberg M (2004) *Protein Sci* 13: 3253-3263
- Paulson JC (1989) *Trends Biochem Sci* 14: 272-276
- Pieper J, Ott K-H, Meyer B (1996) *Nat Struct Biol* 3: 228-232
- Ptitsyn OB (1987) *J Protein Chem* 6: 273-293
- Solís-Mendiola S, Gutiérrez-González LH, Arroyo-Reyna A, Padilla-Zúñiga J, Rojo-Domínguez A, Hernández-Arana A (1998) *Biochim Biophys Acta* 1388: 363-372
- Subramaniam V, Steel DG, Gafni A (1996) *Protein Sci* 5: 2089-2094
- Van Teeffelen AMM, Broersen K, de Jongh HHJ (2005) Glucosylation of  $\beta$ -lactoglobulin lowers the heat capacity change of unfolding, a unique way to affect protein thermodynamics. Accepted for publication in *Protein Sci*.
- Varki A (1993) *Glycobiology* 3: 97-130
- Waldburger CD, Jonsson T, Sauer RT (1996) Barriers to protein folding: *Proc Natl Acad Sci USA* 93: 2629-2634
- Wang C, Eufemi M, Turano C, Giartosio A (1996) *Biochemistry* 35: 7299-7307

Wormald MR, Dwek RA (1999) Structure 7: R155-R160

Yagi M, Sakurai K, Kalidas C, Batt CA, Goto Y (2003) J Biol Chem 278: 47009-47015



# CHAPTER 4

## GLUCOSYLATION OF $\beta$ -LACTOGLOBULIN LOWERS THE HEAT CAPACITY CHANGE OF UNFOLDING, A UNIQUE WAY TO AFFECT PROTEIN THERMODYNAMICS

van Teeffelen AMM, Broersen K, de Jongh HHJ

Protein Science 2005, 14: 2187-2194

### Abstract

Chemical glycosylation of proteins occurs spontaneously *in vivo*, especially under stress-conditions, and has been linked in a number of cases to diseases related to protein unfolding and aggregation. It is the aim of this work to study the origin of the change in thermodynamic properties due to glucosylation of the folded  $\beta$ -lactoglobulin A. Under mild conditions Maillard products can be formed by reaction of  $\epsilon$ -amino groups of lysines with the reducing group of, in this case, glucose. The formed conjugates described here have an average degree of glucosylation of 82%. No impact of the glucosylation on the protein structure is detected, except that the Stokes' Radius was increased by about 3%. Although at ambient temperatures the change in Gibbs energy of unfolding is reduced by 20%, the unfolding temperature is increased by 5°C. Using a combination of circular dichroism, fluorescence and calorimetric approaches it is shown that the change in heat capacity upon unfolding is reduced by 60% due to the glucosylation. Since in the unfolded state the Stokes' Radius of the protein is not significantly smaller for the glucosylated protein, it is suggested that the non-polar residues associate to the covalently linked sugar moiety in the unfolded state, thereby preventing their solvent exposure. In this way coupling of small reducing sugar moieties to solvent exposed groups of proteins offers an efficient and unique tool to deal with protein stability issues, not only relevant in nature, but also for technological applications.

### Keywords:

$\beta$ -lactoglobulin, Maillard reaction, protein stability, heat capacity, glycosylation, thermodynamics

## Introduction

A common mechanism for proteins *in vivo* to cope with stress conditions is to increase the degree of glycosylation (Murakami et al 1997). Good examples of such response are found for heat shock proteins (Venetianer et al 1994, Verma et al 1988). Also from *in vitro* studies it has been reported that glycosylation improves the conformational stability upon heating (Marshall & Rabinowitz 1976, Meldgaard & Svendsen 1994, Wang et al 1996). Alternatively, de-glycosylation of N-linked oligosaccharides has been reported to make some proteins more susceptible to aggregation on heat-treatment (Endo et al 1992). These studies exemplify that carbohydrate moieties attached to the protein surface are an important tool to regulate protein stability and can play an important role in the occurrence of misfolding and aggregation related diseases (see f.e. Ermonval et al 2000). The thermodynamic mechanism behind the protein stabilising role of sugar moieties is, however, unclear.

As demonstrated by Kusters et al (2003), modification of globular protein surface charges with various reagents resulted in all cases studied in a lowering of the Gibbs energy of stabilisation ( $\Delta G$ ). Moreover, the effect on the thermal transition temperature ( $T_t$ ) scaled with a  $\delta(\Delta G)/\delta T_t$  of approximately 0.03 kJ per mol residue per K (Kusters et al 2003), as reported more often for various proteins (Rees & Robertson 2001). Most of the modifications/mutations affect the enthalpic contribution to the protein stability by engineering the charged groups on the protein surface, a strategy often employed by thermophiles (Loladze et al 1999, Grimsley et al 1999, Spector et al 2000, Perl et al 2000, Sanchez-Ruiz & Makhatadze 2001). Studies on glucosylation or fructosylation of  $\beta$ -lactoglobulin suggested, however, a different behaviour (Broersen et al 2004). The reaction of these reducing mono-sugars with primary amino groups (lysine-residues) at slightly elevated temperature (45-55°C) for 3-5 hours, yielded glycosylated proteins with high degree of modification (so-called Maillard-conjugates). These products displayed an increase in  $T_t$  and at the same time a lower Gibbs energy of stabilisation at ambient temperatures. The question arises what the origin is of the altered thermodynamics by glycosylation. Since such (reversible) conjugates are encountered *in vivo*, for example in diabetic cataract formation (Stevens et al 1978), an understanding of the thermodynamic impact of this specific modification is important. While altered enthalpy contributions to the protein Gibbs energy only 'shift' the  $\Delta G(T)$  curve, changes in the  $\Delta C_p$  affect the shape of this dependency.

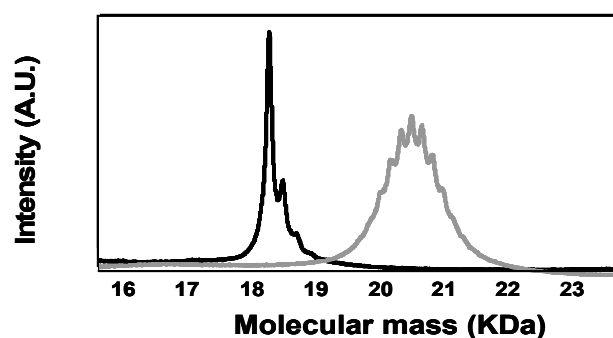
It is the aim of this work to enlighten the origin of the change in thermodynamic properties caused by Maillardation of proteins. For that purpose we selected bovine whey

protein  $\beta$ -lactoglobulin A ( $\beta$ -LGA), a protein which thermodynamic properties are well documented (see f.e. Griko & Privalov 1992). This protein is known to be susceptible *in vivo* to spontaneous lactosylation (Fogliano et al 1998), and in processing (heat-treatment) for technological applications often glycosylated forms of this protein are found. Applying a combination of methodologies to establish thermodynamic properties of the proteins, like the  $\Delta C_p$ , together with determination of Stokes' Radii to evaluate changes in solvent accessible area, provides a new insight in the origin of the altered properties of the protein. The outcome is discussed in the light of both life science and technological consequences.

## Results and Discussion

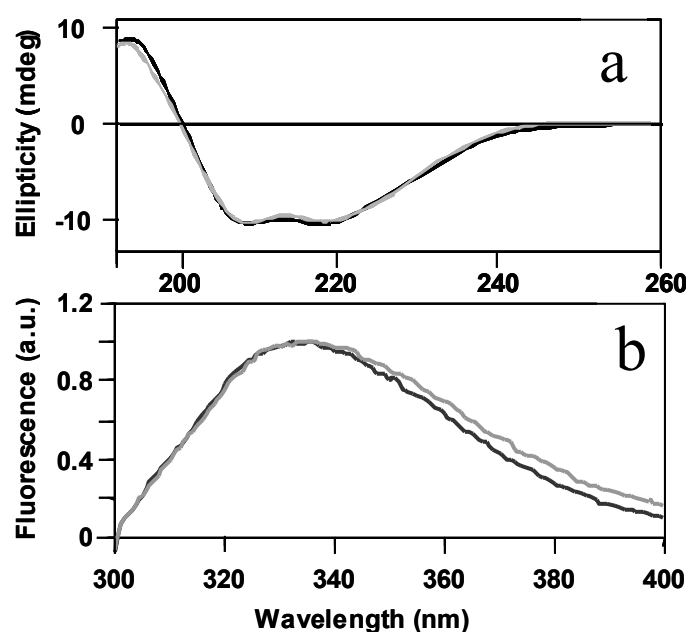
### *Glucosylation of $\beta$ -lactoglobulin A*

Figure 1 shows the mass spectrometric analysis of non-glucosylated and glucosylated  $\beta$ -lactoglobulin A ( $\beta$ -LGA). The mass of  $\beta$ -LGA is 18348 Da, in agreement with the expected mass (de Jongh et al 2001). The small contributions at +223, 446 and 670 Da arise from matrix-adducts (mass of sinapinic acid is 224 Da). Upon formation of Maillard products with glucose a symmetric Gaussian distribution of masses is found, centred around 20613 Da, corresponding with an average degree of modification of 82% (i.e. 14 primary amino groups have reacted with D-glucose). The width of the (symmetric) mass distribution indicates that the heterogeneity in the sample is limited to significant contributions of  $\pm 1$ , 2, and 3 sugar groups per protein. Such heterogeneity is in agreement with previous findings (Broersen et al 2004). The degree of modification was also confirmed by OPA-analysis, as described in detail in the method section. From the latter analysis an ensemble-averaged number of reacted groups of  $14 \pm 1$  was found (results not shown).



**Figure 1** Mass spectrometric analysis of non-glucosylated (black line) and glucosylated  $\beta$ -LGA (grey line).

The impact of the modification on the structure of the protein was studied using far-UV circular dichroism (Figure 2a), near-UV CD (not shown) and tryptophan fluorescence (Figure 2b). As can be seen from Figure 2a the secondary structure was not affected by the glucosylation, based on the resemblance of the two spectra. That the local environment of the aromatic residues was also not significantly affected by the modification is illustrated in Figure 2b by the band positions of the two tryptophans in the protein (residue 19 and 61), that do not appear shifted by the modification. This was confirmed by near-UV CD spectra (not shown). Moreover, from the tyrosine to tryptophan energy transfer efficiency by comparative excitation at 274 and 295 nm (not shown) also no significant differences could be detected. It was already shown previously that glucosylated  $\beta$ -LG was capable of adopting the non-covalent dimer-organisation, just like the intact, folded protein (Broersen et al 2004).



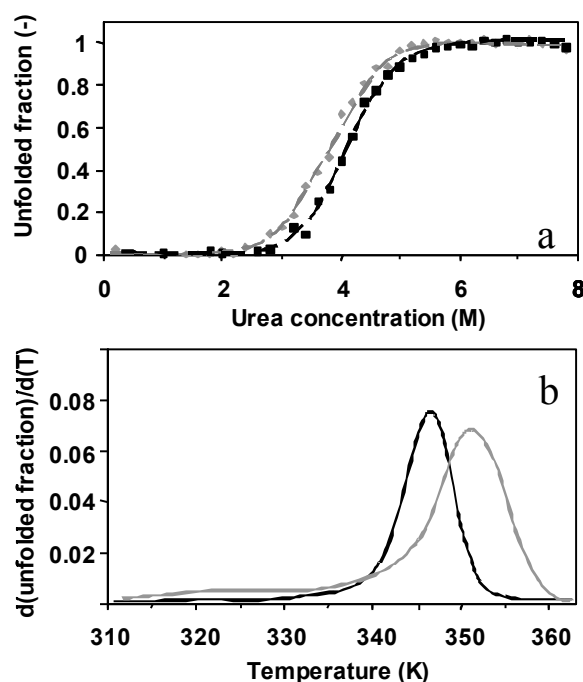
**Figure 2** Conformational properties of 0.1 mg/mL non-glucosylated (black lines) and glucosylated (grey lines)  $\beta$ -LGA in 10 mM phosphate-buffer (pH 7.0) at 20°C at a secondary (far-UV CD, panel a) and tertiary (tryptophan fluorescence, panel b) folding level.

### Conformational stability

The energy change upon unfolding of the non-glucosylated and glucosylated  $\beta$ -LGA was studied using urea unfolding experiments at 20°C, where the unfolding was monitored by detection of the tryptophan fluorescence spectra. Normalisation of the fluorescence intensity at 320 nm and subsequent conversion to fraction unfolded proteins, yield the curves shown in Figure 3a. Upon glucosylation the mid-point of unfolding shifts approximately 0.35 M to

lower urea concentration. Also, the steepness of the curve decreases. Based on the apparent sigmoidality of the curves a two-state unfolding can be assumed (f.e. see Jackson & Fersht 1991). Analysis of these curves using the linear extrapolation model yields then values for the Gibbs energy change ( $\Delta G$ ) of 25.3 and 20.7 ( $\pm 0.3$ ) kJ/mol for the non-glucosylated and glucosylated protein respectively. It is important to note that the heterogeneity in the glucosylated sample (see Figure 1) does not result in a significant broadening of the sigmoidal curve, suggesting that the heterogeneity is not a source of variation in established  $\Delta G$ s.

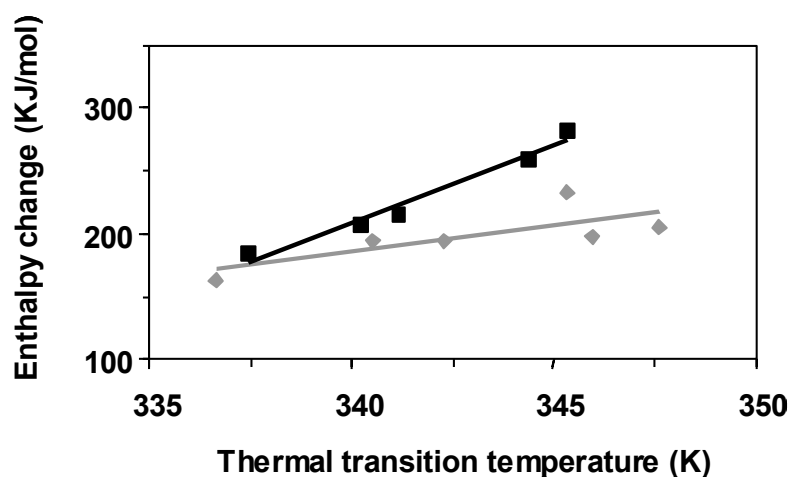
When the thermal transition temperatures of these proteins are evaluated, for example using near-UV CD (Figure 3b) by taking the first derivative of the unfolded fraction over the temperature, it can be concluded that the midpoint of unfolding is shifted from 73 to 78°C due to the glucosylation.



**Figure 3** Evaluation of protein conformational stability. a) Urea denaturation equilibrium study of non-glucosylated (black) and glucosylated (grey)  $\beta$ -LGA as monitored by the tryptophan fluorescence at 320 nm as a function of the urea concentration at 20°C. The lines represent the fits of the two-state unfolding analysis of the data. b) The CD-intensity at 293 nm in the near-UV CD spectra is monitored as function of temperature for the two proteins. Upon re-working the observed intensities in terms of unfolded fraction and taking the derivative over temperature, the curves presented in this panel are obtained.

*Heat capacity changes*

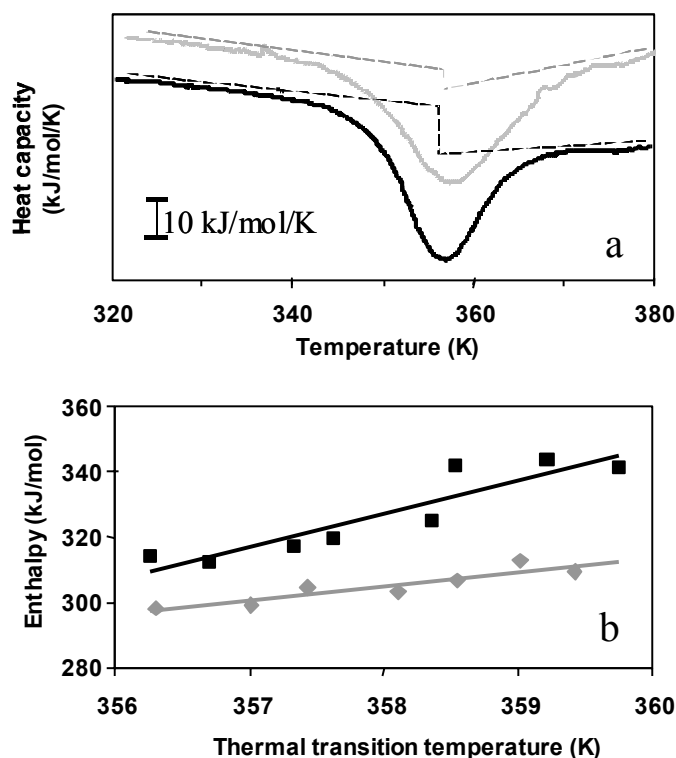
By monitoring the CD ellipticity at 293 nm as a function of temperature in the presence of various denaturant concentrations (not shown), and subsequent analysis of these traces using a two-state unfolding model as described in detail by Pace (1986), a series of enthalpic changes with their corresponding thermal transition temperature can be obtained. The results of these experiments are shown in Figure 4 for non-glucosylated and glucosylated  $\beta$ -LGA. The slope of these lines ( $d(\Delta H)/d(T_i)$ ) reflect the change in heat capacity ( $\Delta C_p$ ) upon unfolding. For non-glucosylated  $\beta$ -LGA a value for  $\Delta C_p$  of 12.4 ( $\pm 0.6$ ) kJ/mol/K is then found, while that of the glucosylated protein is 4.2 ( $\pm 0.3$ ) kJ/mol/K. Performing similar experiments by monitoring the fluorescence intensity as a function of temperature (not shown) provided comparable results (11.8 and 3.8 kJ/mol/K respectively). In literature  $\Delta C_p$ -values for folded  $\beta$ -LG between 10-13 kJ/mol/K have been reported more frequently (f.e. Griko & Privalov 1992).



**Figure 4** *Enthalpy change of non-glucosylated (black symbols) and glucosylated (grey symbols)  $\beta$ -LGA as a function of the thermal transition temperature. Data are obtained by fitting temperature traces of near-UV CD intensities at 293 nm for various denaturant concentrations (0-3.2 M urea).*

Since in the above-described procedure the local environment of the aromatic groups is monitored, and since it cannot be excluded that this provides an incomplete picture of the whole molecule, differential scanning calorimetry was employed to establish  $\Delta C_p$ . This approach is often regarded as the most direct and appropriate technique, but in cases where the thermodynamic reversibility is poor due to extensive aggregation upon unfolding at the

high protein concentrations used, as is the case for  $\beta$ -lactoglobulin, derivation of thermodynamic parameters from the shape of recorded thermograms is complex. Moreover, at high temperatures the formed Maillard-products may react further (thereby changing protein functionalities), resulting in for example browning of the materials.



**Figure 5** a) Differential scanning calorimetric profile of 20 mg/mL non-glucosylated (black line) and glucosylated (grey line)  $\beta$ -LGA in 10 mM phosphate-buffer (pH 7.0). The curves are displaced vertically for clarity. The dashed lines represent the baselines before and after unfolding. b) Plot of the enthalpy change versus the thermal transition temperature for non-glucosylated (black symbols) and glucosylated (grey symbols)  $\beta$ -LGA as obtained using DSC at pH 2.0 and a varying phosphate concentration ranging from 0 to 200 mM.

From the thermograms shown in Figure 5a it is evident that the ‘step’ in  $C_p$  at the midpoint of the enthalpy change is smaller for the glucosylated protein compared to the non-glucosylated one. Alternatively, the change in enthalpy of unfolding as a function of thermal transition temperature under various conditions (like varying pH or phosphate content, see van Teeffelen et al 2005) provides an alternative route. By variation of the phosphate concentration at pH 2.0, where the protein is monomeric but still globular compact and where possible aggregation processes do not occur (in contrast to pH 7) again a series of enthalpy

changes as a function of thermal transition temperatures can be obtained. This is shown for the non-glucosylated protein in Figure 5b. In each of the thermograms recorded (either at pH 2 or 7, where in the latter case this evaluation is sometimes hindered by an exothermic transition related to aggregation) the ‘step’ at  $T_t$  can be established. Values for the  $\Delta C_p$  of 10.7 ( $\pm 0.6$ ) and 10.1 kJ/mol/K are in this way obtained for the non-glucosylated protein using the  $d(\Delta H)/d(T_t)$  and  $C_p$ -step respectively, while for the glucosylated protein in this way values of 4.3 ( $\pm 0.5$ ) kJ/mol/K are found. It has to be noted that other enthalpic processes, like dimer-monomer dissociation or aggregation, that may coincide with the unfolding may foul the interpretation. Moreover, extensive heating of these samples eventually leads to progressive Maillard-products as indicated by browning of the material. This latter process is, however, expected to give rise to a more gradual enthalpic contribution as a function of temperature around the protein thermal transition temperature rather than a cooperative transition. In view of the agreement of the  $\Delta C_p$ -values obtained by DSC and the spectroscopic approaches, it can be concluded that these latter potential processes do not interfere with the evaluation of  $\Delta C_p$  here.

#### *Molecular interpretation of difference in $\Delta C_p$*

How to explain an about 60% decrease in  $\Delta C_p$  upon glucosylation of the protein? An increase in heat capacity upon protein unfolding originates from the ordering of polar solvent molecules around the newly exposed non-polar groups in proteins that were originally buried in the core of the folded structure. Often changes in  $C_p$  are related (with good correlation) to changes in accessible surface area ( $\Delta ASA$ ) (f.e. Baskakov & Bolen 1999). This can for example be established by determination of the Stokes’ Radii using by gel permeation chromatography (see for example Uversky et al 1993). From the elution profiles (not shown) in the absence of denaturant both proteins elute at the mass of the dimer, where the glucosylated protein appears slightly larger than the folded dimer (Table I). Angle-dependent light scattering (Broersen, unpublished results) and fluorescence correlation spectroscopy (Table I) confirmed this small (about 3%) difference in Stokes’ Radius. In the presence of 8M urea both proteins elute as monomer and their Stokes’ Radii do not differ significantly. In 8M urea the proteins display still a certain compactness, as can be expected by the presence of two disulphide bridges that are still intact in the unfolded protein, thereby limiting the conformational space in the unfolded state.



**Table I**      *The Stokes' Radii (in Ångstrom) of non-glucosylated and glucosylated  $\beta$ -LGA as determined by size exclusion chromatography and fluorescence correlation spectroscopy.*

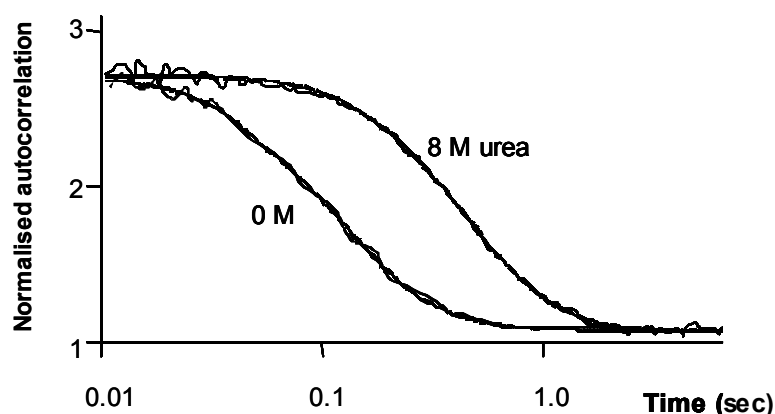
	Non-glucosylated		Glucosylated	
	GPC <sup>a)</sup>	FCS <sup>b)</sup>	GPC <sup>a)</sup>	FCS <sup>b)</sup>
0 M	27.0	30.0	27.9	30.9
8 M	35.7	36.8	36.1	36.2

a) Based on calibration of the size exclusion column using a set of standard proteins with reported Stokes' Radii (f.e. web-site Amersham).  $R^2$ -value of calibration line (elution volume versus Stokes' Radius) was 0.996. Accuracy of established radii was  $\pm 0.3$  Å.

b) The Stokes' Radius was calculated according to  $r = kT/(6\pi\eta D)$ , where  $D$  (diffusion coefficient) was calculated using  $x^2 = 2Dt$  with  $x$  being the radius of the selected volume ( $0.22 \mu\text{m}$ ) and  $t$  the observed diffusion time. The intrinsic viscosity was corrected for the contribution of 8M urea according to the standard formula given by Kawahara and Tanford (1966). The estimated accuracy of the values was  $\pm 0.4$  Å.

Another approach to evaluate the compactness of proteins is by determination of their apparent diffusion coefficient using fluorescence correlation spectroscopy. Figure 6 shows examples of the autocorrelation functions of non-glucosylated  $\beta$ -LGA in the absence and presence of 8M urea. Clearly the diffusion time is doubled in the presence of urea. From the diffusion times the Stokes' radius of the particles can be calculated as presented in Table I both for the non-glucosylated and glucosylated protein in the absence and presence of 8M urea. Compared to size exclusion chromatography the Stokes' Radii in the absence of urea are slightly higher, while those in 8 M urea are well comparable.

When a  $\Delta\text{ASA}$  of  $8.5 \text{ Å}^2$  upon unfolding for non-glucosylated  $\beta$ -LGA is assumed, the  $\Delta\text{ASA}$  of glucosylated protein could then be evaluated to be  $7.5\text{-}8 \text{ Å}^2$  (in view of the 3% larger radius at 0M urea and the similar radius at 8M urea). This might explain then a difference in  $\Delta C_p$  of about  $0.4\text{-}0.5 \text{ kJ/mol/K}$  between the two proteins, according to the equation described by Myers et al (1995). This does, however, not account for the  $6 \text{ kJ/mol/K}$  difference observed. Alternatively, when the non-polar residues that become solvent exposed upon unfolding could associate to the covalently bound sugar moieties, and thereby prevent full solvent exposure this would significantly lower the  $\Delta C_p$ . In this way the non-glucosylated and glucosylated protein might unfold to the same degree, but differ in hydrophobic exposure upon unfolding.

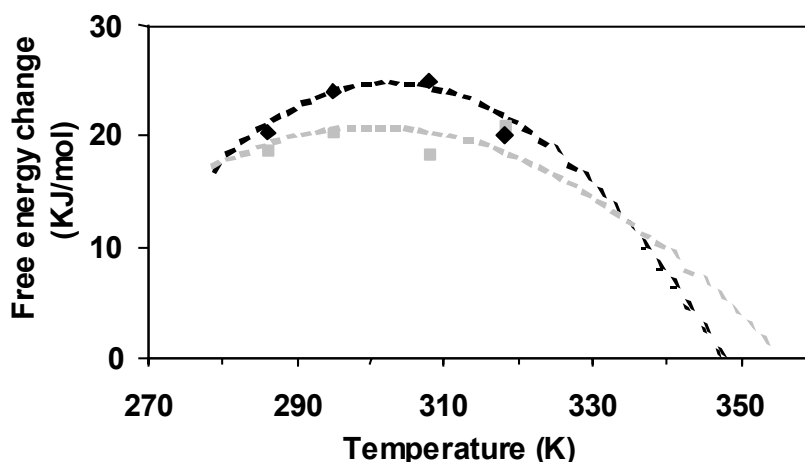


**Figure 6** Diffusion time of non-glucosylated  $\beta$ -LGA in the presence of 0 and 8M urea, as monitored by fluorescence correlation spectroscopy.

In principle, differences in  $\Delta C_p$  could also be the result of alterations in the hydration-patterns of polar groups. We believe that in this case, where glucosylation of  $\beta$ -lactoglobulin also leads to a reduced aggregation propensity (Broersen et al 2005), this is a less dominant mechanism.

#### $\Delta G(T)$ envelopes

Using the data for  $\Delta H$  and  $T_t$  from DSC and  $\Delta C_p$  evaluated as described above, the corresponding  $\Delta G(T)$  envelopes can be constructed, based on the description provided by Greene and Pace (1974). These envelopes are depicted in Figure 7 by the dashed lines for non-glucosylated and glucosylated  $\beta$ -LGA. Alternatively, values for  $\Delta G$  at individual temperatures can be derived by performing urea-titration studies at different temperatures and by monitoring intensity changes in the tryptophan fluorescence spectra (data not shown). These values are represented in Figure 7 by the symbols. Generally, the data points are close to the envelopes drawn. Fitting of the experimental data points to derive estimates for  $\Delta H$ ,  $\Delta C_p$  and  $T_t$ , as often employed in literature, did not give unambiguous results in this case. Despite elaborate attempts using multiple approaches to obtain a quantitative numbers for  $\Delta G$  at given temperatures for this system (for details see van Teeffelen et al 2005), we did not feel confident to use these data to evaluate  $\Delta C_p$  in this way. For example, the established  $\Delta G$  displayed a significant dependence on the frequency of analysis of the fluorescence spectrum.



**Figure 7** Change in Gibbs energy as a function of temperature. The dashed lines (black for non-glucosylated and grey for glucosylated  $\beta$ -LGA) are the established curves according to Greene and Pace (1974) using data obtained by DSC ( $\Delta H$  for non-glucosylated and glucosylated  $\beta$ -LGA are 380 and 340 kJ/mol, with corresponding  $T_s$  of 73 and 78 °C respectively). Values for  $\Delta C_p$  (11.0 and 4.5 kJ/mol for non-glucosylated and glucosylated  $\beta$ -LGA respectively are established as described in this work). The symbols reflect experimental values for non-glucosylated (black) and glucosylated (grey)  $\beta$ -LGA as derived from equilibrium denaturant unfolding studies carried out at different temperatures.

#### Implications of a lowered $\Delta C_p$

A lower  $\Delta C_p$  is inherent to a broader  $\Delta G(T)$ -envelop (Greene & Pace 1974). This provides the situation that at ambient temperature  $\Delta G$  might be lower, but that the protein displays a higher thermostability. Interestingly, when comparing this behaviour with other modifications on surface charges of proteins (see f.e. Kusters et al 2003), it appears that glycosylation with mono-saccharides is thus far the only alteration of functional groups on the protein surface studied that displays this effect. A lower degree of modification with oligo-sugar moieties (as studied in the past for fructose) but with the same total number of sugar moieties per protein, did not display the increased thermostability (Trofimova et al 2004), but showed a positive  $\delta(\Delta G)/\delta T_i$ . In this latter case the flexibility of the attached oligo-saccharide or the constraints on the polypeptide chain are too high to allow an efficient complexation of non-polar residues to prevent solvent exposure.

Generally, a broader  $\Delta G(T)$ -envelop implies that the protein is conformational stable in a larger temperature regime. In nature this provides a useful protection-mechanism against heat or cold-stress conditions. In technological applications this mechanism might lead to

better preservation of enzyme-activity during (heat or cold) processing. A lower aggregation propensity also offers an opportunity to direct the fraction of proteins that is still folded after processing by allowing them to refold before they are prone to aggregate.

## Conclusions

During the relatively mild condition under which Maillard products can be formed (in this study we used 5 h at 60°C, but lower temperatures can be used as well effectively (Broersen et al 2004), the  $\epsilon$ -amino group of lysine reacts with the reducing group of a sugar to form Amadori or Heyn's products via N-substituted glycosylamine. These mild conditions prevent the entering of advanced stages of Amadori and Heyn's rearrangement products that at some point result in irreversible reaction products (responsible for example for brown colouring; Röper et al 1983, Mossine et al 1994). Reversibility of these primary stages of the adducts hinder a proper identification of these products *in situ*. The nature of these post-folding modifications, however, affects in a unique way the thermodynamic properties of the protein, without changing the globular fold. The about 60% lower  $\Delta C_p$  of glucosylated  $\beta$ -LGA is believed to reflect the ability of these proteins to efficiently compensate an extensive hydration of non-polar residues upon unfolding. Therefore, coupling of small reducing sugar moieties to proteins offers nature a unique tool to deal with protein stability issues. The presence of sugar moieties may suppress aggregation by increased net charge (for proteins with an acidic isoelectric point) or increased steric interference. This work poses that the increase of exposed hydrophobicity upon unfolding can be strongly reduced by glycosylation, affecting thereby the aggregation propensity.

## Materials and methods

### Materials

Ortho-phthaldialdehyde (OPA), 3,5-dimethoxy-4-hydroxycinnamic acid, D-(+)-glucose, NaCl, NaOH, L-leucine, NaNO<sub>2</sub>, acetonitrile, trifluoroacetic acid, acrylamide, 8-anilino-1-naphtalenesulfonic acid, Na<sub>2</sub>HPO<sub>4</sub>, NaH<sub>2</sub>PO<sub>4</sub>, D(-)-fructose, and 2-(dimethyl amino) ethanethiol hydrochloride were all of analytical grade and purchased from Sigma-Aldrich. N,N-dimethyl-2-mercaptoethylammonium-chloride (DMA), disodiumtetraboratodecahydrate (borax buffer), and urea were obtained from Merck, and Coomassie Brilliant Blue R-250 from Pharmacia. All chemicals were used without further purification.

$\beta$ -lactoglobulin ( $\beta$ -LG) was isolated and purified (>98% purity) from fresh cow milk (A:B ratio 60:40) using the protocol as described by de Jongh et al (2001). Separation of the two genetic variants was performed on a Source 15Q column (Pharmacia) eluting the proteins in 20 mM bis-Tris (pH 6.0) buffer and using a 0-1M NaCl gradient in 10 column volumes, and detection of the proteins at 214 nm. The two variants displayed a separation at baseline level. The pooled fractions were dialysed extensively against demineralised water and stored at  $-40^{\circ}\text{C}$ .  $\beta$ -lactoglobulin A ( $\beta$ -LGA) had a purity >99% as demonstrated by SDS-PAGE analysis and capillary electrophoresis.

#### *Modification of $\beta$ -lactoglobulin*

500 mg  $\beta$ -LGA was dissolved in 50 mL demineralised water (10 mg/mL) and 500 mg D-glucose was added (molar ratio of primary amino groups:glucose = 1:5). The pH was adjusted to pH 8.0 with 0.2 M NaOH and the solution was mixed and lyophilised. Next, the dried mixture was incubated in a stove for 5 h at  $60^{\circ}\text{C}$  under a  $\text{NaNO}_3$  saturated atmosphere (relative humidity of 65%). After dissolving the material in demineralised water, and extensive dialysis (molecular weight cut-off 12-14 kDa) at  $4^{\circ}\text{C}$  to remove the non-reacted D-glucose, the material was lyophilised and stored at  $-20^{\circ}\text{C}$ .

#### *Analytical methods for product characterisation*

*Chromogenic OPA assay:* The degree of glycosylation (DG) is defined as the mole sugar groups reacted to a mole protein. The DG was determined indirectly by a chromogenic assay described by Kusters et al (2003). This method is based on the specific reaction between ortho-phthaldialdehyde (OPA) and free primary amino groups in proteins in the presence of 2-(dimethyl amino) ethanethiol hydrochloride (DMA), resulting in alkyl-iso-indole derivatives that show absorbency at 340 nm. The OPA-reagent was prepared by dissolving 40 mg OPA in 1 mL methanol, followed by the addition of 25 mL 0.1 M borax buffer, 200 mg DMA and 5 mL 10% SDS. The volume was adjusted to 50 mL with demineralised water. 3 mL reagents and 15  $\mu\text{L}$  0.3 mM protein solution were measured for absorbance at 340 nm in a quartz cuvette. The absorbance of the reagent itself was subtracted from the protein-containing sample. A calibration curve was obtained by adding aliquots of a 2 mM L-leucine (Sigma) solution in water to 3 mL of OPA-reagent. All measurements were performed in triplicate.

*Matrix assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF-MS):* MALDI-TOF mass spectra were acquired in a linear mode on a Voyager-DE<sup>TM</sup> RP (PerSeptive Biosystems Inc., USA), controlled by Voyager RP Software. One  $\mu\text{L}$  of a 0.3 mM  $\beta$ -LGA solution in demineralised water was mixed with 9  $\mu\text{L}$  of a 3,5-dimethoxy-4-hydroxycinnamic acid matrix (10 mg/mL of 3,5-dimethoxy—hydroxy cinnamic acid in acetonitrile/0.3% trifluoroacetic acid/demineralised water (3:1:6, v/v/v). Next, 2  $\mu\text{L}$  of the final mixture was deposited on a 100-well golden plate and air-dried at room temperature to crystallise. Subsequently, the samples were analysed at an accelerating voltage of 25 kV and an initial velocity of 350 m/sec by taking 250 shots per spectrum. The mass range ( $m/z$ ) from 2 to 25 kDa was detected. Prior to each experiment the apparatus was calibrated using a series of reference proteins. All samples were both prepared and measured in duplicate.

#### *Fluorescence experiments*

*Isothermal unfolding experiments.* A 10 M urea stock solution was prepared by dissolving 1.1g of urea per mL of phosphate buffer (0.166 mL 20mM  $\text{Na}_2\text{HPO}_4$  and 0.833mL 20mM  $\text{NaH}_2\text{PO}_4$ ; pH 7.0). A protein stock solution (0.75mg/mL in 10 mM phosphate buffer pH 7.0) was diluted with the urea-solution in phosphate buffer yielding final protein concentrations of 0.05 mg/mL with urea concentrations varying between 0 to 8M. First, the samples were incubated at 20°C for at least 15 h, a condition shown previously to be sufficient to reach an apparent equilibrium in fluorescence intensity (unpublished results). Prior to measurement the samples were incubated for at least 1h in the water bath connected to the fluorescence cell at a selected experiment temperature (between 10 and 45°C). After incubation the sample was transferred to a thermostatted 1 mL quartz cell. The fluorescence emission was recorded between 300 and 400 nm upon excitation at 295 nm using a scan speed of 100 nm/min and slit widths of 5 nm. The experiments were reproduced multiple times.

*Thermal unfolding experiments.* Thermal unfolding of samples of 0.05 mg/mL protein, containing 0-1.2M urea and 1.2-2.8 M urea of non-glucosylated  $\beta$ -LGA and glucosylated  $\beta$ -LGA respectively, were prepared as described above. The difference in the range of the urea-concentrations is related to a matching of the thermal transition temperatures of the two systems. The sample was heated from 15-100°C with a heating rate of 1°C/min. The fluorescence emission was recorded at 360 nm after excitation at 295 nm using slit widths of 5 nm. The experiments were reproduced two times.

### *Circular Dichroism*

*Spectra acquisition.* Far-UV circular dichroism (CD) spectra of 0.1 mg/mL of protein in 10 mM phosphate buffer (pH 7.0) were recorded at 20°C in the spectral range from 190 to 260 nm on a Jasco J-715 spectropolarimeter (Jasco Corporation Japan). Quartz cuvettes were used with an optical path of 0.1 cm. The spectral resolution was 0.5 nm. The scan speed was 100 nm/min and the response time was 0.125 sec with a bandwidth of 1 nm. Typically eight scans were accumulated and subsequently averaged. The spectra were corrected for the corresponding protein-free sample. Near-UV spectra were acquired in a similar way as described above using 1 mg/mL samples and an optical path of 1 cm.

*Thermal unfolding experiments.* The ellipticity of 1 mg/mL protein solutions (containing 0.4-3.2 M urea) in 10 mM phosphate buffer (pH 7.0) was monitored at 293 nm from 20-90°C with a heating-rate of 1°C/min. Quartz cuvettes were used with an optical path of 1 cm. The response time was 8 sec, and the bandwidth 1 nm. The experiments were performed in duplicate.

### *Differential scanning calorimetry*

20 mg/mL protein samples were prepared in a 10 mM phosphate buffer (pH 7.0). Prior to measurements, all samples were degassed and kept at 20°C. After 15 minutes of thermal equilibration of the 0.9 mL sample in the DSC cell (Setaram), the sample was heated up to 120°C with a heating rate of 1°C/min. Using lower heating rates did not affect the obtained results (unpublished data). The heat flow to the cell containing the protein-solution was recorded relative to that of the corresponding buffer. As baseline a scan was recorded where both cells contained buffer. The experiments were performed at least in duplicate.

### *Fluorescence correlation spectroscopy (FCS) experiments*

The FCS set-up has been described in detail elsewhere (Widengren et al 1998). For excitation of Bodipy TMR-labelled proteins (max. 1 per protein) an argon ion laser of 535 nm was used. The laser beam was focused by a lens in front of an epi-illuminated microscope, reflected by a dichroic mirror (Leitz TK580) and refocused on the image plane of a water immersion objective (Zeiss Plan-Neofluar 63 X NA 1.2). The alignment and focusing of the set-up was checked by measuring the correlation function of 0.2 nM Bodipy TMR (Molecular Probes Inc.) in 10 mM buffer (pH 7.0) prior and after each experiment. A detailed description of the theoretical background and the evaluation of the correlation function are given elsewhere

(Widengren et al 1998). The fluctuation in fluorescence intensity in time was analysed using a procedure described in detail elsewhere (Digris et al 1999).

#### *Gel permeation chromatography*

A Superdex HR 75 (10/30) column connected to an Äkta purifier (Amersham Biosciences, Sweden) was used to determine the Stokes' radii ( $R_s$ ) of proteins according to Uversky (1993). Proteins were dissolved in 10 mM phosphate (pH 7.0) and eluted in 10 mM phosphate containing 0.15 M NaCl (pH 7.0). Alternatively, the column was pre-equilibrated with 20 column volumes of the same buffer containing 8M urea and pre-incubated proteins in 8M urea (as described above) were eluted from the column in the corresponding urea-buffer (pH 7.0). Protein detection was at 214 nm. A calibration curve ( $1000/V_{\text{elution}}$  ( $\text{mL}^{-1}$ ) vs. ( $R_s$ )) was made using a series of six proteins with known Stokes' Radii yielding an  $R^2$  of 0.96.

#### **References**

- Baskakov IV, Bolen DW (1999) *Protein Sci* 8: 1314-1319
- Broersen K, Voragen AGJ, Hamer RJ, and de Jongh HHJ (2004) *Biotech Bioeng* 86: 78-87
- Broersen K, Meinders MJB, Hamer RJ, de Jongh HHJ (2005) Glycosylation affects protein unfolding/refolding mechanism and the formation of aggregation-prone intermediates. Manuscript in preparation.
- Digris AV, Skakun VV, Novikov EG, van Hoek A, Claiborne A, Visser AJWG (1999) *Eur Biophys J* 28: 526-531
- de Jongh HHJ, Gröneveld T, de Groot J (2001) *J Dairy Sci* 84: 562-571
- Endo Y, Nagai H, Watanabe Y, Ochi K, Takagi T (1992) *J Biochem (Tokyo)* 112: 700-706
- Ermonval M, Duvet S, Zonneveld D, Cacan R, Buttin G, Braakman I (2000) *Glycobiology* 10: 77-87
- Fogliano V, Monti SM, Visconti A, Randazzo G, Facchiano AM, Colonna G, Ritieni A (1998) *Biochim Biophys Acta* 1388: 295-304
- Greene RF, Pace CN (1974) *J Biol Chem* 249: 5388-5393
- Grimsley GR, Shaw KL, Fee LR, Alston RW, Huyghues-Despointes BM, Thurlkill RL, Scholtz JM, Pace CN (1999) *Protein Sci* 8: 1843-1849
- Griko YV, Privalov PL (1992) *Biochemistry* 31: 8810-8815
- Jackson SE, Fersht AR (1991) *Biochemistry* 30: 10428-10435
- Kawahara K, Tanford C (1966) *J Biol Chem* 241: 3228-3232
- Kosters HA, Broersen K, de Groot J, Simons JWFA, Wierenga PA, de Jongh HHJ (2003) *Biotech Bioeng* 84: 61-70
- Loladze VV, Ibarra-Molero B, Sanchez-Ruiz JM, Makhataдзе GI (1999) *Biochemistry* 38: 16419-16423
- Marshall JJ, Rabinowitz ML (1976) *J Biol Chem* 251:1081-1087
- Meldgaard M, Svendsen I (1994) *Microbiology* 140:159-166
- Mossine VV, Glinsky GV, Feather MS (1994) *Carbohydr Res* 262: 257-270
- Murakami MK, Nishikawa K, Hirakawa E, Murofushi H (1997) *J Biol Chem* 272: 486-489



- Myers JK, Pace CN, Scholtz JM (1995) *Protein Sci* 4: 2138-2148
- Pace CN (1986) *Methods Enzym* 131: 266-280
- Perl D, Mueller U, Heinemann U, Schmid FX (2000) *Nature Struct Biol* 7: 380-383
- Rees DC, Robertson AD (2001) *Protein Sci* 10: 1187-1194
- Röper H, Röper S, Heyns K, Meyer B (1983) *Carbohydr Res* 116: 183-195
- Sanchez-Ruiz JM, Makhatadze GI (2001) *Trends Biotechnol* 19: 132-134
- Spector S, Wang M, Carp SA, Robblee J, Hendsch ZS, Fairman R, Tidor B, Raleigh DP (2000) *Biochemistry* 39: 872-879
- Stevens VJ, Rouzer CA, Monnier VM, Cerami A (1978) *Proc Natl Acad Sci USA* 75: 2918-2922
- Trofimova D, de Jongh HHJ (2004) *Langmuir* 20: 5544-5552
- Uversky VN (1993) *Biochemistry* 32: 13288-13298
- van Teeffelen AMM, Meinders MBJ, de Jongh HHJ (2005) Pitfalls in the analysis of heat capacity changes in proteins illustrated for  $\beta$ -lactoglobulin A. Submitted.
- Venetianer A, Pirity M, Hever SA (1994) *Cell Biol Int* 18: 605-615
- Verma R, Iida H, Pardee AB (1988) *J Biol Chem* 263: 8569-8575
- Wang C, Eufemi M, Turano C, Giartosio A (1996) *Biochemistry* 35: 7299-7307
- Widengren J, Rigler R (1998) *Cell Mol Biol* 44: 857-879



# CHAPTER 5

## THE ORIGINS OF THE EFFECT OF GLUCOSYLATION ON THE AGGREGATION MECHANISM OF $\beta$ -LACTOGLOBULIN

Broersen K, Elshof M, de Groot J, Voragen AGJ, Hamer RJ, de Jongh HHJ

Manuscript in preparation

### **Abstract**

A large number of proteins *in vivo* and *in vitro* are glycosylated. It has been recognised for years that glycosylation can have serious implications on the aggregation process. In this paper we report the results of an investigation to the molecular origins behind the effect of glycosylation on aggregation of  $\beta$ -lactoglobulin. The approach suggested here studies the effect of glycosylation on the different kinetic steps involved in the aggregation process: unfolding and aggregation. It appeared that this approach can pinpoint successfully the kinetic implications of glycosylation in the aggregation process. It was found that glycosylation specifically impedes with the unfolding/refolding step of the aggregation and less with the transition of unfolded proteins into aggregates. Subsequently, two mechanisms were hypothesised to explain this effect, the first related with hydrophobicity and the second with electrostatics. It appeared that neither of these parameters sufficiently explained the effects of glycosylation on aggregation. It was suggested that a previously described effect on the rate of refolding would be primarily responsible for the observed effects, namely the tagging of specific regions of the protein molecule to direct it toward a correct conformation upon folding.

### **Keywords:**

Aggregation, glycosylation,  $\beta$ -lactoglobulin, hydrophobicity, electrostatic repulsion, unfolding/refolding

## Introduction

A common mechanism for proteins *in vivo* to cope with stress conditions is to increase the degree of glycosylation (Murakami et al 1997). Good examples of such glycosylation are found for many heat shock proteins (Venetianer et al 1994, Verma et al 1988). This response has been related to the protective mechanism of glycosylation against unwanted aggregation reactions. The aggregation preventive role was demonstrated by for example, an increased susceptibility of recombinant human erythropoietin on heat treatment upon deglycosylation of N-linked oligosaccharides (Endo et al 1992). Also, calreticulin, located in the endoplasmic reticulum, directs the way to a correct fold through interaction with glycan chains of nascent polypeptide chains and is believed through this mechanism to protect cells by preventing protein aggregation (Heal & McGivan 1998, Helenius et al 1997, Land & Braakman 2001). Posttranslational glycosylation of lens proteins has been shown in the past to induce conformational changes and unfolding of proteins in the lens resulting in aggregation and tissue deposition, leading to cataract (Crabbe 1998). These examples demonstrate that carbohydrate chains may be essential to the thermodynamic parameters of proteins and can play an important role in the occurrence of misfolding- and aggregation-related diseases. The precise molecular mechanism by which glycosylation intervenes with the aggregation reaction of proteins is not known today and is subject of study in this paper.

As both *in vivo* as *in vitro* it was found that glycosylation potentially affects the probability of proteins to self-associate, it was already recognised in the past that the preventive mechanism is likely to be concealed within the molecule. Several studies showed for a wide range of proteins that glycosylation of proteins affects for example the hydrophobicity, net charge, heat-stability, solubility and refolding rate, each determinants of the aggregation process (Kitabatake et al 1985, Wang et al 1996, Nacka et al 1998, Broersen et al 2004, 2005a). Numerous experimental data confirmed that the folding rate of proteins is enhanced by covalent attachment of sugar moieties (see f.e. Fenouillet & Jones 1995, Broersen et al 2005a, Wang et al 1996). The aggregation inhibiting effect of glycosylation has thus been directly related to the kinetic partitioning between folding and aggregation and this may provide a possible explanation of the effect of glycosylation on aggregation. Also, as the associating interactions of proteins to form aggregates are driven by hydrophobic interactions, a second mechanism by which aggregation could be inhibited is a reduced hydrophobicity of the unfolded protein. Van Teeffelen et al (2005) showed that, even though the exposed hydrophobicity of the folded protein was not affected, a decreased hydrophobicity of the unfolded, glycosylated molecule compared to a non-glycosylated counterpart was suggested

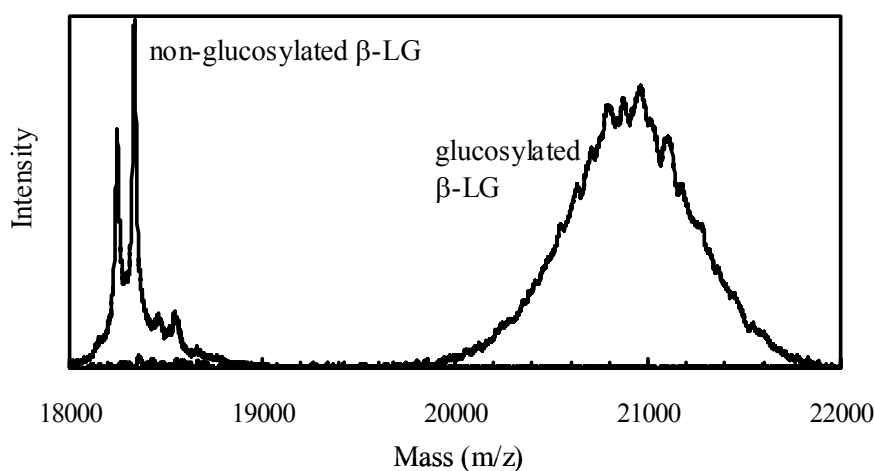
to explain the observed decreased change in heat capacity. Also electrostatics can be affected by the glycosylation reaction in a number of ways. First, the introduction of sugar chains that are actually charged results in a differently charged molecule. Second, the attachment of sugar moieties to the proteins through the Maillard reaction, reduces the number of positive charges and renders the protein with a different net negative charge. Both factors intervene with the net charge of the molecule and affect the electrostatic repulsion and thus have been thought to change the tendency of proteins to aggregate through an effect on the activation barrier for self-association. In view of the wide occurrence of glycosylation reactions *in vivo* (e.g. Wang et al 1996), the recently found relation between protein aggregation and a number of detrimental diseases (Chiti et al 2000, Gosal et al 2004) and its applicability in biotechnological and food systems (Munch et al 1999, Brands et al 2002), the elucidation of the mechanism of aggregation inhibition through glycosylation would provide potential means to understand and/or fine-tune protein aggregation reactions in all these areas.

In this work, monosaccharides were attached to the surface of bovine  $\beta$ -lactoglobulin in order to investigate the effects of glucosylation on the aggregation properties of proteins. The thermodynamic consequences of the procedure employed have been recently investigated in detail for  $\beta$ -lactoglobulin (Broersen et al 2004, 2005a). In this paper we will report our findings on the investigation of the aggregation propensity at a molecular level aiming to provide us with an explanation by which mechanism glycosylation affects the aggregation process of proteins.

## Results and Discussion

In view of numerous previously reported inhibiting effects of glycosylation on protein aggregation (Endo et al 1992, Heal & McGivan 1998, Helenius et al 1997, Land & Braakman 2001), covalently coupled sugar moieties to a protein are expected to change the aggregation propensity of proteins. Our aim was to investigate the molecular origin of the modified aggregation properties through glycosylation and therefore, to compare the aggregation kinetics of two identical proteins where one of the counterparts is glycosylated and the other counterpart is not. The glycosylation procedure employed was the Maillard reaction which has been found valuable previously to obtain glucosylated proteins with intact structural properties under selected incubation conditions, whilst obtaining a significant degree of glycosylation (Broersen et al 2004).

*Degree of glycosylation* - The efficiency of the conjugation of D-glucose with the  $\epsilon$ -amino group or N-terminal  $\alpha$ -amino group of  $\beta$ -lactoglobulin ( $\beta$ -LG), which constitutes 16 possible glycosylation sites, was tested using the OPA assay and mass spectrometry. The OPA assay determines the number of nonreacted amino groups and reflects an ensemble-averaged number. This assay indicated that  $14 \pm 0.5$  out of 16 amino groups of the glycosylated protein were occupied by glucose residues which corresponds to a degree of glycosylation of 88%. As an alternative tool, mass spectrometry was used to analyse the glycosylated  $\beta$ -LG for the degree of modification, providing additional insight in the heterogeneity of the reaction products as well (Figure 1). The spectrum of non-glycosylated protein consists of two main peaks at 18281 and 18372 Da, reflecting the two phenotypes B and A, respectively, in a molar ratio of 40:60. The small peaks detected at 18463 and 18549 Da are the result of sinnapinic acid-protein complexes as was confirmed by the use of a range of matrices for mass spectrometry (results not shown). The glycosylated  $\beta$ -LG spectrum shows a Gaussian distribution of masses, which is centred around 21087 Da (Figure 1) flanked by peaks ranging from 20229 Da to 21283 Da representative of the heterogeneous distribution of  $\beta$ -LG glycosylated to various degrees of modification.

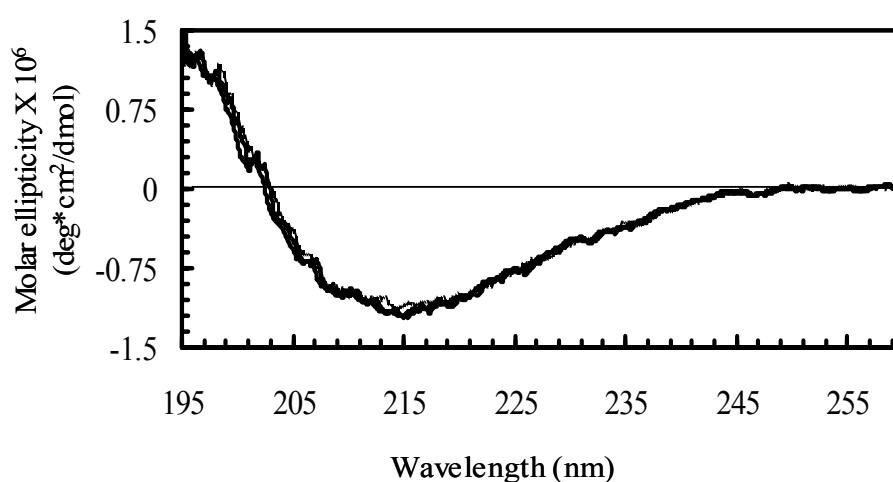


**Figure 1** *Mass (m/z) spectra of glycosylated and non-glycosylated  $\beta$ -LG.*

This heterogeneous mass distribution is consistent with previously reported results on the glycosylation of  $\beta$ -LG and is merely the result of varying association constants for the various lysines present in the primary sequence of  $\beta$ -LG (Broersen et al 2004). Since each glucose moiety attached to a free amino group accounts for 162 Da, the average number of sugar moieties attached to the protein as found with mass spectrometry is 17 indicating a

degree of glucosylation of >100%. This apparent discrepancy between the OPA assay and mass spectrometry analysis can be explained by the differences in principle of these two methods of analysis. It has been reported by Brands (2002) that the  $\epsilon$ -amino group of arginine residues can also be potential glucosylation sites. As each monomeric  $\beta$ -LG molecule contains three arginine residues in the primary sequence, it is likely that this provides additional possibilities for glucose linkage. The OPA assay specifically detects  $\epsilon$ -amino acids from lysine residues as opposed to arginine residues, while mass spectrometry is sensitive to the total molecular mass of the formed molecule irrespective of the location of the modification.

*Structural integrity* – As aggregation involves first the transition from the folded to a (partially) unfolded state, it is important to define the structural impact of glucosylation. While modifying the primary amino groups of  $\beta$ -LG, it was therefore aimed to retain the secondary, tertiary and quaternary structure of the protein. The effect of the glucosylation on the secondary structure of  $\beta$ -LG was studied by far-UV CD (Figure 2).



**Figure 2** Far-UV CD spectra of glucosylated and non-glucosylated  $\beta$ -LG at pH 7.0 in 10 mM sodium phosphate buffer at a protein concentration of 5.5  $\mu$ M. Solid line: non-glucosylated  $\beta$ -LG, dashed line: glucosylated  $\beta$ -LG.

Consistent with previous results (Broersen et al 2004), no significant differences were found upon glucosylation for the secondary structure (Figure 2). The far-UV CD spectra retained a similar zero-crossing (around 203 nm) upon glucosylation, and the shape of the spectrum showed extremes around 195, 210 and 218 nm comparable to previous results for

glycosylation of  $\beta$ -LG (Broersen et al 2004) and other reported spectra (Hirota-Nakaoka & Goto 1999).

Intrinsic tryptophan fluorescence studies (results not shown) showed a similar shape and intensity of the emission spectra as was reported before using the Maillard reaction to glycosylate  $\beta$ -LG (Broersen et al 2004). Hence, the local environment of the tryptophan residues of the glucosylated  $\beta$ -LG was considered to be the same as that in non-glucosylated  $\beta$ -LG. Also near-UV CD spectra indicated that the tertiary structure of the glucosylated protein was comparable to that of non-glucosylated protein (results not shown).

Gel permeation chromatography and native- and SDS-PAGE showed that the quaternary structure remained a non-covalently bound dimer upon glucosylation, similar to the structure of non-glucosylated  $\beta$ -LG (results not shown).

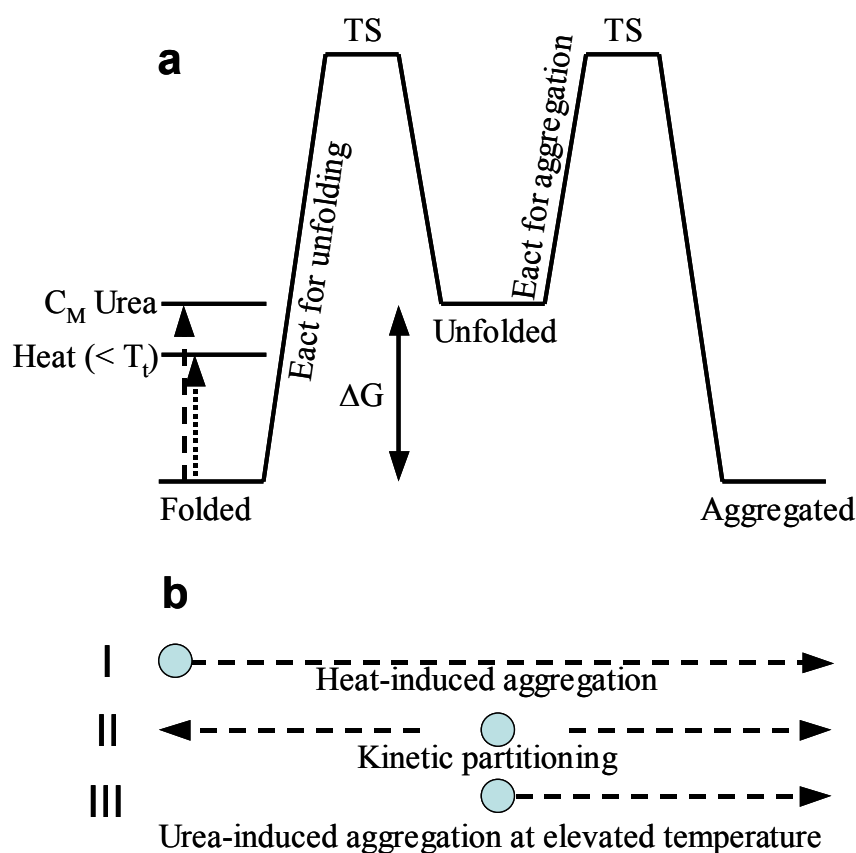
In conclusion, the structural elements of bovine  $\beta$ -LG were not affected by the glucosylation procedure and the Maillard reaction provides thus a useful tool to investigate the effect of glucosylation on the aggregation mechanics of  $\beta$ -LG without affecting the folded protein.

#### *The effect of glucosylation on the rate of protein aggregation*

Figure 3 schematically shows the thermodynamic and kinetic implications of unfolding and aggregation. The figure shows that there is a difference in thermodynamic stability ( $\Delta G$ ) between the folded and the unfolded state. For the sake of simplicity the thermodynamic stability of the folded and aggregated state are presumed to be similar. It can also be found from this figure that the process of unfolding or of aggregation requires an activation energy. The state with maximum energy is called the transition state (TS). The impact of glucosylation on the various states and phases described in Figure 3 were studied. Aggregate formation was first investigated using heat-induced aggregation. From Figure 3b (I) it can be deduced that this aggregation implies the transition from the folded to the aggregated state and thus is a suitable method to detect any differences in the energetic household of the protein upon glucosylation. Secondly, kinetic competition experiments upon refolding of unfolded species through dilution with a buffer, can be used to detect differences in the height of the transition states upon aggregation or folding (Figure 3b, II). Finally, urea-induced aggregation at elevated temperature can be used to investigate whether glucosylation has an effect on the aggregation step (unfolded to aggregated) or the unfolding step (folded to unfolded) (Figure 3b, III). In the latter case the energy level of the folded species is close to



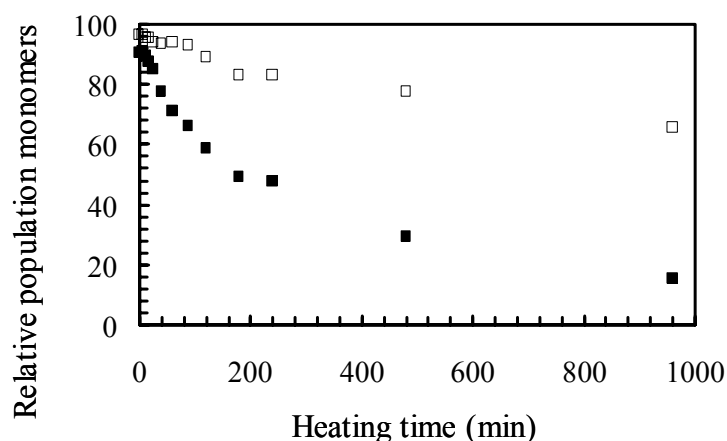
that of the unfolded one assuring that no predominant population of the folded species occurs. The aggregation of glucosylated and non-glucosylated  $\beta$ -LG were compared under these conditions to evaluate the aggregation inhibiting effect of glucosylation.



**Figure 3** a) static representation of the kinetic and thermodynamic implications of unfolding and aggregation in a situation where the folded or aggregated state are preferred, b) various modes selected to induce the aggregation processes. The circles represent the starting point of the process for each mode of aggregation. Note that for the sake of simplicity the thermodynamic stabilities of the folded and aggregated states are presumed similar and the transition states (TS) only show a static representation at ambient temperature and are also presumed similar for the processes of unfolding and aggregation. From the results of the experiments, which will be described, it becomes clear that the activation barriers actually differ upon glycosylation.

**Heat-induced aggregation** – Heat-induced aggregation (just below the thermal transition temperatures determined from DSC) was used to detect any variations through glucosylation in the thermodynamic or kinetic transitions occurring upon aggregation. The possibility of the unfolded molecule to refold to the folded conformation is basically eliminated because the energy level of the folded species is higher or equal to that of the unfolded one. Also, the

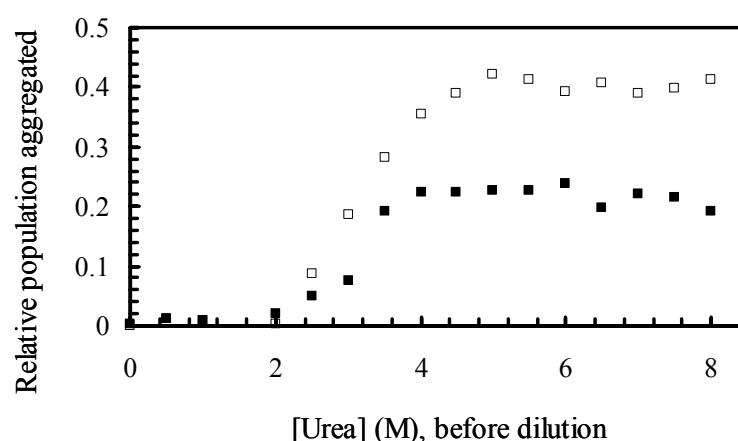
supply of unfolded molecules will be constantly exhausted as the unfolding process continues into the aggregation process. The overall equilibrium between folded and unfolded molecules will thus be constantly shifted towards the unfolded species. The process will be mainly driven toward aggregation and the rate of this process is dictated by the kinetic stability of the folded state as well as the activation barrier for aggregation. A large number of studies have investigated the effects of temperature on the aggregation process of  $\beta$ -LG (Vetri & Militello 2005, Surroca et al 2002, Carrota et al 2001, Galani & Owusu Apenten 1999, Hoffmann & van Mil 1999, Laligant et al 1991). These studies report that partial unfolding of  $\beta$ -LG is required to expose hydrophobic regions and thiol and disulphide groups. Also, the molecular mechanism behind the aggregation of  $\beta$ -LG is well-documented today and involves aggregation following from this (partial) unfolding leading to a combination of hydrophobic and disulphide bond interactions (Surroca et al 2002, Vetri & Militello 2005, Galani & Owusu Apenten 1999). Lately it has been found that hydrophobic interactions are the main driving force for  $\beta$ -LG aggregation and that disulphide bonds primarily provide a stabilising interaction for the formed aggregates (Broersen et al 2005b). Non-glucosylated and glucosylated  $\beta$ -LG were incubated at a fixed temperature below their relative thermal transition temperatures determined by differential scanning calorimetry (results not shown) but close enough to prevent significant population of refolded species. The incubation temperatures were 66°C (non-glucosylated) and 77°C (glucosylated  $\beta$ -LG). This approach to correct the rate of aggregation for heat stability has been validated before (Weijers et al 2005, Broersen et al 2005c). Previous literature describes that, upon unfolding, the dimeric  $\beta$ -LG molecule first dissociates before aggregates are formed (Galani & Owusu Apenten 2000). Figure 4 therefore shows the loss of the dissociated monomeric protein using gel permeation chromatography as a function of the incubation time using heat-induced aggregation. Clearly, the loss of the monomeric fraction of glucosylated  $\beta$ -LG is significantly faster ( $\log k = -2.37 \text{ min}^{-1}$ ) compared to the loss of the non-glucosylated protein ( $\log k = -2.76 \text{ min}^{-1}$ ) (Figure 4) (The Arrhenius equation describes the relationship between the activation energy and the logarithm of the rate of unfolding). Posttranslational glucosylation apparently does not always provide a protecting effect against aggregation as was also found before. Crabbe (1998) found that glycosylation of lens proteins induces conformational changes and unfolding of proteins in the lens and the prolonged incubation over years results in aggregation and tissue deposition, leading to cataract.



**Figure 4** Heat-induced aggregation of glucosylated and non-glucosylated  $\beta$ -LG analysed by gel permeation chromatography. The rate of disappearance of monomeric protein was analysed by comparing the rates at an ionic strength of 4.3 mM. ( $\square$ ) non-glucosylated  $\beta$ -LG ( $\blacksquare$ ) glucosylated  $\beta$ -LG.

*Aggregation through kinetic partitioning* - Figure 5 shows the formation of aggregates of 10 mg/mL glucosylated and non-glucosylated  $\beta$ -LG at pH 7.0 as a result of kinetic partitioning from the unfolded state incubated in various concentrations of urea (0-8M) and subsequent dilution in a 10 mM sodium phosphate buffer (without urea) resulting in either aggregation or refolding. Preliminary incubation studies using intrinsic tryptophan fluorescence indicated that the overnight incubation procedure at 20°C allowed sufficient time for the folding/unfolding transition to reach an equilibrium dictated by their Boltzmann distribution. Upon incubation of the proteins at a urea concentration below 2M, all molecules remained in the folded state and no aggregation occurs (Figure 5). Previous studies indicated that below 2M urea the protein molecule is structurally fully intact (Broersen et al 2005a) and thus that at 2M urea no unfolding and subsequent aggregation will take place. Above a concentration of 2M urea a sufficient concentration of protein molecules is (partially) unfolded to drive the kinetic partitioning behaviour toward aggregation in a urea concentration dependent manner. Around 4-5M urea a plateau region is reached beyond which an increasing urea concentration does not change the refolding yield anymore for non-glucosylated as well as glucosylated  $\beta$ -LG (Figure 5). The refolded fraction of proteins was tested for refolded structure using far-UV CD and urea titrations recorded with fluorescence. It was found that this refolded fraction did not differ in  $\Delta G$  and secondary structure elements from the folded  $\beta$ -LG (results not shown). From this experiment it becomes clear that the

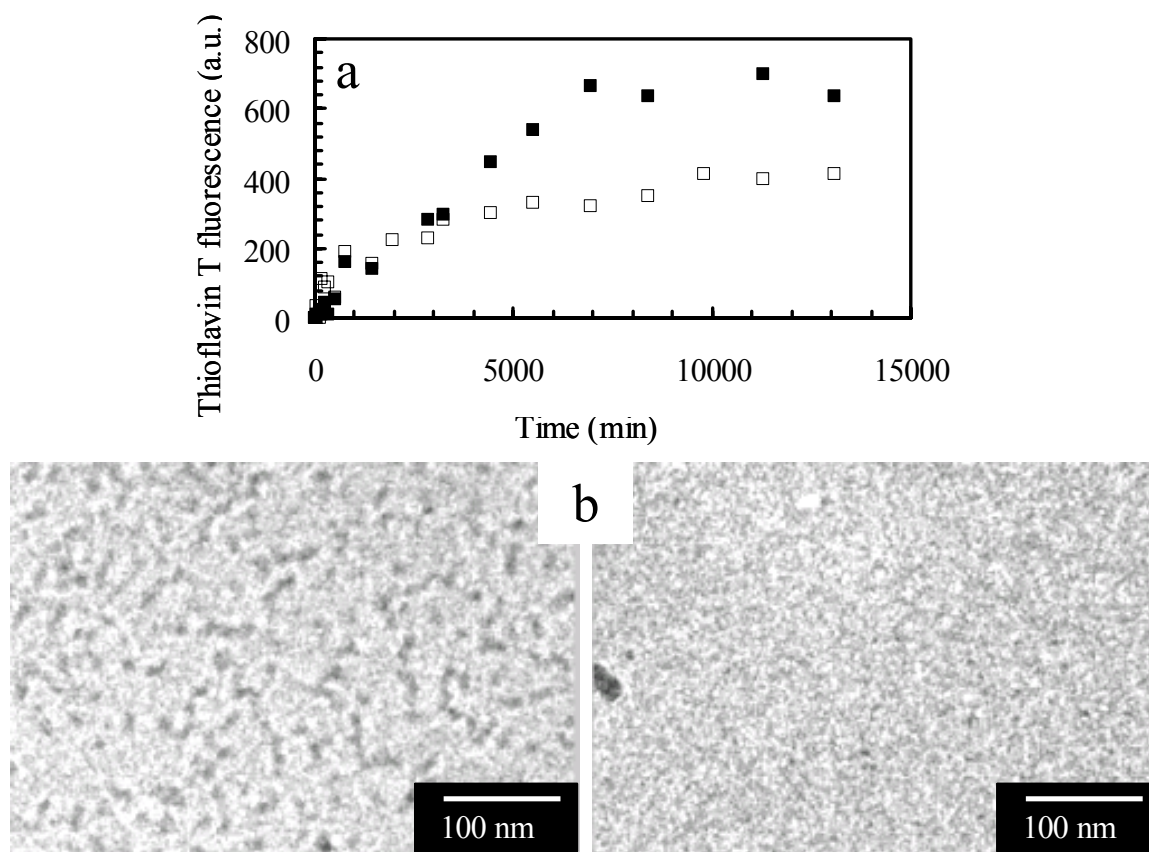
refolding yield through kinetic partitioning, under the conditions applied, is limited to approximately 57% for non-glucosylated  $\beta$ -LG compared to 80% for glucosylated  $\beta$ -LG. As was shown in Figure 3, studying aggregation through kinetic partitioning yields information regarding the relative kinetics of the refolding reaction compared to the aggregation reaction. It is clear that one or both of these terms are affected by glucosylation, but this experiment does not provide information to enable distinction between the aggregation or unfolding reaction.



**Figure 5** Urea-induced aggregate formation of glucosylated and non-glucosylated  $\beta$ -LG analysed by gel permeation chromatography. Samples were incubated in various urea-concentrations and subsequently diluted ten times with a buffer solution. The diluted samples were analysed using gel permeation chromatography. ( $\square$ ) non-glucosylated  $\beta$ -LG ( $\blacksquare$ ) glucosylated  $\beta$ -LG.

*Urea-induced aggregation through incubation at the midpoint of unfolding* – To investigate the effect of glucosylation on the activation energy for aggregation from the unfolded state, we used a method previously described by Hamada and Dobson (2002). They investigated the formation of fibrillar aggregates of  $\beta$ -LG in the presence of a urea concentration corresponding to the midpoint of unfolding using intrinsic tryptophan fluorescence. Instead of investigating the kinetic partitioning behaviour between refolding and aggregation from the unfolded state, the method described by Hamada and Dobson (2002) is based on a shift in equilibrium from a folded to a (partially) unfolded state resulting in a gradual exhaustion of the supply of folded proteins driven towards aggregation. We incubated non-glucosylated and glucosylated  $\beta$ -LG at 37°C at the concentration urea corresponding to the midpoint of unfolding (4.5 M for both non-glucosylated and glucosylated  $\beta$ -LG, Broersen et al 2005a). Plots of Thioflavin T fluorescence as a function of incubation time produce

curves that indicate growth of aggregates as suggested through the increasing Thioflavin T fluorescence over time (Figure 6a). The Thioflavin T fluorescence data also revealed that the rate of aggregation of  $\beta$ -LG was promoted by the glucosylation as glucosylation increased the Thioflavin T fluorescence faster compared to non-glucosylated  $\beta$ -LG.



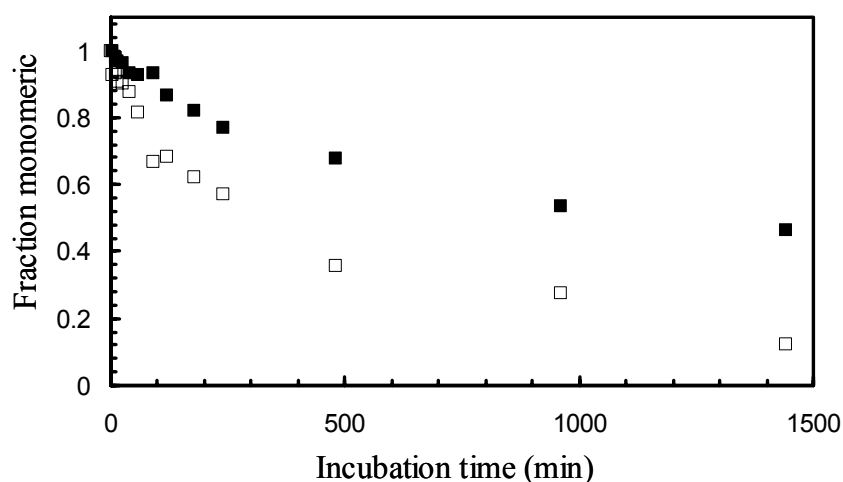
**Figure 6** Aggregate formation of glucosylated and non-glucosylated  $\beta$ -LG incubated at 4.5 M urea (midpoint of unfolding obtained through urea-induced unfolding measurements) for various time intervals and determined by different methods. a) Thioflavin T fluorescence of fibrils from glucosylated and non-glucosylated  $\beta$ -LG determined by incubation of the formed samples with Thioflavin T and evaluated using fluorescence. ( $\square$ ) non-glucosylated  $\beta$ -LG ( $\blacksquare$ ) glucosylated  $\beta$ -LG. b) Transmission electron micrographs of aggregates formed after incubation for 16 days of glucosylated (right panel) and non-glucosylated  $\beta$ -LG (left panel).

The incubated  $\beta$ -LG solutions at 37°C containing 4.5 M urea were also probed for the presence of aggregates using TEM. When the solutions were incubated for 16 days, the solutions turned opalescent, particularly in the non-glucosylated protein solution. The supramolecular structure of the aggregates were investigated using TEM (Figure 6b). The electron micrographs show that the aggregates formed in non-glucosylated  $\beta$ -LG are

unbranched, short aggregates with diameters of 4-5 nm. Remarkably, in contrast to the Thioflavin T results where upon incubation for 14000 min a clearly higher Thioflavin T fluorescence was observed for glucosylated  $\beta$ -LG compared to non-glucosylated  $\beta$ -LG, upon probing the presence of aggregates using TEM, no visible aggregates were observed. Le Vine (1995) found that the affinity of aggregates from different protein sources varied. It is possible that the glucosylated aggregates are having a higher affinity for the Thioflavin T probe compared to the aggregates formed of non-glucosylated  $\beta$ -LG.

#### *Refolding kinetics as a possible term for aggregation prevention*

The refolding efficiency of  $\beta$ -LG to the unfolded state is dictated by the rate of refolding and the rate of intermolecular association. It was found before that the rate of refolding can be increased by glycosylation (Fenouillet & Jones 1995, Broersen et al 2005a, Wang et al 1996). Also for the glucosylated  $\beta$ -LG variant used in this study it has been reported before that the attachment glucose moieties to the protein significantly increased the refolding rate as well as the refolding efficiency tested at a protein concentration of 10 mg/mL (Broersen et al 2005a for refolding rates and Figure 5 for refolding efficiency). As the refolding efficiency can significantly limit the aggregation upon kinetic partitioning experiments, we investigated the aggregation rate upon eliminating the possibility for the protein to refold. The reasoning behind this is that it was found before that glucosylation leads to a decrease of the activation energy for unfolding from 400 kJ mol<sup>-1</sup> to 175 kJ mol<sup>-1</sup> (Broersen et al 2005a). We investigated the aggregation process from a fully unfolded state, in 6 M GuHCl. It has been verified before that the protein is fully unfolded in 6M GuHCl through equilibrium studies. It may be clear that stability parameters have been significantly affected by the glucosylation. This difference obviously has implications for the transition of folded to unfolded molecules, the first step in the aggregation process. Investigation of the aggregation rate from a fully unfolded state (in 6M GuHCl) eliminates the differences in the unfolding step and isolates the effect of glucosylation on the second step of the aggregation process, the transition of molecules from an unfolded state into an aggregated state.



**Figure 7** Aggregation kinetics at 20 °C of non-glucosylated (□) and glucosylated (■)  $\beta$ -LG in 6M GuHCl.

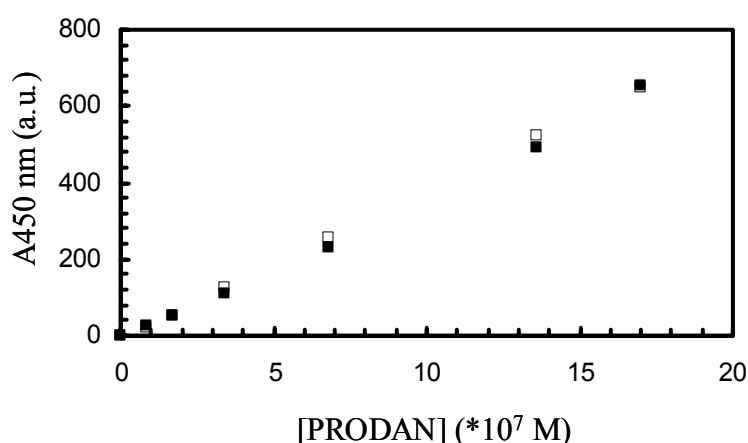
In Figure 7 the fraction monomeric molecules determined by gel permeation chromatography is plotted against the incubation time at 20°C. The aggregation rate ( $\log k$ ) for non-glucosylated  $\beta$ -LG is  $-2.55 \text{ min}^{-1}$  and for glucosylated  $\beta$ -LG is  $-2.73 \text{ min}^{-1}$ . The aggregation rates in Figure 7 are clearly not coinciding, a small difference in the aggregation rate upon glucosylation remains. It seems however that the rates of unfolding and refolding at least partially explain the reported effects of glucosylation on the aggregation rate as the differences found in Figure 7 are clearly smaller upon eliminating the effects of unfolding and refolding.

Summarising, the combined investigation of aggregation rates under various conditions suggests that the observed effect of glycosylation on protein aggregation strongly relies on the species from which the aggregation process is started. Also, the comparison of the effect of various aggregation conditions already suggests that the kinetic parameters of refolding and unfolding potentially provide an (partial) explanation for the observed effects of glucosylation on the aggregation process. To elucidate the molecular details of the effect of glucosylation on the aggregation process the effects of hydrophobicity and electrostatics on the aggregation process have been further investigated.

#### *Hydrophobicity as a possible term for aggregation prevention*

Hydrophobicity has been appointed in the past by various research groups as one of the major terms by which the aggregation process of proteins can be controlled (Laligant et al

1991, Galani & Owusu Apenten 1999, Sava et al 2005). As glycosylation has been found to readily affect the hydrophobicity of proteins, through direct interactions of the sugar chain with the hydrophobic backbone or side residues (Pieper et al 1996, Jentoft 1990, Kato 2000) this possible mechanism of aggregation prevention has been explored in the present study in various ways. Using two different methods an estimation was derived of the exposed hydrophobicity of glucosylated and non-glucosylated  $\beta$ -LG. 6-propionyl-2-(*N,N*-dimethylaminonaphthalene) PRODAN is an uncharged hydrophobic ligand that, upon binding to the hydrophobic regions of a protein forms a fluorescent complex. Plotting the change in relative fluorescence intensity of glucosylated and non-glucosylated  $\beta$ -LG at 450 nm as a function of PRODAN concentration in the range of 0 to  $1.80 \times 10^{-6}$  M suggested no significant difference in exposed hydrophobicity of folded glucosylated and non-glucosylated  $\beta$ -LG (Figure 8).



**Figure 8** Hydrophobicity of glucosylated and non-glucosylated  $\beta$ -LG evaluated using PRODAN fluorescence at pH 7.0 of non-glucosylated and glucosylated  $\beta$ -LG, ( $\square$ ) non-glucosylated  $\beta$ -LG ( $\blacksquare$ ) glucosylated  $\beta$ -LG.

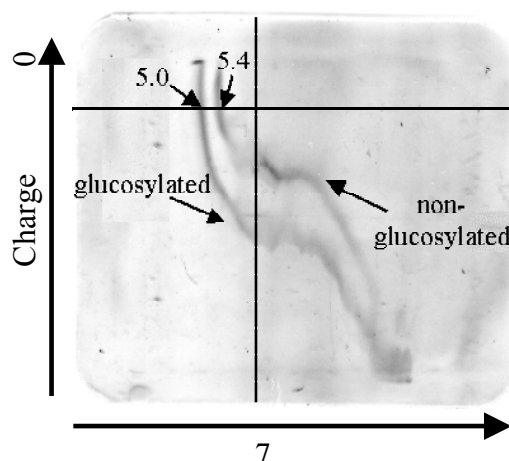
It needs to be noted that even though the PRODAN ligand has been found very useful to detect differences in hydrophobic exposure of protein molecules, which exert differences in net charge (Haskard & Li-Chan 1998), the specificity and sensitivity of the PRODAN ligand have not been unambiguously demonstrated and was therefore used in combination with Hydrophobic Interaction Chromatography. The latter method has been shown to be a valuable method before sensitive to differences in protein surface hydrophobicity in a qualitative manner (unpublished results). Non-glucosylated  $\beta$ -LG and glucosylated  $\beta$ -LG both show an



intensity maximum at an elution volume of 48 mL (results not shown). It can be concluded that glucosylation does not significantly affect the hydrophobicity of the folded state. It was found before that only very minor unfolding is required for the aggregation of glucosylated and non-glucosylated  $\beta$ -LG (Broersen et al 2005a) and it was therefore proposed that the hydrophobicity of the folded state most likely reflected the hydrophobicity of the aggregation prone molecule. To confirm that glucosylation indeed did not affect the aggregation mechanism by modification of the hydrophobicity, heat-induced aggregate formation was tested by native and SDS-PAGE. It was found that the rate of aggregation by non-covalent interactions was not affected by glucosylation (results not shown). The final proportion of non-covalently bound aggregates after heating for 960 min was not affected by the glucosylation (about 50% of the formed aggregates was non-covalently bound, results not shown) providing support for our other results that the opportunities for non-covalent aggregation are not affected by glucosylation. Further evidence was provided by viscometry experiments to determine the intrinsic viscosity of non-glucosylated and glucosylated  $\beta$ -LG at pH 7.0. The Huggins constant ( $K'$ ) calculated from the results showed an intrinsic viscosity of 0.97 mL/mg for glucosylated  $\beta$ -LG compared to 0.99 mL/mg for non-glucosylated  $\beta$ -LG. These values did not indicate significant differences suggesting that the 'stickiness' of the folded molecules was not affected by the glucosylation (results not shown).

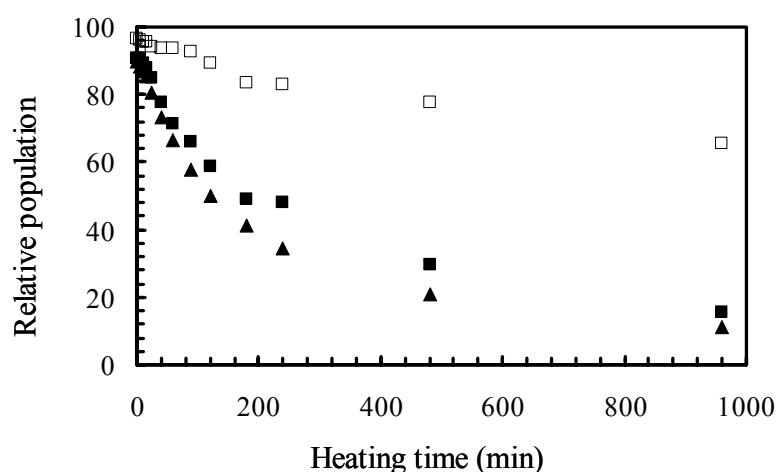
#### *Electrostatics as a possible term for aggregation prevention*

Electrostatic repulsion significantly affects the activation barrier for self-association on a long range (Lopez de la Paz 2002, Zurdo et al 2001) as well as intrinsic protein stability by e.g. stabilising ion pairs or destabilising repulsion (Frankenberg et al 1999). The isoelectric point of the protein upon glucosylation as determined by isoelectric focussing changed to 5.0 compared to 5.4 for non-glucosylated  $\beta$ -LG (Figure 9, also tested with isoelectric focussing, results not shown). However, at pH 7.0, non-glucosylated  $\beta$ -LG has a net charge of  $-7$  compared to  $-20$  upon glucosylation displaying a significant difference in net charge (calculated using Swissprot application for calculating titration curves with sequence 1BEB for non-glucosylated  $\beta$ -LG and replacing all lysines for uncharged amino acids of 1BEB for glucosylated  $\beta$ -LG).



**Figure 9** Electrophoretic titration curves of non-glycosylated and glycosylated  $\beta$ -LG

To investigate whether the difference in net charge found upon glycosylation through the Maillard reaction accounts for the observed differences in the aggregation rates, the aggregation rates were first compared at a similar ionic strength of 4.3 mM NaCl (Figure 4). A significant difference in the aggregation rate was observed. Using the calculation of Wu et al (1998), using input on the molecular radius provided by dynamic light scattering and SAXS measurements (results not shown), the electrostatic repulsion could be corrected for the differences in net charge as calculated from theoretical titration curves. 18 mM NaCl was required to correct the electrostatic repulsion of the glycosylated  $\beta$ -LG to a similar value ( $1.12 \times 10^{-31} \text{ C}^2 \text{ m}^{-1}$ ) as non-glycosylated  $\beta$ -LG. Figure 10 shows the results of the kinetics of the disappearance of the monomeric peak upon increasing incubation time by heating analysed by gel permeation chromatography.



**Figure 10** The effect of electrostatics on heat-induced aggregation of glycosylated and non-glycosylated  $\beta$ -LG analysed by gel permeation chromatography. The relative population of the

*monomer peak decreases with increasing heating time. The rates of disappearance of monomeric protein were analysed by comparing the rates at similar ionic strength of 4.3 mM: (□) non-glucosylated  $\beta$ -LG, (■) glucosylated  $\beta$ -LG or similar electrostatic repulsion of  $1.21 \cdot 10^{-31} \text{ C}^2 \text{ m}^{-1}$ : (▲) glucosylated  $\beta$ -LG.*

It may be clear that correction of the electrostatic repulsion does not result in coinciding aggregation rates for glucosylated and non-glucosylated  $\beta$ -LG. In fact, the rate of aggregation upon correction for electrostatic repulsion ( $\log k = -2.24 \text{ min}^{-1}$ ) deviates even more from the non-glucosylated protein ( $\log k = -2.76 \text{ min}^{-1}$ ) compared to the non-corrected sample ( $\log k = -2.37 \text{ min}^{-1}$ ). Thus, even though in the past strong electrostatic repulsion has been reported as a significant inhibitor of the aggregation process, the inhibiting effect of glucosylation on the aggregation reaction cannot be explained by this mechanism. Apparently, the differences in electrostatics determined by isoelectric focussing and titration curves are not responsible for the observed differences in the rate of aggregation.

## Conclusions

The results reported in this paper partially explain the previously reported apparent discrepancies in aggregation rates of glucosylated  $\beta$ -LG. The approach suggested here implies the investigation of the effect of glucosylation on the kinetics of the various steps involved in the aggregation process. It seemed that this approach can pinpoint successfully the kinetic implications of glucosylation in the aggregation process. It was found that glucosylation specifically interacts with the unfolding/refolding step of the aggregation process. This is consistent with the already reported increase in the refolding rate upon glucosylation of  $\beta$ -LG (Broersen et al 2005a). Subsequently, two hypothesised molecular mechanisms were investigated for their relative importance in the process: the effect of hydrophobicity and the effect of electrostatics. It was suggested that either a decreased hydrophobicity or an increased electrostatic repulsion, both effects of the glucosylation, would be the origin of the increased refolding yield upon glucosylation. It appeared that neither of these two parameters sufficiently explained the effects of glucosylation on aggregation. As has been reported previously, the fast refolding for glucosylated  $\beta$ -LG is probably due to the tagging of specific regions of the unfolded molecule, which should be exposed in the folded protein (Broersen et al 2005a). In other words, the covalent attachment of sugar molecules to the protein results in an efficient direction of the protein towards a correctly folded state.

## Materials and methods

### *Protein purification and modification*

*Isolation and purification* -  $\beta$ -LG was isolated and purified (>98% purity) from fresh cow milk (A:B ratio 60:40) using the protocol as described by de Jongh et al (2001).

*Preparation of glucosylated proteins* – Typically, 0.273 mM protein was dissolved in demineralised water. Per mL protein solution, 27.8 mM D-glucose (Sigma) was added. The pH was adjusted to 8.0 using 1M NaOH and the mixture was lyophilised. Next, the mixed powders were incubated at 60°C for 5 h at a relative humidity of 65% (atmosphere equilibrated with supersaturated NaNO<sub>2</sub> solution). The incubated samples were next dissolved in demineralised water and dialysed against demineralised water to remove all the non-reacted sugar from the solution. See also Broersen et al (2004) for details of this procedure. Subsequently, the solutions were lyophilised and stored at -20°C until use.

### *Protein characterisation*

*Matrix-Assisted-Laser-Desorption/Adsorption Ionisation-time-of-flight spectroscopy (MALDI-tof MS)* – MALDI-TOF mass spectra were acquired to determine the degree of modification as reported previously (Broersen et al 2004) on a Voyager-DE<sup>TM</sup> RP (PerSeptive Biosystems Inc., USA) using a 3,5-dimethoxy-4-hydroxycinnamic acid matrix.

*Chromogenic OPA Assay* – The degree of modification of the primary amino groups was determined indirectly based on the specific reaction between ortho-phthaldialdehyde (OPA) and free primary amino groups in proteins by a chromogenic assay described by Church et al (1983). All measurements were performed in duplicate and corrected for protein-free sample. The protein concentration was determined by the adsorption at 280 nm using an extinction coefficient of 0.659 M<sup>-1</sup> cm<sup>-1</sup>.

### *Structural analysis*

*Circular Dichroism (CD) spectroscopy* - Far- and near-UV CD spectra of 5.5  $\mu$ M and 0.55  $\mu$ M, respectively, glucosylated and non-glucosylated  $\beta$ -LG in a 10 mM sodium phosphate buffer (pH 7.0) were recorded on a Jasco J-715 spectropolarimeter (Jasco Corporation Japan) at 25°C in the spectral range from 260 to 190 nm (far-UV CD) or 250 to 350 nm with a spectral resolution of 0.5 nm. Spectra were recorded as averages of 16 scans. The scanning speed was 100 nm/min and the response time was 0.50 s with a bandwidth of

1.0 nm. Quartz cuvettes with an optical path of 10 mm (near-UV CD) and 1 mm (far-UV CD) were used. The spectra were corrected for the corresponding protein-free sample.

*Fluorescence spectroscopy* - Emission and excitation fluorescence spectra of glucosylated and non-glucosylated protein solutions of 2.7  $\mu\text{M}$  in a phosphate buffer (10 mM, pH 7.0) were obtained using a Cary Med Eclipse (Varian) fluorimeter at 25°C. The emission spectra from 300 to 400 nm were determined using excitation at 295 nm. Excitation and emission slit widths were 5 nm and a scan speed of 100 nm/min was used. All spectra were recorded in duplo and corrected for the corresponding protein-free sample.

*Isoelectric focusing (IEF)* – The apparent isoelectric points of the glucosylated and non-glucosylated  $\beta$ -LG were determined using the Phast System (Pharmacia). 4  $\mu\text{L}$  of 1.0 mg/mL protein solutions were applied to IEF gels with a pH-gradient ranging from 2.5 to 6.5 (Pharmacia) and from 3 to 10 (Pharmacia). The gels were fixed with 20% trichloric acid, stained using Coomassie Brilliant blue (R-250) and destained in 30% methanol/10% acetic acid.

*Electrophoretic titration (ETC)* – Glucosylated and non-glucosylated  $\beta$ -LG were dissolved at a protein concentration of 54.6  $\mu\text{M}$  in demineralised water. A pH gradient was generated on a ready-to-use IEF gel (3-9, Amersham) on a PhastSystem using 2000 V, 2.0 mA, 3.5 W and 75 Vh at 15°C. After rotating the gel 90° for the second dimension, the samples were applied to the gel and the gel separated the fractions at the pH range at 200V, 2.0 mA, 3.5 W for 15 Vh at 15°C. The gels were fixed with 20% trichloric acid, washed with destain solution (30% methanol, 10% acetic acid v/v), stained with 0.02% Coomassie brilliant blue R 350 (10% 0.2% Coomassie brilliant blue R350 (Pharmacia) in 90% destain solution) and subsequently destained with destain solution.

*Dynamic light scattering* – Dynamic light scattering measurements were performed using an ALV-5000 multi-bit multi-tau correlator and a Spectra Physics solid state laser operating with vertically polarised light with wave length  $\lambda$  532 nm. The range of scattering wave vectors covered  $3.0 \times 10^{-3} < q < 3.5 \times 10^{-2} \text{ nm}^{-1}$  ( $q = 4\pi n_s \sin(\theta/2)/\lambda$ , where  $\theta$  is the angle of observation and  $n_s$  is the refractive index of the solution). The temperature was controlled by a thermostat bath at  $20 \pm 0.1^\circ\text{C}$ .

*Hydrophobic Interaction Chromatography (HIC)* – Glucosylated and non-glucosylated  $\beta$ -LG were dissolved at a concentration of 0.16 mM in 1M ammonium sulphate and 10 mM sodium phosphate buffer (pH 7.0). 500  $\mu\text{L}$  sample was applied to a 5 mL HiTrap PhenylFF column (Pharmacia) using an Akta Purifier (Amersham) and unbound sample was

washed off with 5 column volumes of 1M ammonium sulphate and 10 mM sodium phosphate buffer (pH 7.0). The bound protein was eluted from the column using a linear salt gradient from 100% 1M ammonium sulphate and 10 mM sodium phosphate buffer (pH 7.0) to 100% 10 mM sodium phosphate buffer (pH 7.0) at a flow rate of 2.0 mL/min and detection of the protein at 280 nm. Subsequently, the column was washed with 5 column volumes of 10 mM sodium phosphate buffer (pH 7.0).

*Intrinsic viscosity* – The flow time of a filtered solution of glucosylated and non-glucosylated  $\beta$ -LG solutions in 10 mM sodium phosphate buffer (pH 7.0) were measured at  $20 \pm 0.1^\circ\text{C}$  using an Ubbelohde capillary viscometer. The relative viscosity ( $\eta_{\text{rel}}$  = flow time of sample/flow time of solvent) and specific viscosity ( $\eta_{\text{sp}} = \eta_{\text{rel}} - 1$ ) were calculated using the viscosity of a 10mM sodium phosphate buffer (pH 7.0) as solvent. Reduced viscosity ( $\eta_{\text{red}} = \eta_{\text{sp}}/\text{sample concentration}$ ) was plotted against the sample concentration which was varied from 0 to 10 mg/mL protein, and extrapolated to zero concentration using the Martin equation (Huggins 1942) to determine the intrinsic viscosity ( $[\eta]$ , mL/g). The so-called Huggins constant ( $K'$ ) is a characteristic of a given solute-solvent system (Huggins 1942) and can be calculated from the slope divided by  $[\eta]$  (Sakai 1968).

*PRODAN assay* - A stock solution of  $4 \times 10^{-4}\text{M}$  PRODAN (6-propionyl-2-(*N,N*-dimethylamino)-naphthalene) (Fluka) was prepared in methanol and protected from light. The concentration of PRODAN was determined spectrophotometrically at 360 nm, using a molar absorption coefficient of  $1.8 \times 10^4 \text{ M}^{-1} \cdot \text{cm}^{-1}$ . 2.5  $\mu\text{L}$  PRODAN solution (or methanol: blank) was added to 1 mL of protein sample (containing 0-0.44mM protein). Samples were incubated for 15 min in the dark and the fluorescence intensity was measured at  $20^\circ\text{C}$  using a Varian Cary Eclipse Fluorescence Spectrophotometer at an excitation wavelength of 365 nm, an emission wavelength range of 400-550 nm and a slit width of 5 nm. The fluorescence intensities measured were corrected for a blank containing protein and methanol. The initial slope ( $S_0$ ) of the net fluorescence intensity versus protein concentration (%) plot was calculated by linear regression analysis and used as a value to express the protein surface hydrophobicity.

### *Unfolding and stability*

*Change in free energy by urea titration* – Emission fluorescence spectra were recorded using a Cary Med Eclipse (Varian) fluorimeter at a protein concentration of 2.7  $\mu\text{M}$  in a phosphate buffer (10 mM, pH 7.0). The samples were equilibrated overnight at a range of 0 –

8 M urea at 25°C in a 10 mM phosphate buffer pH 7.0. The fluorescence emission was recorded between 300 and 400 nm upon excitation at 295 nm at a scan speed of 100 nm/min and an excitation and emission slit width of 5 nm. Denaturant titration curves were obtained by plotting the band intensity (selected at the maximum intensity of the native protein emission spectrum) for a fixed emission wavelength as a function of denaturant concentration. The change in free energy ( $\Delta G$ ) as described previously (Broersen et al 2005a).

*Differential Scanning Calorimetry (DSC)* – 0.22 mM protein solutions in 10 mM phosphate buffer pH 7.0 were sealed in the cell of a VP-DSC MicroCalorimeter (MicroCal Inc., Northampton USA). A 10 mM phosphate buffer was used as reference sample. The heat flow was recorded in duplicate from 25 to 110 °C at a heating rate of 1°C/min. The data were analysed using the MicroCal Origin software.

#### *Aggregation kinetics*

Various methods to determine the aggregation kinetics of the glucosylated and non-glucosylated  $\beta$ -LG were employed.

*Urea-induced denaturation at various urea concentrations* - Proteins were incubated at a range of urea concentrations (0 – 8M) at pH 7.0 for 24 h at 25°C. Then the protein/denaturant mixtures were diluted tenfold in 10 mM phosphate buffer (pH 7.0). To quantify the refolding efficiency gel permeation chromatography (GPC) was performed on the diluted samples. The quaternary structure of non-glucosylated and glucosylated refolded/aggregated protein in a 10 mM phosphate buffer (pH 7.0) containing 50 mM NaCl at 25°C was analysed by GPC using a 24 mL Superdex 75 column HR 10/30 (Amersham) at a flow rate of 0.4 mL/min and detection of protein in the eluent at 280 nm. The monomeric and aggregated fractions were collected separately. The refolded fraction of the proteins was obtained and tested for structural elements using far-UV CD and urea titrations with fluorescence as described above.

*Urea-induced aggregation at various time intervals* – Aggregates were formed by incubation of 10 mg/mL protein solutions in 4.5 M urea in 10 mM sodium phosphate buffer (pH 7.0) in autoclaved 1.5 mL eppendorf tubes of which the lids were tightly sealed to prevent evaporation. The samples were placed in an eppendorf incubator at 37°C. For each incubation time, varying from 0 to 15000 mins, 10  $\mu$ L aliquots were removed and rapidly cooled in liquid nitrogen.

*Heat-induced aggregation* – Heat-induced aggregates of glucosylated and non-glucosylated  $\beta$ -LG were formed at 7°C below their thermal transition temperatures at pH 7.0 and a protein concentration of 0.5 mM and similar ionic strength of 4.2 mM or similar electrostatic repulsion of  $1.21 \times 10^{-31} \text{ C}^2 \text{ m}^{-1}$  corrected with 1M NaCl. Samples were collected at various time intervals between 0 and 960 min and rapidly cooled in ice. Subsequently, the samples were stored at 4°C until further analysis.

*SDS-PAGE* – Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to a protocol described by Leammli (1970). A 15% (w/v) acrylamide separating gel and a 4% (w/v) acrylamide stacking gel containing 0.1% SDS (w/v) were run using a Mini-PROTEAN II Electrophoresis Cell (Biorad). Samples of 0.1 mM protein were prepared in sample buffer containing 10% SDS (w/v) and 1.25%  $\beta$ -mercaptoethanol (w/v). Gels were stained with Coomassie Brilliant Blue R-250 and destained with 30% methanol/10% acetic acid in water (v/v). The fractions of monomeric and polymeric material were evaluated using densitometric analysis.

*Gel permeation chromatography (GPC)* – Aggregates formed by heat- or urea-induced aggregation of glucosylated and non-glucosylated  $\beta$ -LG were dissolved in 10 mM sodium phosphate buffer (pH 7.0) and 0.15M NaCl at a concentration of 1 mg/mL. Upon application of 100  $\mu$ L protein solution to the column (25 mL Superdex 75, Amersham), the protein was eluted from the column at a flow rate was 0.75 mL/min and the eluted fractions were detected at 280 nm. All measurements were performed in duplo.

*Transmission electron microscopy* – 1.1 mM protein solutions were heated for 60 min at 90°C in a 10 mM phosphate buffer pH 7.0 in the absence and presence of 0.15 M NaCl and subsequently cooled to 20°C. Then, the solutions were diluted to 0.27 mM and placed on a perforated carbon film, supported by a 200 mesh copper grid and subsequently blotted. The liquid film was vitrified by rapidly plunging the grid into liquid propane at a temperature of -170°C. The specimens were stored in liquid nitrogen until transfer into the cryo-holder. The images were recorded digitally using a Philips CM12 transmission electron microscope operating at 80 kV with a Gatan 791 CCD camera using Digital Micrograph software.

*Thioflavin T assay* – Solutions of 10 mg/mL protein in 10 mM sodium phosphate buffer (pH 7.0) and various concentrations of urea were prepared in 1.5 mL eppendorf tubes (autoclaved before use) of which the lids were tightly sealed to prevent evaporation. The samples were placed in an eppendorf incubator at 37°C. For each incubation time, varying from 0 to 15000 mins, 10  $\mu$ L aliquots were taken and mixed with 50  $\mu$ L 250  $\mu$ M Thioflavin T



in 10 mM sodium phosphate buffer (pH 7.0) and 940  $\mu$ L 10 mM sodium phosphate buffer (pH 7.0). The fluorescence at 440-600 nm was monitored using 10 mm cuvettes. The excitation wavelength was 440 nm.

### Acknowledgements

The authors acknowledge the assistance of Mireille Weijers, Igor Dolbnya and Wim Bras for their technical assistance at the DUBBLE (Grenoble, France) and the Netherlands Organization for the Advancement of Research (NWO) for providing financial support for performing measurements at the DUBBLE. Also we acknowledge Esther Kiezebrink for conducting the Ubbelohde experiments and Jan Klok for assistance with the dynamic light scattering experiments.

### References

- Brands SMJ (2002) Kinetic modelling of the Maillard reaction between proteins and sugars. Thesis of Wageningen University, Wageningen, The Netherlands
- Broersen K, Voragen AGJ, Hamer RJ, de Jongh HHJ (2004) *Biotech Bioeng* 86: 78-87
- Broersen K, Meinders MBJ, Hamer RJ, Voragen AGJ, de Jongh HHJ (2005a) Glycosylation affects protein unfolding/refolding mechanism and the formation of aggregation-prone intermediates. Manuscript in preparation.
- Broersen K, van Teeffelen A, Vries A, Voragen AGJ, Hamer RJ, de Jongh HHJ (2005b) Do sulfhydryl groups affect aggregation and gelation properties of ovalbumin? Manuscript in preparation.
- Broersen K, Weijers M, de Groot J, Hamer RJ, de Jongh HHJ (2005c) Electrostatics controls fibril formation – Part I: Charge engineering of proteins affects unfolding. Manuscript in preparation.
- Carrotta R, Bauer R, Waninge R, Rischel C (2001) *Protein Sci* 10: 1312-1318
- Chiti F, Taddei N, Bucciantini M, White P, Ramponi G, Dobson CM (2000) *EMBO J* 19: 1441-1449
- Church FC, Swaisgood HE, Porter DH, Catignani GL (1983) *J Dairy Sci* 66: 1219-1227
- Crabbe MJ (1998) *Cell Mol Biol (Noisy-le-Grand)* 44: 1047-1050
- de Jongh HHJ, Gröneveld T, de Groot J (2001) *J Dairy Sci* 84: 562-571
- Endo Y, Nagai H, Watanabe Y, Ochi K, Takagi T (1992) *J Biochem (Tokyo)* 112: 700-706
- Fenouillet E, Jones IM (1995) *J Gen Virol* 76: 1509-1514
- Frankenberg N, Welker C, Jaenicke R (1999) *FEBS Lett* 454: 299-302
- Galani D, Owusu Apenten RK (1999) *Int J Food Sci Tech* 34: 467-476
- Galani D, Owusu Apenten RK (2000) *Thermochim Acta* 363: 137-142
- Gosal WS, Clark AH, Ross-Murphy SB (2004) *Biomacromolecules* 5: 2408-2419
- Hamada D, Dobson CM (2002) *Protein Sci* 11: 2417-2426
- Haskard CA, Li-Chan ECY (1998) *J Agric Food Chem* 46: 2671-2677
- Heal R, McGivan J (1998) *Biochem J* 329: 389-394
- Helenius A, Trombetta ES, Hebert DN, Simons JF (1997) *Trends Cell Biol* 7: 193-200

- Hirota-Nakaoka N, Goto Y (1999) *Bioorg Med Chem* 7: 67-73
- Huggins M (1942) *J Am Chem Soc* 64: 2716-2718
- Jentoft N (1990) *Trends Biochem Sci* 15: 291-294
- Kato A, Nakamura S, Ban M, Azakami H, Yutani K (2000) *Biochim Biophys Acta* 1481: 88-96
- Kitabatake N, Cuq JL, Cheftel JC (1985) *J Agric Food Chem* 33: 125-130
- Laligant A, Dumay E, Valencia CC, Cuq J-L, Cheftel J-C (1991) *J Agric Food Chem* 39: 2147-2155
- Land A, Braakman I (2001) *Biochimie* 83: 783-790
- Leammler UK (1970) *Nature* 227: 680-685
- Levine H (1995) *Amyloid: Int J Exp Clin Invest* 2: 1-6
- Lopez de la Paz M, Goldie K, Zurdo J, Lacroix E, Dobson CM, Hoenger A, Serrano L (2002) *Proc Natl Acad Sci USA* 99: 16052-16057
- Munch G, Schick Tanz D, Behme A, Gerlach M, Riederer P, Palm D, Schinzel R (1999) *Nat Biotech* 17: 1006-1010
- Murakami MK, Nishikawa K, Hirakawa E, Murofushi H (1997) *J Biol Chem* 272: 486-489
- Nacka F, Chobert JM, Burova T, Leonil J, Haertle T (1998) *J Protein Chem* 17: 495-503
- Pieper J, Ott K-H, Meyer B (1996) *Nat Struct Biol* 3: 228-232
- Sakai T (1968) *J Polymer Sci* 6: 1659-1672
- Sava N, van der Plancken I, Claeys W, Hendrickx M (2005) *J Dairy Sci* 88: 1646-1653
- Surroca Y, Haverkamp J, Heck AJR (2002) *J Chrom A* 970: 275-285
- van Teeffelen AMM, Broersen K, de Jongh HHJ (2005) Glycosylation of  $\beta$ -lactoglobulin lowers the heat capacity change of unfolding: a unique way to affect protein thermodynamics. Accepted for publication in *Protein Sci*.
- Venetianer A, Purity M, Hever SA (1994) *Cell Biol Int* 18: 605-615
- Verma R, Iida H, Pardee AB (1988) *J Biol Chem* 263: 8569-8575
- Vetri V, Militello V (2005) *Biophys Chem* 113: 83-91
- Wang C, Eufemi M, Turano C, Giartosio A (1996) *Biochemistry* 35: 7299-7307
- Weijers M, Broersen K, Barneveld PA, Cohen Stuart MA, Hamer RJ, de Jongh HHJ, Visschers RW (2005) Electrostatics controls fibril formation – Part II: Net charge can affect fibril morphology and gel properties of ovalbumin. Manuscript in preparation.
- Wu J, Bratko D, Prausnitz JM (1998) *Proc Natl Acad Sci USA* 95: 15196-15172
- Zurdo J, Gujjarro JJ, Jiménez JL, Saibil HR, Dobson CM (2001) *J Mol Biol* 311: 325-340

## **PART II**

### **ELECTROSTATICS AND PROTEIN STABILITY**



# CHAPTER 6

## ELECTROSTATICS CONTROLS FIBRIL FORMATION. PART I: CHARGE ENGINEERING OF PROTEINS AFFECTS UNFOLDING

Broersen K, Weijers M, de Groot J, Hamer RJ, de Jongh HHJ

Manuscript in preparation

### **Abstract**

This study aims to describe how random charge modification affects the unfolding process of ovalbumin in the context of aggregate formation. To this end, a series of ovalbumin variants were produced that were chemically engineered to vary their net charges. Heat- and denaturant-induced unfolding studies suggested that the protein deals with charge modification by varying the stability. It was found that charge modification significantly affected the thermal transition temperature as well as the Gibbs free energy. The unfolding processes of all the variants were irreversible, without exception. A two-state irreversible unfolding transition model accurately describes the heat-induced unfolding process of all the engineered proteins as was shown by differential scanning calorimetry studies. Only an ovalbumin variant with a net charge of  $-1$  deviated from this behaviour and formed thermodynamically stable intermediates. These intermediates were also observed in equilibrium unfolding studies using urea-induced unfolding. It is possible that charges that normally destabilise the intermediate and prevent accumulation by electrostatic repulsion were eliminated in the ovalbumin variant with a net charge of  $-1$ . Population of these unfolding intermediates and subsequent refolding led to the formation of an aggregated fraction, suggesting a kinetic preference for aggregation over refolding. Other aggregation related physical and chemical parameters were therefore investigated and it was found that hydrophobicity or disulphide bond formation were not affected by charge modification. Our results lead to the conclusion that the thermodynamic properties of ovalbumin were significantly affected by the charge modification but that the possibility to form aggregation-prone molecules was by no means reduced. The observed differences in the rate of unfolding upon charge modification can have important consequences for the kinetics of aggregation.

### **Keywords:**

Ovalbumin, net charge, unfolding, stability, intermediate, aggregation

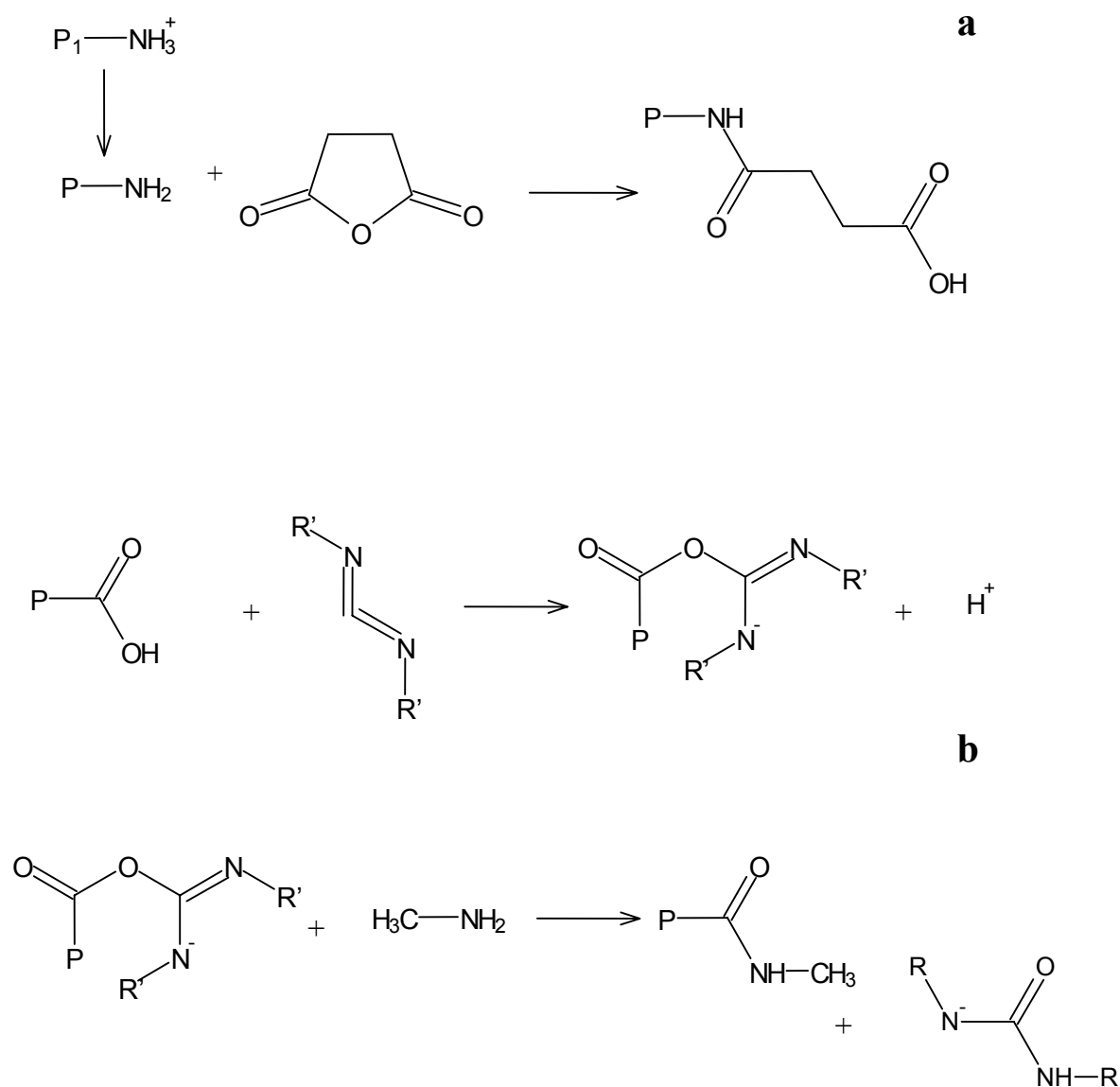
## Introduction

The process of protein aggregation enjoys a shared interest from a (bio-) technological point of view (Sun et al 2002) as well as that it has been associated with a number of amyloid-related diseases (Narang 1980, Merz et al 1983). Within this perspective it is important to improve our current understanding of the background of protein aggregation. Factors determining the aggregation process have been investigated in detail during the last decade (Kelly 1998, Lansbury 1999, Uversky & Fink 2004, Fandrich & Dobson 2001, Hwang et al 2004, Malisauskas et al 2003, de la Paz et al 2004, Kitabatake 1995). One important determinant for aggregation that has been recently identified is the electrostatic contribution (Lopez de la Paz et al 2002, Chiti et al 2002, 2003, Zurdo et al 2001). Upon varying charge interactions of (partly) unfolded molecules by changing the solvent conditions, the propensity of polypeptides to form aggregates can be varied.

Even though the contribution of electrostatics has been strongly related to aggregate formation, a clue as to the mechanism underlying this relation is still missing. Aggregation requires first (partial) unfolding of the polypeptide molecule, and second, the ability of these molecules to approach each other close enough to enable intermolecular bonding. It can be questioned which of these two steps is affected by varying protein charge. Zurdo et al (2001) for example, showed that fibrils for many proteins are formed only at a pH that promotes partial unfolding of the protein to expose hydrophobic residues (see also Guijarro et al 1998, Villegas et al 2000, Frankenberg et al 1999). On the other hand, ionic strength has been found to affect the rate of aggregation by modulating the repulsion effect (e.g. Pots et al 1999). The effect of charge, by charge engineering of ovalbumin, on the rate of aggregation and aggregate morphology is discussed by Weijers et al (2005).

The required (partial) unfolding of the polypeptide chain is governed by the conformational stability of the molecule. A large difference in free energy of the folded and the unfolded polypeptide chain as well as a high kinetic barrier to unfolding will limit the generation of aggregation prone protein molecules. It has been reported before that electrostatic repulsion between particles acts as a prime determinant for the rate of aggregation (Nakamura et al 1978). Affecting unfolding kinetics and thus the availability of aggregation-prone molecules by introducing or eliminating charged residues in a protein molecule could provide an alternative explanation for the observed effect of electrostatics on aggregation. Also, nature has used this approach during evolution as becomes evident when comparing various thermophilic and mesophilic proteins. Among these proteins, the thermodynamic stability is modulated by varying the close-range electrostatic interactions.

Also, work carried out by the group of Sanchez-Ruiz confirms that optimisation of charge-charge interactions on the surface of proteins can be a useful strategy in the design of thermostable proteins with a high resistance to unfolding (Loladze et al 1999). It appears thus very likely that, apart from affecting the rate at which (partially) unfolded protein molecules can approach each other, also the rate at which (partially) unfolded molecules are formed, will be affected by charge modification.



**Figure 1** a) Succinylation reaction,  $P$  = protein, b) Methylation reaction,  $P$  = protein

Modification of charge-charge interactions of a protein can be achieved in two ways. With point mutations the newly expressed amino acid sequence is allowed to fold in an

expression system after the modification. However, this approach often results in alternative folding of the protein compared to the wild type protein. Random chemical modification (such as succinylation and methylation, Figure 1) uses the folded conformation of the protein as a starting point and adds new properties. This approach of chemical modification has been followed by many researchers in the past by for example Gounaris and Perlmann (1967) and Nakagawa et al (1972).

The model system used in this study is ovalbumin, a well-characterised serpin with a molecular size of 45 kDa that is known to form aggregates under various conditions (Koseki et al 1989, Veerman et al 2003). Ovalbumin is widely used in the food industry and its unfolding and aggregation behaviour are well documented (e.g. Weijers et al 2003). In this work, the charge of ovalbumin has been chemically modified using methylation and succinylation reactions and the impact of this on its unfolding behaviour is documented in this paper.

## Results

*Preparation and selection of materials* - The succinylation (Figure 1a) and methylation reaction (Figure 1b) were used to introduce or to eliminate charges in ovalbumin, respectively. With these materials we aim to investigate the effect of charge on the thermodynamic properties of proteins related to the generation of molecules that can undergo aggregation as well as some key parameters related to protein aggregation such as hydrophobicity and sulfhydryl-disulphide exchange reactions. To obtain a range of differently charged ovalbumins, parameters, such as incubation time, temperature and reactant: protein ratio can be varied as mentioned in the Materials and Methods section. These variants are OVA-26 to OVA-1 in which the number denotes the net charge as calculated from OPA and Woodward experiments (Table I). It is important to retain an intact globular structure after the modification to allow comparison of stability and aggregation properties. We therefore evaluated the effect of charge modification on the structural properties. For this, the produced ovalbumin variants were tested for their secondary and tertiary structure using intrinsic tryptophan fluorescence (Figure 2a), near-UV circular dichroism (CD) (results not shown) and far-UV CD (Figure 2b). It was found that there are limitations as to which extent the degree of modification can be varied without affecting the structural properties of the molecules (results not shown). As shown in Figure 2a, all selected OVA variants showed a  $\lambda_{\text{max}}$  around 340 nm, comparable to the wild type ovalbumin (OVA-12). The results of the



molecular characterisation are summarised in Table I. The four selected variants, ranging in net charge from  $-1$  to  $-26$  (OVA-1, OVA-5, OVA-12 and OVA-26), had secondary (Figure 2b), tertiary (Figure 2a) and quaternary (gel permeation chromatography, results not shown) structures comparable to OVA-12.

**Table I**      *Modification and electrostatic description of selected ovalbumin variants*

Ovalbumin variant	# groups available/molecule			Charge/molecule at pH 7.0 <sup>d</sup>	Isoelectric point
	Carboxyl <sup>a</sup> groups/molecule	Lysine <sup>b</sup> groups/molecule	Number of modified groups <sup>c</sup>	Calculated	
OVA-26	$58 \pm 2$	$15 \pm 1$	$7 \pm 1$	$-26 \pm 2$	3.5
OVA-12	$50(48)^* \pm 2$	$21(21)^* \pm 1$	-	-12	4.7
OVA-5	$43 \pm 2$	$22 \pm 1$	$7 \pm 2$	$-5 \pm 2$	5.0
OVA-1	$39 \pm 2$	$22 \pm 1$	$11 \pm 2$	$-1 \pm 2$	5-7

\*Between brackets: theoretical values of carboxyl groups and lysine groups derived from the primary sequence of ovalbumin (PDB-code 1OVA).

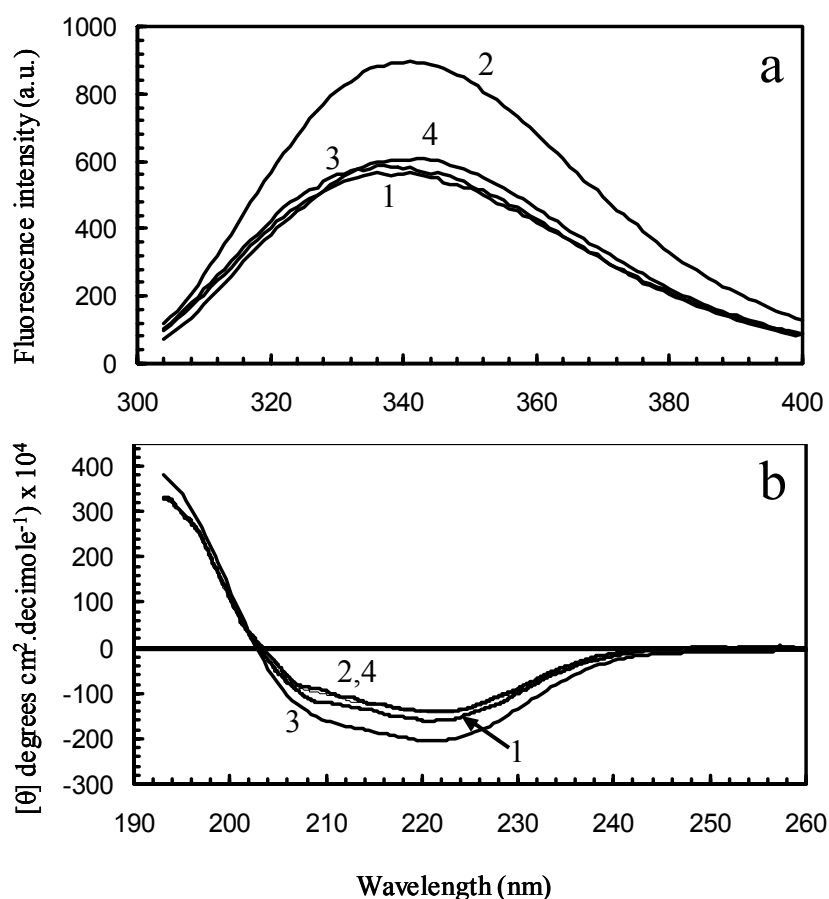
<sup>a</sup>Number of carboxyl groups determined using chromogenic Woodward K assay.

<sup>b</sup>Number of lysine groups determined using chromogenic OPA assay.

<sup>c</sup>Number of modified groups determined using the results of the Woodward K assay and OPA assay. The number indicates the number of succinyl (in case of succinylation) or methyl groups (in case of methylation) attached to the protein.

<sup>d</sup>Charge is calculated on the basis of primary sequence (1OVA) input in Swiss-Prot.

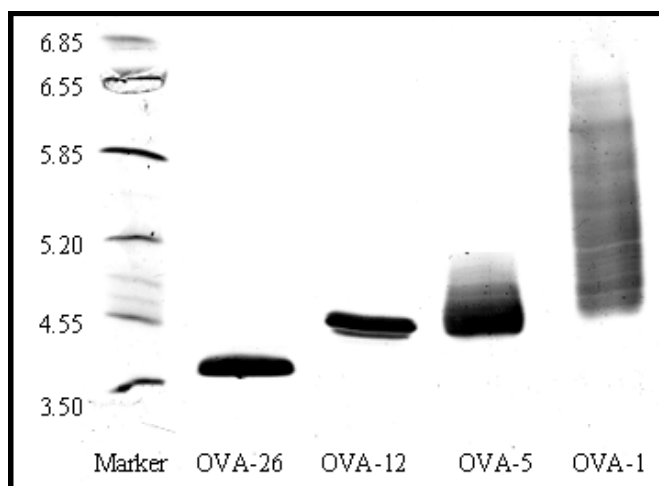
*Characterisation of physico-chemical properties of ovalbumin variants* - Figure 3 and Table I illustrate that the apparent isoelectric point (IEP) changed in the expected direction from 3.5 for OVA-26 up to  $\sim 7$  for OVA-1. OVA-1 showed a large range of iso electric points due to a higher heterogeneity resulting from the modification procedure compared to the other ovalbumin variants.



**Figure 2** Intrinsic tryptophan fluorescence spectra a) and far-UV circular dichroism spectra b) of unmodified and modified variants. The protein concentrations are  $2.7 \mu\text{M}$  for tryptophan fluorescence and  $5.5 \mu\text{M}$  for far-UV CD in  $10 \text{ mM}$  potassium phosphate pH 7.0 and the data are recorded at  $25^\circ\text{C}$ . The tryptophan fluorescence and far-UV circular dichroism spectra obtained are: OVA-26 (1), OVA-12 (2), OVA-5 (3) and OVA-1 (4).

Intermolecular interactions are, apart from electrostatics, dictated by the hydrophobicity of the protein and the presence of reactive cysteines. To isolate the effect of charge on aggregation properties it is important that hydrophobicity as well as the presence of reactive cysteines are not changed as a result of the charge modification.

Using PRODAN, an uncharged probe that, upon binding to hydrophobic areas on the protein surface results in fluorescence, the impact of the charge modification on the hydrophobic exposure of the ovalbumins was determined (Table II). The surface hydrophobicity ( $S_0$ ) determined in this way was not different for all the ovalbumin variants (Table II). All variants had an  $S_0$  between 59 and 68. For comparison, the  $S_0$  value of another folded globular protein,  $\beta$ -lactoglobulin, was reported to be 375 (Alizadeh-Pasdar & Li-Chan 2000).



**Figure 3** *Isoelectric focusing of charge variants of ovalbumin.*

To enable covalent aggregation through disulphide interactions, in the first place reactive thiol groups from cysteine have to be present in the primary sequence of the protein. Using Ellman's assay it was found that the number of thiol groups available in the ovalbumin was not affected by charge modification (Table II). The reactivity of the thiol groups can be determined using the sulfhydryl-disulphide exchange index (SEI). It was found that the reactivity of the thiol groups at 25°C was not affected by the charge modification (Table II) within the standard error of 5% for this experiment. Also at elevated temperature (~ 58°C), no differences in reactivity could be detected, except for OVA-1 which showed a slightly higher reactivity (SEI = 2.81) compared to the other ovalbumin variants (SEI = 3.14-3.28, Table II).

*Effect of charge on denaturant-induced unfolding* - Figure 4 shows the profiles of the normalised fluorescence intensity changes versus denaturant concentration from 0 to 8M urea for the equilibrium urea-induced unfolding of the differently charged ovalbumin variants. Two unfolding steps can be distinguished. For the highly charged OVA-26, the first transition takes place around a midpoint of 3.5 M urea (Table III). This transition is succeeded by another transition with a midpoint of 4.2 M urea. A lower net negative charge of the protein (OVA-12 and OVA-5) shifted this second transition to a higher denaturant concentration of 5.5 M urea. Even though all variants show the presence of a thermodynamically stable intermediate upon unfolding, this is even more explicitly shown for OVA-1. Calculation of the free energy of the unfolding transition ( $\Delta G$ ) is not valid for a number of reasons. First, the mentioned thermodynamically stable intermediate renders the transition a multiple state

process, which cannot be described as a two-state process. Second, the reversibility of the unfolding process was tested for all variants by refolding of the proteins from 8M urea by dilution in a buffer in the absence of urea. Using fluorescence, CD and SDS-PAGE, it was found that the transition is irreversible at the protein concentrations tested (data not shown). An important observation that can be deduced from these data is that upon changing the electrostatics of a protein, the equilibrium between folded, intermediate and unfolded molecules can be significantly affected by for example changing the thermodynamic stability of an intermediate state.

**Table II** Number of thiol groups, surface hydrophobicity ( $S_0$ ) and sulphhydryl-disulphide exchange index (SEI) for ovalbumin variants.

Ovalbumin variant	Number of thiol groups		$S_0^c$ Initial slope	SEI*	
	folded <sup>a</sup>	unfolded <sup>b</sup>		SEI <sub>25°C</sub>	SEI <sub>58°C</sub> <sup>d</sup>
OVA-26	0	5 ± 1	66	3.14	3.28
OVA-12	0	4 ± 1	68	3.43	3.21
OVA-5	0	4 ± 1	65	3.35	3.14
OVA-1	0	5 ± 1	59	3.15	2.81

<sup>a</sup>Number of SH groups detected using Ellman Assay

<sup>b</sup>Number of SH groups detected using Ellman Assay in the presence of 10% SDS

<sup>c</sup>Initial slope ( $S_0$ ), PRODAN fluorescence intensity/% protein, was calculated from fluorescence intensity vs. protein concentration plot

<sup>d</sup>OVA-26 and OVA-5 variants were heated at 54°C and 56°C, respectively

*Effect of charge on the formation of unfolding intermediates* – From urea-induced unfolding studies of OVA-1, a clear plateau region with significant intensity becomes apparent between 4 and 5 M urea (Figure 4). The change in the thermodynamics of an intermediate state has been further studied by investigating the  $\lambda_{\text{max}}$  shift as a function of the urea-concentration. It was found that for OVA-12, OVA-26 and OVA-5 the fluorescence intensity at 330 nm and the change in peak position ( $\lambda_{\text{max}}$ ) as a function of urea concentration coincide (see in Figure 5a for OVA-5). For the OVA-1 variant a transition midpoint at a urea concentration of 5.2 M was obtained by monitoring the fluorescence

intensity at 330 nm (Figure 5b, ♦) compared to a midpoint of 6.2 M when plotting the peak position as a function of urea concentration (Figure 5b, ◇). These results confirm that the presence of a thermodynamically stable intermediate in OVA-1 is much more pronounced compared to the OVA variants with a different net charge.

**Table III**      *Thermodynamic description of selected ovalbumin variants*

Ovalbumin variant	T <sub>t</sub> (°C) DSC <sup>a</sup>	E <sub>act</sub> <sup>b,c</sup> (kJ mol <sup>-1</sup> )	k of unfolding (min <sup>-1</sup> ) <sup>b</sup>	midpoint of unfolding (M urea) at 330 nm <sup>d</sup>
OVA-26	73.8	296	0.0402	4.2
OVA-12	78.8	338	0.0323	5.5
OVA-5	78.2	293	0.0345	5.5
OVA-1	76.2	208	0.0442	5.2

<sup>a</sup>The thermal transition temperature was described as the peak temperature in the DSC thermogram.

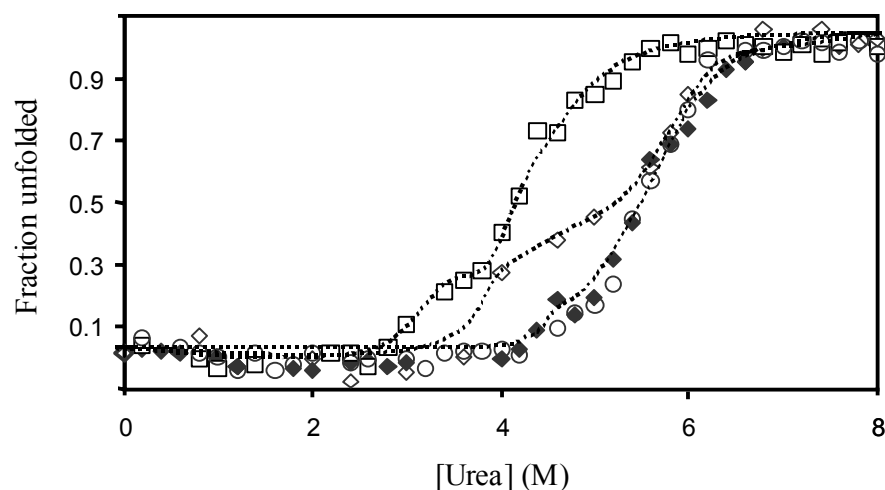
<sup>b</sup>The *k* of unfolding was calculated by a theoretical description of the temperature dependence of the frequency factor and the activation energy.

<sup>c</sup>The DSC curves are described using first-order irreversible unfolding models (Sanchez-Ruiz et al

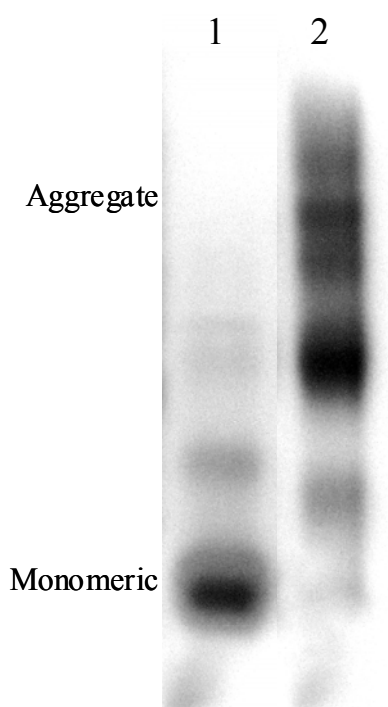
1988), resulting in the definition of an activation energy of unfolding.  $\frac{\nu}{T_t^2} = \frac{A \text{Re}^{-E/RT_t}}{E}$

where  $\nu$  is the scan rate, *E* the energy of activation, *A* the frequency factor of the Arrhenius equation, *R* the gas constant and *T<sub>t</sub>* is the thermal transition temperature.

<sup>d</sup>Urea-concentration at which 50% of the molecules reside in a folded state.



**Figure 4** Urea-induced unfolding equilibrium of ovalbumin variants. The protein concentration is  $1.17 \mu\text{M}$  in  $10 \text{ mM}$  potassium phosphate pH 7.0 and the data are recorded at  $25^\circ\text{C}$ . The fluorescence emission intensity data recorded at  $330 \text{ nm}$  are normalised and expressed as normalised fluorescence intensity change versus denaturant concentration. OVA-12 (O), OVA-5 ( $\blacklozenge$ ), OVA-1 ( $\diamond$ ), OVA-26 ( $\square$ ).

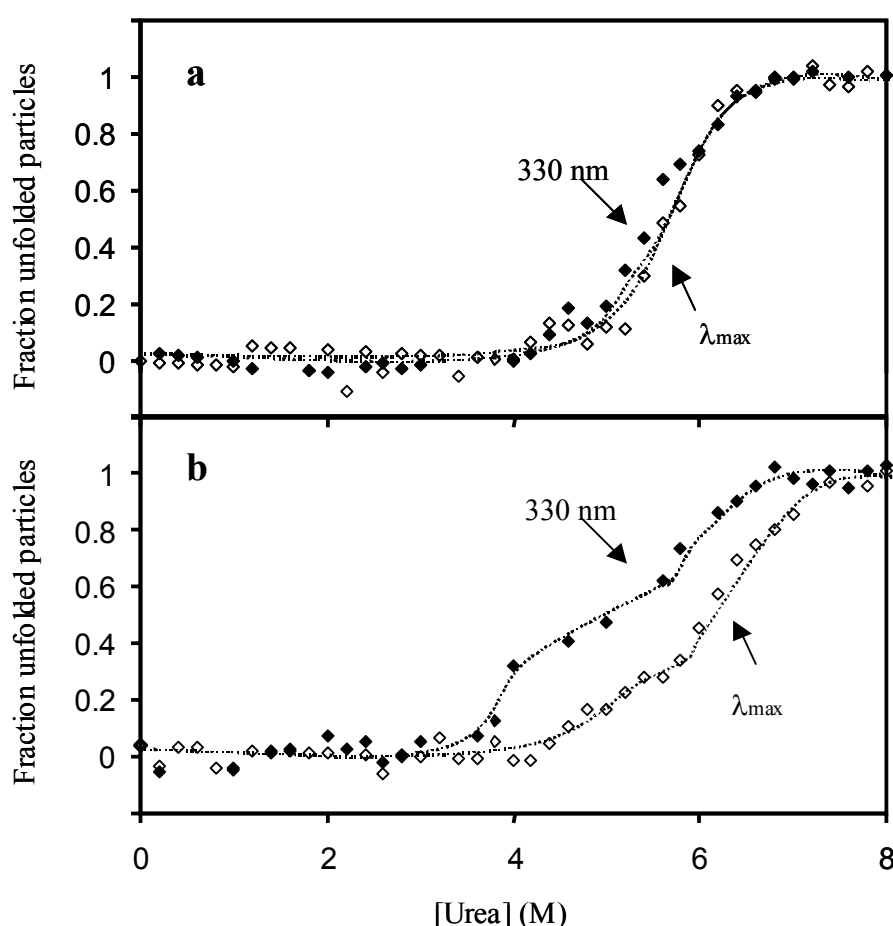


**Figure 8** Native-PAGE electrophoresis of OVA-1 incubated in  $4\text{M}$  urea (lane 1) and upon dilution with  $10 \text{ mM}$  potassium phosphate buffer at pH 7.0 in the absence of urea (lane 2).

*The effect of electrostatics on kinetic partitioning upon refolding* – The formation of a thermodynamically stable intermediate could give rise to a significant difference in the kinetic partitioning between aggregation and refolding. The contribution of the observed OVA-1 unfolding intermediate to the refolding process has been determined by urea- and heat-induced unfolding studies.

Using far-UV CD it was found that heat-induced unfolding at a protein-concentration of  $2.3 \mu\text{M}$  is an irreversible process for all variants tested (shown in Figure 6 for OVA-12 and OVA-1). The CD spectra at  $90^\circ\text{C}$  and subsequent refolding to  $20^\circ\text{C}$  appear very similar. These results are consistent with the finding that upon heating to  $90^\circ\text{C}$  by DSC (Figure 7) and

subsequent cooling to 20°C it was observed that OVA-1 as well as the other OVA variants do not refold to the original folded state (results not shown). Native gel electrophoresis confirmed that, upon refolding from a heat-induced unfolded state, all protein molecules proceeded into an intermolecular associated state (results not shown), indicating that the kinetic partitioning is driven toward aggregation. As also for OVA-1 all unfolded proteins adopt an aggregated state by both heat-induced and urea-induced unfolding, it can be assumed that the existence of a thermodynamically stable intermediate does not provide a supporting measure for the protein to refolding. To verify this, the OVA-1 intermediate was populated at 4M urea (see Figure 4). Upon dilution in buffer it was found, using native gel electrophoresis (Figure 8) that also from the intermediate state the preferable route is into an aggregated state.



**Figure 5** Urea-induced unfolding of a) OVA-5 and b) OVA-1 in 10 mM potassium phosphate at pH 7.0 and 25°C. Data obtained by monitoring the fluorescence intensity at 330 nm ( $\diamond$ ) and the frequency shift of the fluorescence maximum ( $\blacklozenge$ ) in the presence of 1.17  $\mu$ M of ovalbumin. The fluorescence emission intensity data recorded at 330 nm are normalised and expressed as normalised fluorescence intensity change versus denaturant concentration.

**Table IV** Volumetric and density parameters of selected ovalbumin variants

Ovalbumin variant	Stokes' Radius (Å) <sup>a</sup>		Density (kg m <sup>-3</sup> ) <sup>b</sup>	
	Folded*	Unfolded	Folded*	Unfolded
OVA-26	37.1	67.1	1171 <sup>c</sup>	648
OVA-12	32.0	64.1	1313	667
OVA-5	32.0	63.5	1339	675
OVA-1	31.1	62.2	1379 <sup>c</sup>	690

<sup>a</sup>Stokes' Radius determined from a standard curve with thyroglobulin, ferritin, catalase, aldolase, ovalbumin,  $\alpha$ -lactalbumin and cytochrome c using GPC and a slope of  $1000/V_{el} = 1.0368 R_s + 34.712$

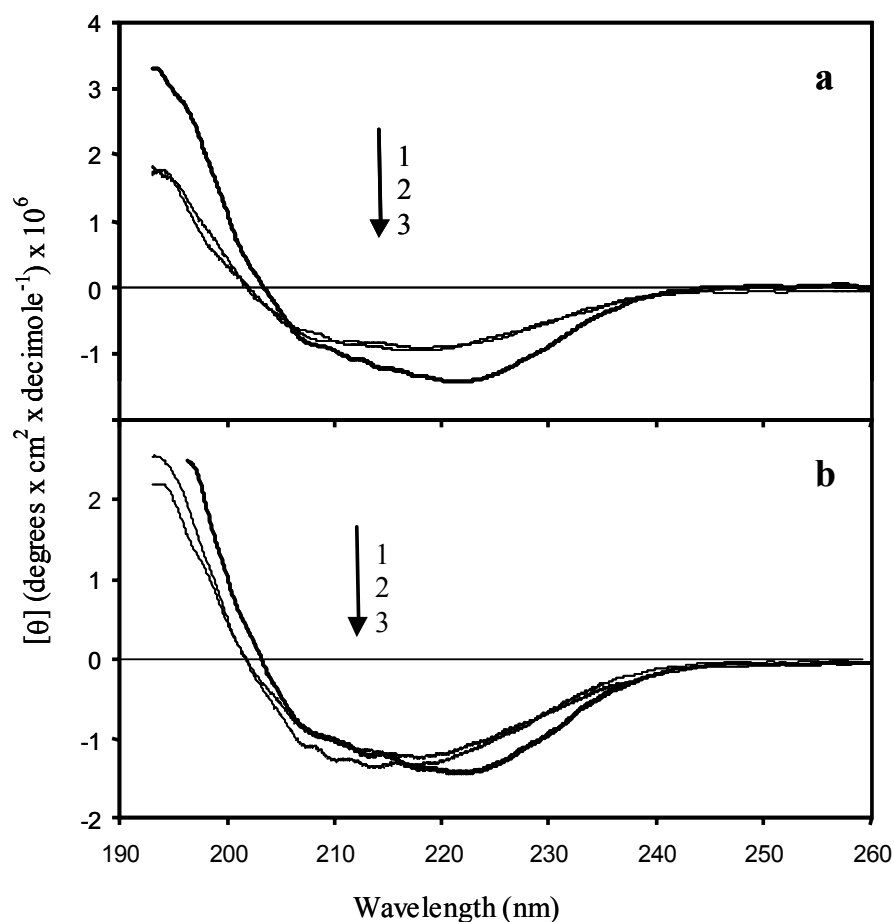
\*Based on duplicate experiment the standard error is 2%.

<sup>b</sup>Density of folded and unfolded molecules in 8M urea calculated from the Stokes' Radius obtained by GPC and calculated molecular weight based on the number of modification as reported in Table I. Molecular weights are OVA-26 43451 Da, OVA-12 42750 Da, OVA-5 42841 Da, OVA-1 42893 Da.

<sup>c</sup>Based on a 2% average experimental deviation (results not shown), these figures are significantly different from the density of the unmodified molecule.

*Effect of electrostatic interactions on thermal stability* - Figure 7 shows the DSC enthalpic unfolding profiles of the OVA variants. OVA-26 shows a transition peak between 65 and 80°C and has a peak temperature ( $T_i$ ) of 73.8°C (Table III). A decrease in net charge to -12 shifts the  $T_i$  up to 78.8°C. Upon decreasing the net charge more down to -5, the peak temperature decreases to 78.2°C and the peak becomes broader. For OVA-1, the peak temperature decreases further down to 76.2°C. Since the OVA variants (apart from the wild type protein) are to some extent heterogeneous in net charge (Figure 3), it was determined if charge heterogeneity would affect the thermal transition. However, no correlation was found between peak broadening and sample heterogeneity.

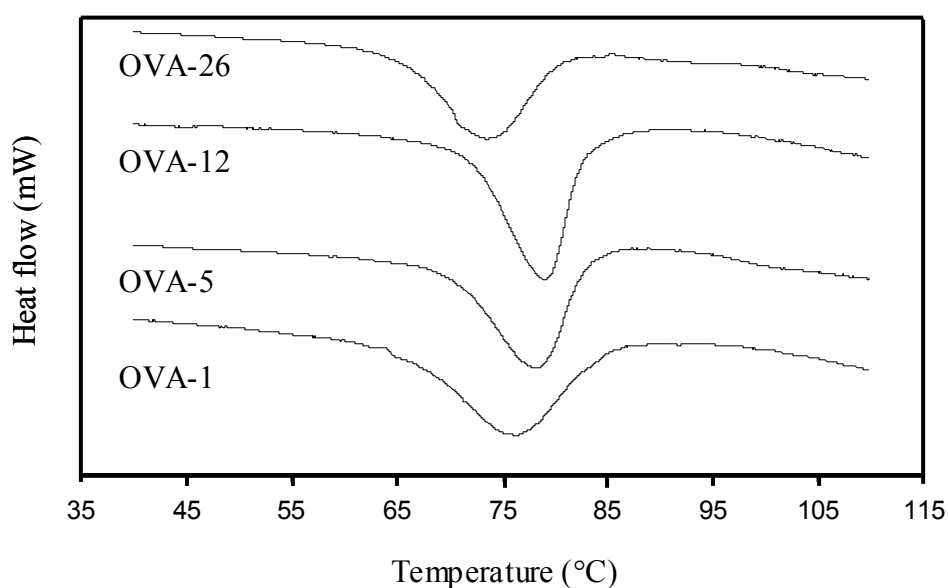




**Figure 6** Folded, unfolded and refolded OVA-12 a) and OVA-1 b). The spectra were obtained at 20°C (folded) (spectrum 3), 90°C (unfolded) (spectrum 1) and at 20°C after cooling (refolded) (spectrum 2). The protein concentration is 2.3  $\mu\text{M}$  in 10 mM potassium phosphate at pH 7.0.

*Effect of charge on the kinetics of heat-induced unfolding* - The activation energies of the rate-limiting step upon unfolding as deduced by fitting the DSC profiles (Figure 7) by an equation using first-order irreversible unfolding models (Sanchez-Ruiz et al 1998) are tabulated in Table III. OVA-26 has an activation energy of 296  $\text{kJ mol}^{-1}$ . Decreasing the net negative charge from this point leads to an increase of the activation energy: OVA-12 has an activation energy of 338  $\text{kJ mol}^{-1}$ . A further decrease in net negative charge leads to a decreased activation energy of 293  $\text{kJ mol}^{-1}$  for OVA-5 and 208  $\text{kJ mol}^{-1}$  for OVA-1. It should be noted that the unfolding transition is an irreversible transition and that the final ‘unfolded’ state cannot be defined. Therefore it is only plausible to ascribe the found activation energies to the transition with the highest activation energy, which is similar to the rate-limiting step. Even though these values do not always describe a well-defined unfolding transition, it becomes clear that the thermodynamics of the unfolding transition is

significantly altered upon charge modification. It was found that the two-state irreversible first-order model for unfolding fits the unfolding behaviour of all variants except for OVA-1 (Weijers et al 2005). This may be due to the pronounced population of a thermodynamically stable intermediate, also observed for urea-induced unfolding, affecting the unfolding kinetics (Figures 4 and 5b). Apparently, the population of thermodynamically stable intermediates observed for urea-induced unfolding for the other variants (OVA-26, OVA-12 and OVA-5) does not affect the possibility to fit the DSC data with a two-state irreversible first-order model.



**Figure 7** Temperature-induced unfolding of ovalbumin variants by differential scanning calorimetry. The protein concentration is 0.55 mM in 10 mM potassium phosphate pH 7.0 for DSC. The heating-rate was 1 °C/min.

*Effect of charge on the density of folded and unfolded ovalbumin variants* - A change in stability upon denaturant-induced unfolding could be interpreted either as a destabilisation of the folded state or as a stabilisation of the unfolded state of the molecule. To obtain an impression on the extent by which charge modification either affects the stability of the unfolded or the folded state of ovalbumin, we used the previously established knowledge that the surface accessible area (SAA) of a protein molecule is correlated to the heat capacity ( $C_p$ ) of the proteins (Myers et al 1995). The Stokes' Radius can be used as a direct measure for the solvent accessible area by determination of the molecular density (Richards 1977). The Stokes' Radius varies from 31.1 Å for OVA-1 to 37.1 Å for OVA-26 (Table IV). The

unfolded proteins have Stokes' Radii from 62.2 Å for OVA-1 to 67.1 Å for OVA-26. From these data the average molecular density can be calculated dividing the molecular mass over the volume of the molecule. The molecular mass is calculated from the degree of substitution as experimentally determined using the OPA assay (for the OVA-26 variant) and the Woodward assay (for the OVA-5 and OVA-1 variants) (Table I). The molecular density of the folded OVA-26 is  $1171 \pm 21 \text{ kg m}^{-3}$  (Table IV). Decreasing the net negative charge results in an increasing molecular density of the folded protein molecule:  $1313 \pm 16 \text{ kg m}^{-3}$  for OVA-12,  $1339 \pm 17 \text{ kg m}^{-3}$  for OVA-5 and  $1379 \pm 35 \text{ kg m}^{-3}$  for OVA-1 (Table IV). These values are all around the reported average value for folded protein molecules of  $1300 \text{ kg m}^{-3}$  (Fischer et al 2004). Only the OVA-26 variant has a slightly swollen structure in the folded state. These changes must be of a very local nature as the energy transfer between tryptophan and tyrosine in the folded ovalbumin variants using fluorescence (results not shown) did not differ upon charge modification. In the unfolded state, all densities appear comparable to the OVA-12 (wild type) variant of  $667 \pm 12 \text{ kg m}^{-3}$  (Table IV).

## Discussion

*Net charge modification affects protein stability* - Succinylation and methylation of ovalbumin have been used in this work to investigate the role of charge on the unfolding and refolding behaviour of proteins in the context of aggregation. Aggregate formation is generally constrained by the availability of a high concentration of (partly) unfolded molecules (conformational) and the absence of forces opposing aggregation (colloidal). Previous work mostly concentrated on one of these constraints in isolation. For example, Frankenberg et al (1999) and Sali et al (1991) showed that intramolecular electrostatic factors can contribute significantly to the conformational stability. Ma and Holme (1982) and Gossett and Baker (1983) indicated that increased repulsion caused by introducing additional net charge decreases protein-protein interactions and thus changes the colloidal stability. A combined approach will reveal the relative importance of both colloidal and conformational stability toward the aggregation phenomenon (see Weijers et al (2005)).

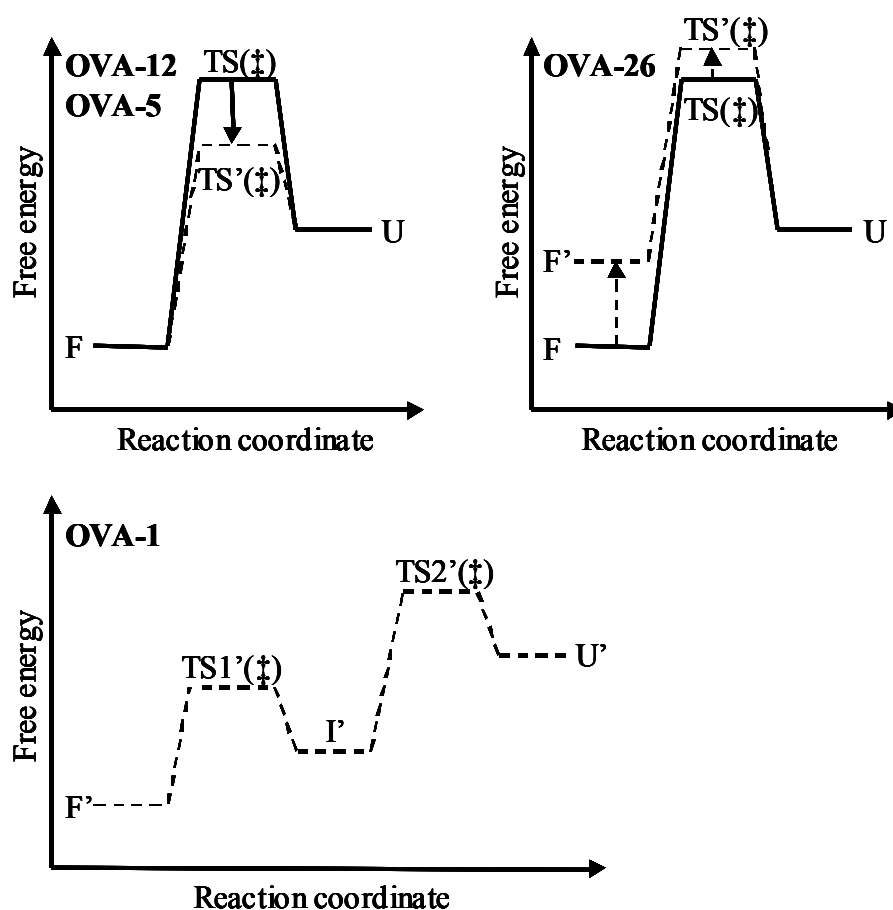
To understand the role protein net charge, brought about by chemical engineering, plays in the unfolding process in the context of aggregation propensity, it is important to also investigate to which extent other factors that determine aggregation properties are affected by the modification. These include hydrophobic interactions, sulfhydryl/disulphide exchange reactions. We found that it is possible to modify protein charge by succinylation or

methylation without affecting exposed surface hydrophobicity. The thiol reactivity was changed slightly for OVA-1 (at elevated temperature) and for the other variants it was not different (Table II).

Charge modification was found to have far-reaching implications on the stability of the variants (Table III, Figures 4 and 7). The consequences of this for protein thermodynamics are captured in a model depicted in Figure 9. This is first illustrated for urea-induced unfolding of OVA-26 which occurs at a significantly lower urea-concentration midpoint compared to OVA-12 (Figure 4, Table III). Also the peak temperature for unfolding is significantly lower for OVA-26 compared to OVA-12 (Figure 7, Table III). No linear relationship was found when relating the net charge with the stability of the variants determined by urea- or heat-induced unfolding. This can be explained by the fact that next to introducing an additional possibility for the molecule to form new ion pairs through charge introduction, this also introduces additional internal electrostatic repulsion leading to loss of intramolecular stability. This has been shown effectively by experiments by Sanchez-Ruiz and co-workers relating specific point mutations of various protein primary sequences to stability parameters. They showed convincingly that the contribution of electrostatic forces to stability is related to the specific location of ion pairs and the geometry of ion pair networks (Xiao & Honig 1999, Godoy-Ruiz et al 2004, Loladze et al 1999). As in this report we achieved charge modification through random charge engineering, it is not possible to directly relate the obtained results on stability with the specific location of the modified residues but we have clearly shown that charge modification can have a large thermodynamic impact on protein stability.

*Changing the net charge of the protein changes the rate of unfolding* – The rate of unfolding of all OVA-variants tested is a rate-limiting step in the aggregation process (Weijers et al 2005). An important condition to drive the aggregation process is a fast generation of a sufficient population of (partly) unfolded protein molecules. This rate of generation is determined by the activation energy of unfolding. It was found that charge reversal or elimination decreases the activation energy for unfolding compared to the wild type protein (OVA-12) (Table III). No relation was found between the net charge and the activation energy of unfolding. Cavagnero et al (1998) also found no correlation between protein net charge and unfolding rate. They showed that the unfolding rate of rubredoxin from two different organisms can differ significantly, even though the proteins carried a similar net charge. In the thermodynamic model proposed in Figure 9, the activation energies of

unfolding are shown to decrease for OVA-26 and OVA-1 having implications for the free energy content of the transition state. It was suggested that the location of ion pairs as well as the geometry of the ion pair networks are playing a larger role in determining the rate of unfolding than the total net charge. Cavagnero et al (1998) proposed a mechanism for the observed differences, which may also provide an explanation for our observations. They found that slow unfolding rates of a protein with intact salt bridges may be the result of surface ‘clamping’ exerted by ion pairs in the native structure. This leads to localised structural constraints on the protein surface that make the unfolding transition of the protein energetically less accessible. However, as the use of the succinylation or methylation reaction imply that the modification takes place randomly along the protein sequence, it is not possible to discuss the effects of this charge modification more specifically.



**Figure 9** Model representing how charge modification in a protein molecule can affect the energetics of the various species involved in the unfolding transition under native conditions.  $F$  and  $U$  represent the unmodified folded and unfolded species.  $F'$ ,  $I'$  and  $U'$  represent the modified folded and unfolded species.  $TS(\ddagger)$  represents the transition state.

*Changing the net charge of a protein has implications for the folded state* – In general, protein aggregation requires (partial) protein unfolding to allow intermolecular interactions to take place. To identify how charge variation affects protein stability, it is important to have an indication whether the modified thermal and urea-induced stabilities observed are a consequence of a stability change of either the native state or the unfolded state. An indication for this can be provided by the molecular density of the folded and unfolded protein molecule (Table IV). It was found that at the level of the folded molecule more profound swelling occurred for OVA-26 compared to the other ovalbumin variants (Table IV). On the other hand, the density of the unfolded molecule was hardly affected by the modification. Both the entropic as well as the enthalpic contributions may have been affected by the density change of the folded protein. An increase in volume generally is related to an increased surface accessible area and leads to an increased requirement of water molecules to arrange into a so-called “iceberg cage” surrounding the hydrophobic regions of the protein molecule. However, from PRODAN-binding experiments we found no indication that the hydrophobic exposure of OVA-26 was significantly higher than that of the other samples (Table II). The entropy of the folded system will therefore be affected by the swelling. Next to the increased entropic penalty, swelling of the folded molecule may result in disruption of ion pairs and hydrogen bonds resulting in a decreased enthalpy of the folded molecule, consistent with DSC results. A destabilisation of the folded state, via an enthalpic penalty, thus seems a likely way for the molecule to deal with additional negative charges (Figure 9).

*Changing the net charge of a protein may affect the unfolding/folding pathway* – The unfolding process of most OVA variants can be described by a two-state irreversible unfolding mechanism. The unfolding process of OVA-1 showed different behaviour, as it cannot be described by a two-state unfolding mechanism. Figure 5b shows that the expression of the fraction of folded molecules as a function of the shift of  $\lambda_{\text{max}}$  and the fluorescence intensity at 330 nm do not coincide. Also in Figure 4 it can be found that unfolding of OVA-1 shows a significant population of a thermodynamically stable intermediate state. For the other OVA variants tested  $\lambda_{\text{max}}$  and fluorescence intensity at 330 nm as a function of urea concentration were found to overlap. Even though a small inconsistency (larger for OVA-26) in the unfolding curves for the other ovalbumin variants was found in Figure 4, the intermediates formed do not seem to populate to the extent as it does for OVA-1. In Figure 9 the stable intermediate state has been schematically drawn for OVA-1 whereas the unfolding

transition for OVA-26, OVA-12 and OVA-5 are presented as a single transition. It is clear from these results that charge modification of ovalbumin may induce the formation of thermodynamically stable intermediates. It is possible that these intermediates can either direct the protein into a correctly folded state or into an aggregated state affecting the reversibility of the unfolding transition. However, as it was found that the unfolding transition of OVA-1 was irreversible, comparable to the other OVA variants tested, the OVA-1 intermediate thus does not play a role in the refolding of the protein. Even more, the accumulation of this intermediate state at moderate urea concentrations and subsequent dilution in urea-deficient buffer, led to the formation of aggregates (Figure 8). The intermediate should thus be much alike the state which is involved in the aggregation process. It is suggested that the appearance of a thermodynamically stable intermediate upon charge modification can be the result of repulsive forces that would otherwise be responsible for the fast transition of partially into fully unfolded molecules which are now screened and enable population of a visual intermediate (Figure 9).

Summarising, in this work we have shown that, next to the ability of protein molecules to overcome opposing electrostatic repulsive forces counteracting the aggregation, also the rate of unfolding is a crucial factor determining aggregation properties and are strongly affected by charge modification. The results have been presented into a model describing the effect of charge variation on protein thermodynamics (Figure 9). The results suggest that, when considering charge effects on the aggregation properties of proteins that it is crucial to combine the results of conformational stability studies as well as colloidal stability studies to evaluate the mechanism by which aggregation is controlled.

## Materials and Methods

### *Materials*

Ovalbumin was purified from fresh hen eggs according to a previously described purification protocol (Kosters et al 2003).

Carboxylic acid groups in ovalbumin were methylated using the procedure described by Kosters et al (2003) based on a method described by Hoare and Koshland (1967). A 1 L solution of 0.22 mM ovalbumin containing 1M methylamine was set to pH 4.6 using 1M HCl and stirred at room temperature. Subsequently, 300 mg solid N-(3-Dimethylaminopropyl)-N-ethylcarbodiimide hydrochloride (EDC) was added stepwise in four portions to give a final concentration of 100 mM, while the pH was kept at 4.6 by the addition of 1 M NaOH using a pH-stat (Metrohm). When all added EDC was dissolved, the solution was stirred for another 2

h. Next, the pH was raised to 8.0 by drop wise addition of NaOH (1M). After another 1 h of stirring at room temperature the mixture was dialysed extensively against demineralised water at 4°C followed by freeze-drying. This procedure resulted in the generation of OVA-1, OVA-5 was prepared by stirring for 1 h after the addition of EDC and keeping all other conditions constant. The material was stored at -20°C.

Primary amino groups in ovalbumin were succinylated essentially as described by Kusters et al (2003) based on the procedure described by Klotz (1967). 400 mL solution of 55 mM ovalbumin in demineralised water was set at pH 8.0 by the addition of 3 M NaOH using a pH-stat (Metrohm) at room temperature. Then 500 mg of solid succinic anhydride was added in small aliquots during 30 min yielding a final concentration of 30 mM succinic anhydride. During the addition of succinic anhydride the pH was kept at 8.0 ( $\pm 0.1$ ) by the addition of 1 M NaOH using pH-stat equipment. Next, the solution was stirred for another 2 h, followed by extensive dialysis against demineralised water at 4°C and subsequent freeze-drying. The material was stored at -20°C until use.

#### *Material characterisation – chemical characterisation*

The purity of ovalbumin, before and after modification was checked using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) according to Laemmli (1970).

The apparent iso-electric point (IEP) of the ovalbumin variants was determined as described previously using iso-electric focusing (Broersen et al 2004).

The degree of succinylation was determined by the method described by Church et al (1983) based on the specific reaction between ortho-phthaldialdehyde (OPA) and free primary amino groups in proteins. In the presence of 2-(dimethyl amino) ethanethiol hydrochloride (DMA) these amino groups react to alkyl-iso-indole derivatives that show absorption at 340 nm. All measurements were performed in triplicate.

The available number of carboxylic acid groups on the protein was determined by a spectrophotometric method using Woodward reagent K as described by Kusters and de Jongh (2003). This method is based on the reaction between Woodward reagent K and free carboxylic acid groups resulting in an increase in absorbance at 269 nm. All measurements were performed in duplicate.

*Structural characterisation* - To investigate protein structure, emission and excitation fluorescence spectra of 2.7  $\mu$ M (modified) ovalbumin solutions in a phosphate buffer (10



mM, pH 7.0) were analysed using a Perkin Elmer (LS 50 B) luminescence spectrometer at 20°C ( $\pm 0.1^\circ\text{C}$ ). The emission spectra from 300 to 400 nm were recorded at an excitation wavelength of 274 nm and 295 nm. Excitation and emission slit widths were 5 nm and a scan speed of 120 nm/min was used. All spectra were corrected for the corresponding protein-free sample.

Far- and near UV-circular dichroism (CD) experiments were performed as described previously (Broersen et al 2004) using a Jasco J-715 spectropolarimeter (Jasco Corp., Japan). Spectra were recorded as averages of sixteen spectra. The scanning speed was 100 nm/min and the response time was 0.5 s with a bandwidth of 1.0 nm. The protein variants were dissolved in 10 mM phosphate buffer at pH 7.0 at 2.3  $\mu\text{M}$  for far-UV CD and 23.4  $\mu\text{M}$  for near-UV CD. Cells with path lengths of 0.1 and 1 cm were used for far-UV CD and near-UV CD, respectively.

The number of free thiol groups was determined using the chromogenic Ellman's assay as has been published before (Ellman et al 1959). The amount of accessible thiol groups of the ovalbumin variants in the presence and absence of 10% SDS using 2-nitro-5-mercaptobenzoic acid (DTNB).

The sulphhydryl-disulphide exchange index (SEI) was determined according to Owusu-Apenten et al (2003). 0.3 mL of  $2.2 \times 10^{-4}$  M ovalbumin solutions were reacted with 2.7 mL  $2.5 \times 10^{-4}$  M 2-pyridine disulphide in a 10 mM sodium phosphate buffer (pH 7.0) at 25°C and at 58°C. The formation of 2-thiopyridine (2-TP) was spectrophotometrically followed at 343 nm as a function of time. The equation for this 2<sup>nd</sup> order reaction was published by Owusu-Apenten et al (2000). The rate constant for pyridine disulphide reaction with protein ( $k$ ) or GSH ( $k^*$ ) is given in mole protein per second and was standardised by the rate constant for glutathione reaction with 2-pyridine disulphide determined under identical conditions and expressed as the Sulphydryl-disulphide Exchange Index (SEI) (1).

$$\text{SEI} = -\log(k/k^*) \quad (1)$$

*Thermal unfolding* - DSC scans were obtained using a SETARAM (Caluire, France) micro DSC III with stainless steel 1 mL sample cells. Solutions were prepared with protein concentrations of 0.23 mM, pH 7.0 and also at 2.2 mM for OVA-26 at pH 7.0. The protein-free solution was used in the reference cell. The temperature was varied from 25 to 120°C at a scan rate of  $1^\circ\text{C min}^{-1}$ . Reference-reference baselines were obtained at the same scanning

conditions and subtracted from the sample curves. The measurements were performed in duplicate.

*Equilibrium unfolding* - Urea titration experiments were performed to monitor the change in fluorescence intensity upon the addition of urea. A 10 mM phosphate buffer solution at pH 7.0 was prepared using reagent grade salts  $\text{NaH}_2\text{PO}_4$  and  $\text{Na}_2\text{HPO}_4$  (Merck) in demineralised water. An 8 M urea stock solution was prepared by dissolving 0.755 g of urea (Merck) per mL in a 10 mM phosphate buffer solution. The protein stock solutions were prepared by diluting the protein with 10 mM phosphate buffer pH 7.0 at 23.9  $\mu\text{M}$ . One hundred  $\mu\text{L}$  of the protein stock solution was then diluted with 2 mL of urea-solution with final concentrations ranging from 0-6 M. Prior to the measurement the samples were incubated for 24 h at room temperature which is required to reach stable fluorescence intensity (unpublished results). After incubation, the sample was placed into a thermostatted 1 mL quartz cell at 25°C. The fluorescence emission was recorded between 300 and 450 nm upon excitation at 295 nm on a Varian Cary Eclipse Spectrofluorimeter. The slit widths were set at 5 nm.

*Exposed hydrophobicity* - Fluorescence of N,N-Dimethyl-6-propionyl-2-naphthylamine (PRODAN; BioChemika) was measured at 20°C using a Varian Cary Eclipse Spectrofluorimeter. 0.19 mM PRODAN dissolved in methanol. The final solution was filtered and the concentration was determined at 365 nm using an extinction coefficient of  $\epsilon_{365} = 1.45 * 10^4 \text{ M}^{-1} \text{ cm}^{-1}$  (Weber & Farris 1979). Fluorescence spectra were acquired at  $\lambda_{\text{ex}}$  361 nm and  $\lambda_{\text{em}}$  380-620 nm at a constant probe concentration of  $9.2 * 10^{-7} \text{ M}$  and a varying protein concentration ranging from  $7.1 * 10^{-6}$  to  $1.4 * 10^{-4} \text{ M}$ . 1 mL quartz cuvettes were used.

*Stokes' Radius determination* - The Stokes' Radii of the ovalbumin variants were determined with gel permeation chromatography (GPC). Proteins (0.23 mM) in 10 mM phosphate (pH 7.0) in the presence and absence of 6 M urea were incubated overnight at 4°C. A Superdex 200 column (10/30) connected to an Äkta Purifier (Amersham Biosciences) was used and equilibrated either with 10 mM phosphate (pH 7.0) containing 0.15 M NaCl (for the folded protein) or 10 mM phosphate (pH 7.0), 0.15 NaCl and 6 M urea (for the unfolded protein). The protein was eluted at 20°C and detected at 280 nm. The measurements were duplicated.

## Acknowledgements

The authors thank Prof. Sanchez-Ruiz, Dr. Ton Marcelis and Prof. Willem Norde for stimulating discussions, Prof. Voragen for careful reading of the manuscript and Ir. Hans Kusters for carrying out the chromogenic Ellman's assay.

## References

- Alizadeh-Pasdar N, Li-Chan ECY (2000) *J Agric Food Chem* 48: 328-334
- Broersen K, Voragen AGJ, Hamer RJ, de Jongh HHJ (2004) *Biotechnol Bioeng* 86: 78-87
- Cavagnero S, Debe DA, Zhou ZH, Adams MWW, Chan SI (1998) *Biochemistry* 37: 3369-3376
- Chiti F, Stefani M, Taddei N, Ramponi G, Dobson CM (2003) *Nature* 424: 805-808
- Chiti F, Calamai M, Taddei N, Stefani M, Ramponi G, Dobson CM (2002) *Proc Natl Acad Sci USA* 99: 16419-16426
- Church FC, Swaisgood HE, Porter DH, Catignani GL (1983) *J Dairy Sci* 66: 1219-1227
- de la Paz ML, Serrano L (2004) *Proc Natl Acad Sci USA* 101: 87-92
- Ellman GL (1959) *Arch Biochem Biophys* 82: 70-77
- Fandrich M, Dobson CM (2001) *Amyloid* 8: 26-26
- Fischer H, Polikarpov I, Craievich AF (2004) *Protein Sci* 13: 2825-2828
- Frankenberg N, Welker C, Jaenicke R (1999) *FEBS Lett* 454: 299-302
- Godoy-Ruiz R, Perez-Jimenez R, Ibarra-Molero B, Sanchez-Ruiz JM (2004) *J Mol Biol* 336: 313-318
- Gossett PW, Baker RC (1983) *J Food Sci* 48: 1391-1394
- Gounaris AD, Perlmann GE (1967) *J Biol Chem* 242: 2739-2745
- Guijarro JJ, Sunde M, Jones JA, Campbell ID, Dobson CM (1998) *Proc Natl Acad Sci* 95: 4224-4228
- Hoare DG, Koshland DE (1967) *J Biol Chem* 242: 2447-2453
- Hwang W, Zhang S, Kamm RD, Karplus M (2004) *Proc Natl Acad Sci USA* 101: 12916-12921
- Kelly JW (1998) *Curr Opin Struct Biol* 8: 101-106
- Kitabatake N, Kinekawa Y-I (1995) *Food Chem* 54: 201-203
- Klotz IM (1967) *Methods Enzymol* 11: 567-580
- Koseki T, Kitabatake N, Doi E (1989) *Food Hydr* 3: 123-134
- Kusters H, Broersen K, de Groot J, Simons J-WFA, Wierenga P, de Jongh HHJ (2003) *Biotechnol Bioeng* 84: 61-70
- Kusters H, de Jongh HHJ (2003) *Anal Chem* 75: 2512-2516
- Laemmli UK (1970) *Nature* 227: 680-685
- Lansbury PT (1999) *Proc Natl Acad Sci USA* 96: 3342-3344
- Loladze VV, Ibarra-Molero B, Sanchez-Ruiz JM, Makhataдзе GI (1999) *Biochemistry* 38: 16419-16423
- Lopez De La Paz M, Goldie K, Zurdo J, Lacroix E, Dobson CM, Hoenger A, Serrano L (2002) *Proc Natl Acad Sci U S A* 99: 16052-16057
- Ma C-Y, Holme J (1982) *J Food Sci* 47: 1454-1459
- Malisauskas M, Zamotin V, Jass J, Noppe W, Dobson CM, Morozova-Roche LA (2003) *J Mol Biol* 330: 879-890

- Merz PA, Wisniewski HM, Somerville RA, Bobin SA, Masters CL, Iqbal K (1983) *Acta Neuropathol (Berl)* 60: 113-124
- Myers JK, Pace CN, Scholtz JM (1995) *Protein Sci* 4: 2138-2148
- Nakamura R, Sugiyama H, Sato Y (1978) *Agric Biol Chem* 42: 819-824
- Nakagawa Y, Capetillo S, Jirgensons B (1972) *J Biol Chem* 247: 5703-5708
- Narang HK (1980) *J Neuropathol Exp Neurol* 39: 621-631
- Owusu-Apenten RK, Galani D (2000) *J Sci Food Agric* 80: 447-452
- Owusu-Apenten RK, Chee C, Hwee OP (2003) *Food Chem* 83: 541-545
- Pots AM, ten Grotenhuis E, Gruppen H, Voragen AGJ, de Kruif KG (1999) *J Agric Food Chem* 47: 4600-4605
- Richards FM (1977) *Ann Rev Biophys Bioeng* 6: 151-176
- Sali D, Bycroft M, Fersht AR (1991) *J Mol Biol* 220: 779-788
- Sanchez-Ruiz JM, Lopez-Lacombe JL, Cortijo M, Mateo PL (1988) *Biochemistry* 27: 1648-1652
- Sun Y, Hayakawa S (2002) *J Agric Food Chem* 50: 1636-1642
- Uversky VN, Fink AL (2004) *Biochim Biophys Acta* 1698: 131-153
- Veerman C, de Schiffart G, Sagis LMC, van der Linden E (2003) *Int J Biol Macromolecules* 33: 121-127
- Villegas V, Zurdo J, Filimonov VV, Aviles FX, Dobson CM, Serrano L (2000) *Protein Sci* 9: 1700-1708
- Weber G, Farris FJ (1979) *Biochemistry* 18: 3075-3078
- Weijers M, Barneveld PA, Cohen Stuart MA, Visschers RW (2003) *Protein Sci* 12: 2693-2703
- Weijers M (2005) Electrostatics controls fibril formation. Part II: denaturation, aggregation and gelation of charged ovalbumin variants. In: *Aggregate morphology and network properties of ovalbumin*. Thesis from Wageningen University, pp 118-138
- Xiao L, Honig B (1999) *J Mol Biol* 289: 1435-1444
- Zurdo J, Gujjarro JJ, Jiménez JL, Saibil HR, Dobson CM (2001) *J Mol Biol* 311: 325-340

# CHAPTER 7

## ELECTROSTATICS CONTROLS FIBRIL FORMATION. PART II: NET CHARGE CAN AFFECT FIBRIL MORPHOLOGY AND GEL PROPERTIES OF OVALBUMIN

Weijers M, Broersen K, Barneveld PA, Cohen Stuart MA, Hamer RJ, de Jongh HHJ,  
Visschers RW

Manuscript in preparation

### **Abstract**

The effect of electrostatic interactions on aggregation rate, fibril morphology and gel properties of ovalbumin were investigated using chromatography, electrophoresis, electron microscopy and turbidity measurements. A range of differently charged ovalbumin variants were produced ranging in net charge from  $-1$  to  $-26$ . The rate of the heat-induced generation of unfolded molecules was significantly affected by net charge. Upon correcting for differences in thermal transition temperature, the apparent differences in the rate of unfolding were eliminated. Even though the rate of unfolding after this correction was not affected by the modified electrostatics, net charge significantly affected the morphology of the formed aggregates. At a net charge of  $-5$  and lower, branched and clustered aggregates were formed and with increasing charge, the aggregates became stiffer and more fibrillar with less branching. Also the visual appearance of the gels was affected. Gels from the aggregates with a low net charge were more turbid, whereas the highly charged aggregates formed gels which were transparent, also upon significantly increasing the ionic strength. These results demonstrate that the charge density on the protein molecule plays a major role in the formation of fibrillar structures and in the resulting visual appearance of gels. By controlling the charge density on the protein molecules, the magnitude of electrostatic repulsion can be changed, which allows us to control fibril morphology and therefore the turbidity and rheological properties of ovalbumin networks.

### **Keywords:**

Ovalbumin, electrostatics, fibril morphology, gelation

## Introduction

Fibril formation by aggregating proteins has been a major topic for a number of decades in biology, medicine and the food industry (Gosal et al 2002, Kallberg et al 2001, Koo et al 1999, Shah & Janmey 1997, Ellis & Pinheiro 2002). In biological systems, the first event leading to protein aggregation is often related to enzymatic cleavage or misfolding, whereas in food systems it is usually heat-induced. For many proteins unfolding has been found to be an irreversible process (Donovan & Beardslee 1975, Sanchez-Ruiz et al 1988, La Rosa et al 2002). This observation has now been firmly related to the occurrence of a subsequent aggregation process (Klibanov & Ahern 1987). The irreversibility of the unfolding process is largely governed by the balance between hydrophobic attraction and electrostatic repulsion between unfolded proteins. Electrostatics has thus been identified to importantly affect the aggregation process (Raman et al 2005, Lopez de la Paz et al 2002, Zurdo et al 2001, Zbilut et al 2004). To investigate the role of electrostatics in the process of aggregation, electrostatic repulsion can be altered by chemical modification or introduction of the charged residues of a polypeptide chain. In a number of papers it was shown that heat-induced aggregation of ovalbumin and egg white protein is retarded if the charge on the protein was increased (Nakamura 1978, Ma & Holme 1982, 1986, Mine 1996). It was also found that aggregate morphology and the visual and textural aspects of the formed gel may be somehow interrelated (Mine 1996). Mine (1996) showed for example that increasing the net charge changed the transparency and elasticity of ovalbumin gels. The mechanism for this has however not been clarified. The effect of electrostatic interactions on fibril morphology has been considered by Odijk (Odijk & Mandel 1978, Odijk 1978, 1977). An essential term to take into consideration upon charge engineering is the effect the modification can have on the overall conformation of the protein. This is essential in the discussion of the effect of charge on aggregate and gel properties. A series of papers have shown that charge engineering can have serious implications for the conformational structure and stability (Stein et al 1991, Xiao & Honig 1999, Godoy-Ruiz et al 2004, Loladze et al 1999). It is thus important to evaluate the effects of the modification procedure on the protein conformation. In addition, a set of criteria can be formulated by which the engineered proteins can be selected.

As discussed earlier, protein aggregation is often associated with an irreversible unfolding process. It may thus be evident that the two steps of unfolding and aggregation involved require a full investigation. The first step, namely the unfolding process is described in detail in the paper by Broersen et al (2005). This paper showed that it is possible, under selected conditions to use charge engineering of proteins, although random, in a non-

destructive manner with respect to the folded structure of the wild type protein. Using succinylation and methylation, four ovalbumin variants were generated ranging in charge from -1 (OVA-1), -5 (OVA-5), -12 (OVA-12) and -26 (OVA-26). The numbers denote the net charges calculated at pH 7.0. In the preceding paper (Broersen et al 2005), it was demonstrated that these variants showed no changes in the overall topology, i.e., in protein conformation, surface-exposed hydrophobicity or disulphide interaction potential while their stability was significantly affected. Therefore we consider this series an excellent model system to study the effect of electrostatic interactions on aggregation. Also, it was found that, as the modification did not affect the structure of the unfolded molecule apart from introducing a significant charge variation, this method can serve as a useful approach to investigate the effect of charge variation on aggregate morphology. In the work described in this publication, the effect of protein net charge on aggregate formation and morphology and subsequent gel properties were investigated, using ovalbumin, a chicken egg white protein, as a model system.

## Results

Broersen et al (2005) prepared a series of ovalbumin (OVA) variants with ranging net charge by chemical modification (either succinylation of lysine groups or methylation of carboxyl groups). A full structural and thermodynamic characterisation of these proteins can be found in the preceding paper (Broersen et al 2005). In that paper, the variants were characterised in terms of size, hydrophobicity, sulfhydryl-disulphide exchange index (SEI), and structural elements. Based on the results of this characterisation, four variants were selected with intact structures that have comparable molecular characteristics but differ in their net charge (OVA-1, OVA-5, OVA-12 and OVA-26, in which the numbers -1 to -26 represent the calculated net charges of the proteins at pH 7.0). A summary of the physico chemical characteristics of these variants can be found in Table I. Using these engineered OVA variants we expect to elucidate in which way net charge affects fibril formation, fibril morphology and gel formation.

**Table I** *Iso-electric point (isoelectric focusing) and net charge at pH 7.0 (calculated using the primary sequence PDB: 1OVA) for ovalbumin variants after chemical modification.*

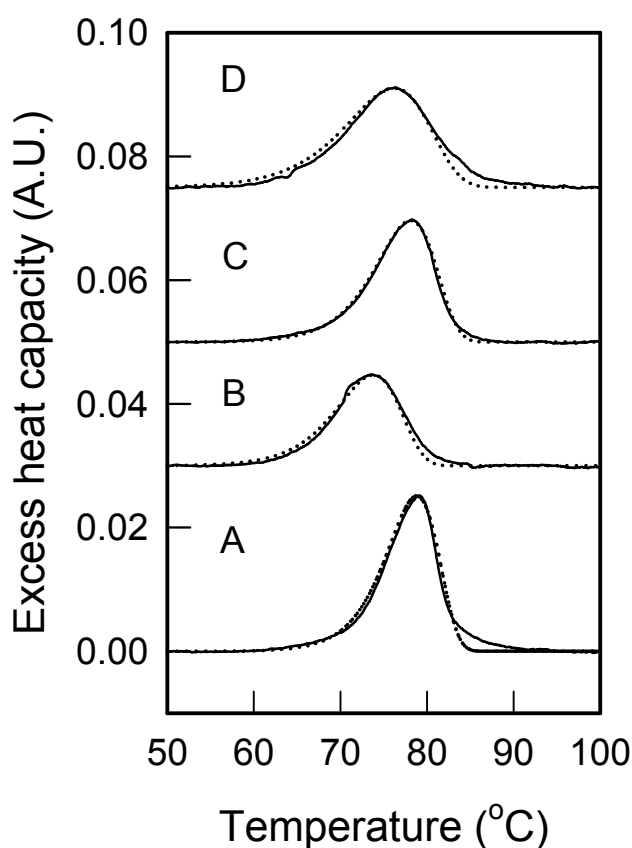
Ovalbumin variant	pI (-)	Net charge at pH 7.0
OVA-26	3.5	-26
OVA-12	4.7	-12
OVA-5	5.0	-5
OVA-1	5-7	-1

### *Unfolding and aggregation kinetics*

#### *Effect of surface charge on the unfolding rate and irreversibility*

Broersen et al (2005) already showed that charge variation can have significant consequences for the thermodynamic behaviour of ovalbumin. As the thermodynamic behaviour of proteins has been related to the aggregation probability of proteins in the past, we first investigated the effect of charge variation on the heat-induced unfolding rates of ovalbumin. Figure 1 shows that the peak temperatures determined from DSC thermograms of the differently charged OVA variants decreased for OVA-26, OVA-5 and OVA-1 compared to OVA-12. This finding is consistent with previous observations that net charge can significantly affect protein stability (Xiao & Honig 1999, Godoy-Ruiz et al 2004, Loladze et al 1999). It was also already found in the preceding paper that the denaturant-induced unfolding process is irreversible for all charge variants (Broersen et al 2005). In the present paper we found that heat-induced unfolding is also irreversible and that this irreversibility is net charge-independent under the conditions tested. Reheating showed no endothermic transition for any of the charge variants (results not shown). Therefore, charge engineering did not improve the refolding yield upon cooling from heat-induced unfolding under these conditions. Previously, we have shown that the heat-induced unfolding of OVA-12 can be well described by a two-state irreversible unfolding model ( $F \rightarrow U$ ) (for details see Weijers et al 2003). In this paper, we used this two-state irreversible model to fit the experimental DSC profiles for the differently charged OVA variants at 1.0°C/min (Figure 1).





**Figure 1** Differential scanning calorimetry thermograms and the best fit of the two-state irreversible unfolding model for the OVA variants in solution at a scan rate of 1.0 °C/min (solid lines), pH 7.0 and protein concentrations of 10 mg/mL. (A) OVA-12; (B) OVA-26; (C) OVA-5 and (D) OVA-1. Dashed lines represent irreversible first-order fits and continuous lines represent the data obtained by experiments.

As can be judged from Figure 1, the two-state irreversible model successfully fits the experimental data in the whole temperature range for OVA-26, OVA-12 and OVA-5. The experimental data of OVA-1 (Figure 1d) are less well described by this model. These observations are consistent with the finding in the preceding paper (Broersen et al 2005) that ovalbumin with a net charge of -1 led to a significant population of a thermodynamically stable intermediate upon denaturant-unfolding. It can be questioned whether heat- and denaturant-induced data are interchangeable. Interestingly, for both unfolding-inducing modes, inconsistencies with a two-state unfolding process were observed.

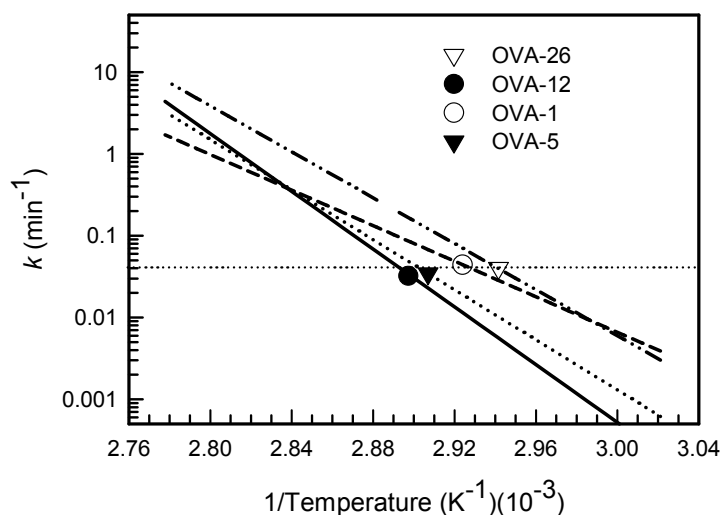
From the fits obtained by fitting the experimental DSC data to the two-state irreversible model, the energy of activation of the unfolding process ( $E_a$ ) and the frequency

factor of the Arrhenius equation ( $A$ ) can be calculated (Table II). From these data, the rate constant ( $k$ ) of the rate-limiting step was calculated using the Arrhenius equation (Figure 2). Figure 2 shows the calculated rate constants in the temperature range of 58 to 87°C for the OVA variants using the values tabulated in Table II. It was shown before for OVA-12 that the rate constant shows a linear dependency with temperature within this temperature region (Weijers et al 2003).

**Table II** *Thermodynamic parameters of the engineered OVA variants obtained from fitting experimental DSC data (Figure 2).*

Ovalbumin variant	$T_t$ (°C) <sup>a</sup>	$E_{act}$ (kJ/mole)	$\ln(A \text{ (min}^{-1}\text{)})$
OVA-26	73.8	269	92
OVA-12	78.8	338	114
OVA-5	78.1	293	99
OVA-1	76.2	208	70

<sup>a</sup> $T_t$  is the peak temperature (thermal transition temperature) at a scan rate of 1°C/min,  $E_{act}$  the energy of activation and  $A$  the frequency factor of the Arrhenius equation.



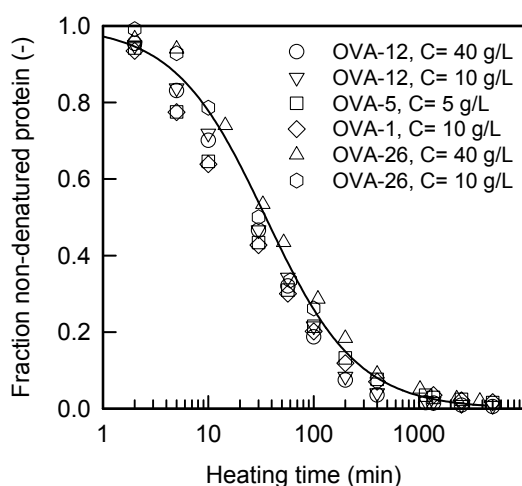
**Figure 2** *Rate constant ( $k$ ) for different OVA variants as a function of temperature. First-order rate constants, calculated from DSC results, change with temperature as described in the Arrhenius equation; frequency factor and energy of activation are obtained from Table II. The symbols represent the experimentally observed rate constants with HPSEC at which the solutions of OVA variants were heated (7 °C below  $T_t$ ) as presented in Figure 3.*

*Effect of surface charge on heat-induced generation of unfolded molecules*

To experimentally test the effect of net charge on the kinetics of unfolding, the fraction of folded monomers as a function of heating time for the OVA variants was determined with HPSEC (Figure 3). While preliminary experiments indicated that at the peak temperature found by DSC the high rate of unfolding was difficult to follow experimentally, we selected a temperature at a fixed temperature (in this case  $-7^{\circ}\text{C}$ ) below the peak temperature to test if the difference in peak temperature observed by DSC (Table II) was responsible for the observed differences in rate of unfolding. This shift to below the thermal transition temperature also ensures an experimentally accessible rate of unfolding.

The outcome of this experiment led to two major observations. First, incubation of the charge engineered variants at different concentrations does not result in differences in the observed rate of unfolding. As aggregation from the unfolded state has been previously shown to be a concentration-dependent process, whereas unfolding is not, this observation suggests that unfolding is the rate-limiting step for all the OVA variants tested. Moreover, the first-order behaviour as shown previously for OVA-12 (Weijers et al 2003), also applies to the OVA variants used in this study. The small variation in the loss of folded proteins upon heating for the OVA variants shown in Figure 3 are the result of small variations in the exact heating conditions and underline the strong dependence of the kinetics of unfolding on the incubation temperature.

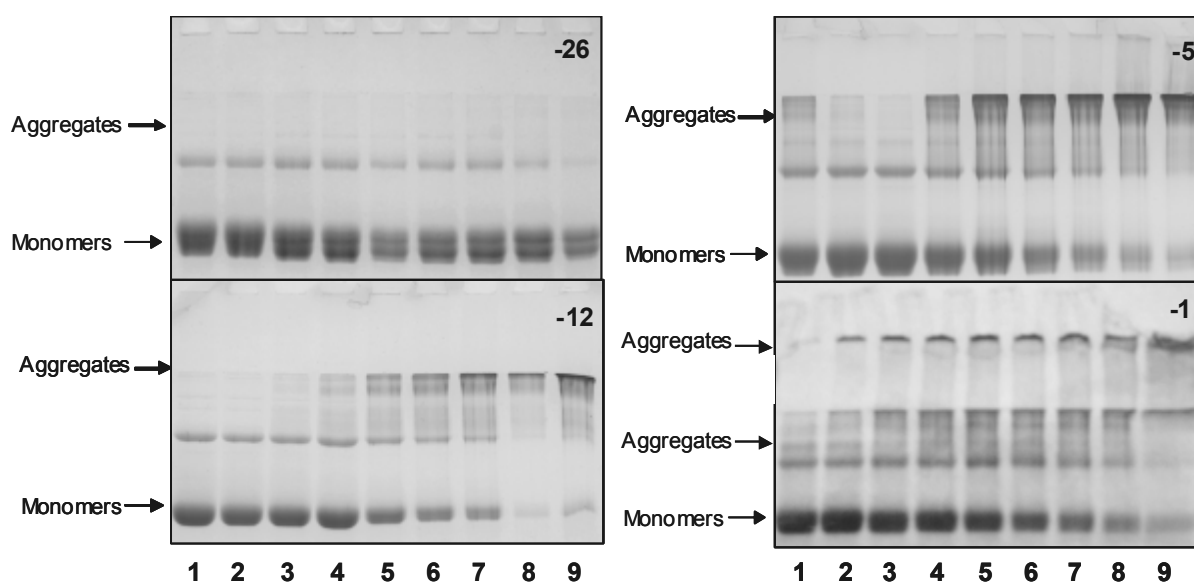
Second, correction of the rate of unfolding (Figure 3) for the observed thermal transition temperature led to the suggestion that variations in net charge do not affect the unfolding rate of ovalbumin other than by affecting the thermodynamic stability. The rates of unfolding, for the OVA variants, calculated from the Arrhenius equation (Figure 2) were in line with the experimentally determined rate constants obtained from the data in Figure 3 (see Weijers et al 2003 for calculations). The theoretical  $k$  values calculated from DSC analysis are in agreement with the experimental  $k$  values obtained from HPSEC. The symbols in Figure 2 represent the experimental  $k$  values at  $7^{\circ}\text{C}$  below  $T_i$  obtained from Figure 3. This finding therefore suggests that the unfolding of all the OVA variants follows similar irreversible first-order kinetics.



**Figure 3** Heating time dependence of the fraction folded OVA for charge engineered OVA variants determined with HPSEC. The OVA solutions were heated at 7°C below their peak temperature ( $T_i$ ).  $T_i$  is tabulated in Table II. Solid line is a guide to the eye.

#### Characterisation of the physicochemical properties of protein aggregates

In the preceding publication by Broersen et al (2005), we already reported that the sulfhydryl-disulphide exchange index and the exposed hydrophobicity of the folded proteins was not affected by the modification. This validates the assumption that the potential for chemical cross-linking of the reactive particles generated by heating was not affected by charge engineering.

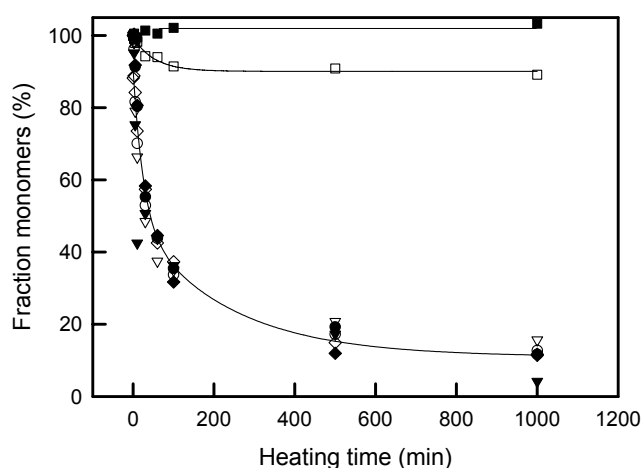


**Figure 4** Non-reducing SDS PAGE of unheated and heated OVA solutions (5 mg/mL, pH 7.0, 72 °C). Lane 1, unheated OVA; lane 2-9, OVA heated for 2, 5, 10, 30, 60, 100, 500 and 1000 minutes.

*Note that the additional protein band in the untreated OVA sample are ovalbumin dimers (this band disappears under reducing conditions).*

### *Covalent interaction upon charge modification*

The combination of non-reducing SDS-PAGE and SDS agarose electrophoresis can provide useful information on the kinetics of the formation of covalent interactions during the aggregation process. Protein aggregates of all OVA variants were prepared by heating 5 mg/mL protein solutions at pH 7.0. Figure 4 shows the kinetics of aggregation using non-reducing SDS-PAGE (Surroca et al 2002, Bartkowski et al 2002). Upon incubation at a fixed temperature of 7°C below the thermal transition temperature determined using DSC, it was observed that for the charge variants OVA-12, OVA-5 and OVA-1 the monomeric protein band decreased in time due to the formation of disulphide linked aggregates. The highly charged OVA-26 showed almost no decrease of the monomeric protein band during heating (upper panel, Figure 4). Apparently, no disulphide cross-linked oligomers were formed for OVA-26 under these conditions. Densitometric analysis of the data in Figure 4, i.e. the fraction non-disulphide cross-linked monomers of the total fraction protein visible in each lane is presented as a function of heating time in Figure 5. The represented kinetics of disulphide bond formation shown in this figure primarily suggests that all OVA variants follow similar covalent aggregation kinetics except for OVA-26. At a higher protein concentration of 10 mg/mL, OVA-26 also shows disulphide bond formation within the time frame studied (SDS-PAGE, results not shown). Figure 4 also shows qualitatively that with decreasing net charge the formed aggregates become significantly larger. For OVA-5, OVA-12 and OVA-26 it was found that during the entire incubation time, all aggregates were still able to enter the gel indicating that the size of the aggregates was generally smaller than 250 kDa. In contrast, the OVA-1 variant also shows aggregates that were not able to enter the gel after heating for only 2 minutes. As it is difficult to analyse the fraction of aggregates that has not entered the gel matrix, this aspect was further studied using SDS agarose electrophoresis, a suitable method to detect much larger aggregates compared to SDS-PAGE (Alting et al 2004). The results from SDS agarose electrophoresis confirmed the data obtained by SDS PAGE that the aggregates formed upon neutralisation of the net charge (OVA-1) were significantly larger than the aggregates formed by the other charge engineered ovalbumins (results not shown).

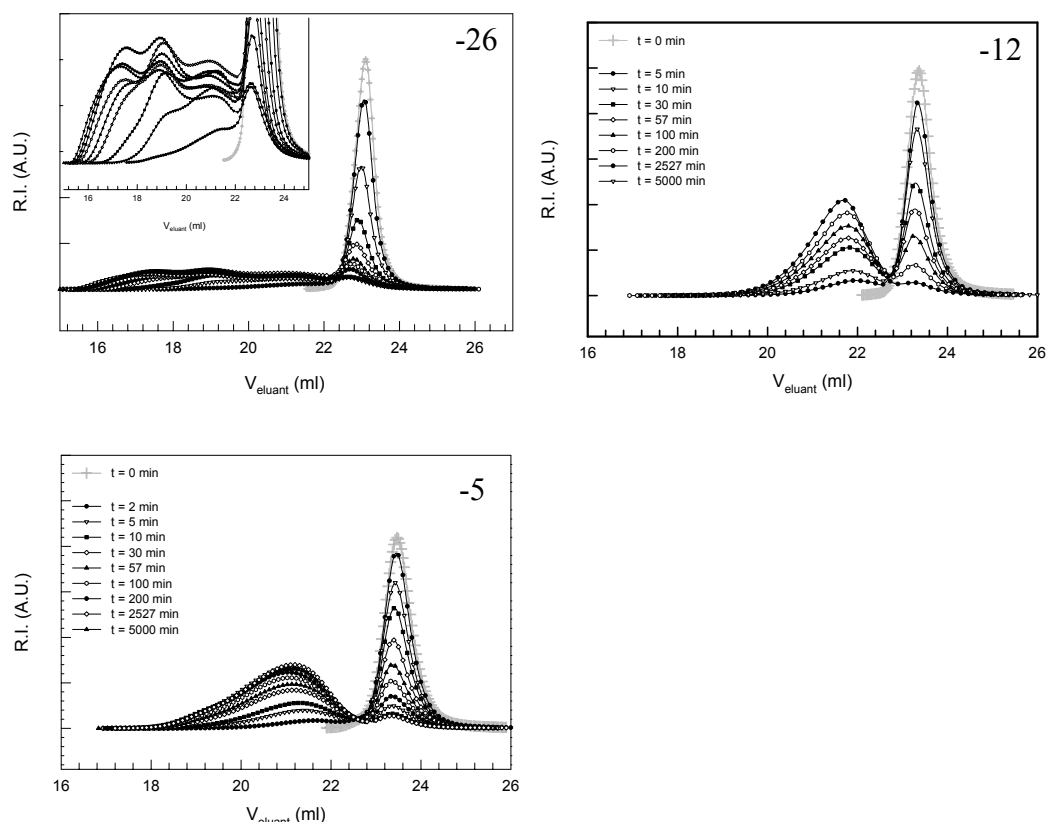


**Figure 5** Densitometric analysis of non-reducing SDS-PAGE data as shown in Figure 4. Filled and open symbols represent OVA variants heated at a protein concentration of 5 mg/mL and 10 mg/mL, respectively. OVA variants: OVA-12 (●,○); OVA-26 (■,□); OVA-5 (▼,▽); OVA-1 (◆,◇). Solid lines are guides to the eye.

The size of the formed aggregates was thus significantly affected by charge variation. The results from gel electrophoresis experiments suggested that a very high net charge can result in a significant delay in disulphide linked aggregation during heat-induced aggregation and that the size of the formed aggregates is also affected by net charge.

#### *Polydispersity of formed aggregates*

Size-exclusion chromatography was used to study the effect of net charge on the polydispersity of the formed aggregates. All OVA variants were heated at 72°C and aliquots were withdrawn at various heating times ranging from 0 to 5000 minutes and analysed by size-exclusion chromatography. Figure 6 shows elution profiles of heated solutions of the OVA variants (OVA-26, OVA-12, and OVA-5). Unfortunately, it was not possible to perform this experiment for the OVA-1 variant.



**Figure 6** Chromatograms of different OVA solutions at 10 mg/mL (pH 7.0) (OVA-26 at 40 mg/mL) at different heating times at 72 °C. The inset is an enlargement of the SEC-MALLS data from the chromatogram of OVA-26. The peak at elution volumes between 22.5 and 24.5 mL corresponds to monomers.

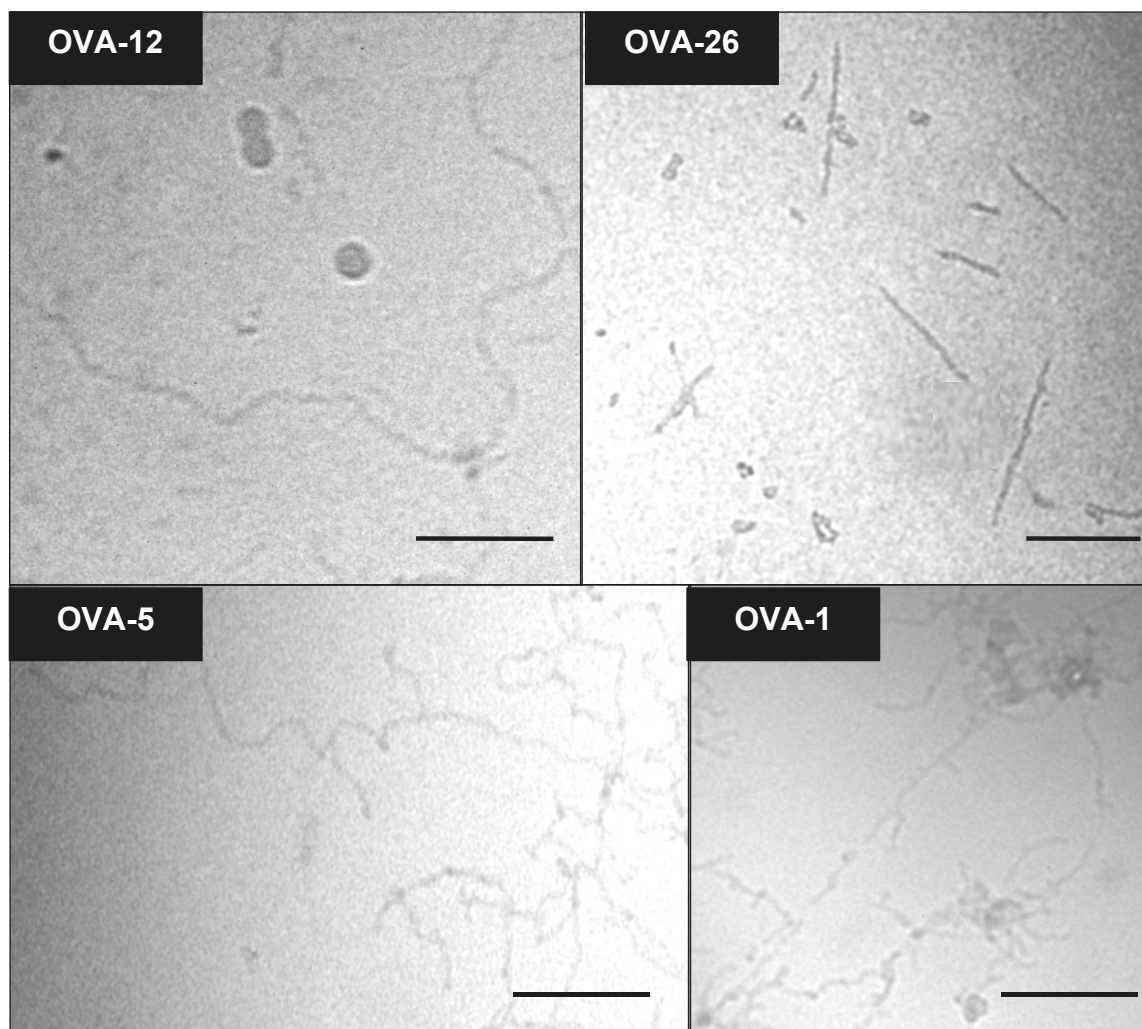
Aggregates of OVA showed a wide size distribution and elute at volumes between 16.0 and 22.5 mL. With increasing heating time, the fraction of monomers decreased as was shown by the decreasing peak height corresponding to monomers and peaks appearing at lower elution volumes. This indicated that the size and fraction of the aggregates increases as a function of heating time. It was observed in the elution profile of OVA-12 heated at 10 mg/mL that the aggregates formed are baseline separated from the monomeric peak. This suggests that under these conditions large, more discrete aggregates were formed and no or only very small fractions of dimers and trimers were formed. This is also the case for OVA-5. However, the aggregates formed for OVA-5 show a larger polydispersity ( $M_w/M_n \sim 3$ ) in size compared to OVA-12. This was shown by the wide range of elution volumes the aggregates are eluted at for OVA-5: from 18.0 to 22.4 mL and for OVA-12 from 20.0 to 22. mL. In the leftupper panel, chromatograms of OVA-26 solutions heated at 40 mg/mL are shown. The

inset demonstrates that no clear distinction can be made between monomers and aggregates, which indicates the formation of oligomers, possibly consisting of dimers and trimers.

#### *Microstructural organisation of aggregates*

Figure 7 shows cryo-TEM images of the microstructural organisation of the aggregates obtained from heating the OVA variants. It was found that net charge has an effect on fibril morphology and contour length ( $l_c$ ). The characteristics derived from analysing the cryo-TEM images of the formed aggregates and the effects of net charge are summarised in Table III. For OVA-1 and OVA-5 the aggregates appear very curved and in some cases, particularly for OVA-1, organised in clusters (lower panel of Figure 7). The aggregates formed by these variants also show a significant degree of branching. Upon increasing the net charge to  $-12$  (OVA-12) the aggregates retain some of their curvature, although it is clearly decreased compared to the OVA-1 and OVA-5 aggregates, but the degree of branching is nearly negligible. The aggregates prepared from OVA-26 were again less curved and appear more or less as short stiff rods without any branching observed. These results suggest that the morphology of aggregates is clearly affected by the net charge of the protein. Apparently, decreasing the net charge results in an increased degree of branching and a decreasing persistence length, which characterises the curvature. The contour length of the aggregates shows an apparent maximum between a net charge of  $-12$  and  $-5$ . At higher net charge the aggregates were significantly shorter and at lower net charge the aggregates became clustered. Therefore a reliable determination of the contour length was not possible.



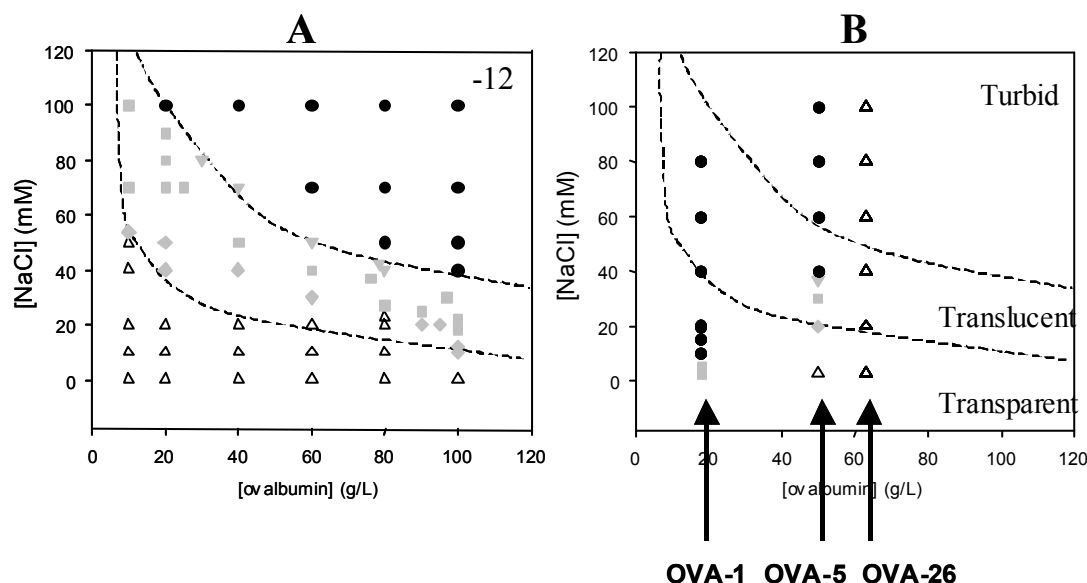


**Figure 7** Microstructural organisation of aggregates formed of OVA aggregates with varying net charge formed investigated by cryo-TEM after heating for 24 h at 72 °C at pH 7.0. OVA variants were initially heated at: 55 mg/mL (OVA-12); 42 mg/mL (OVA-26); 52 mg/mL (OVA-5) and 25 mg/mL (OVA-1), and diluted to 1 mg/mL before cryo-TEM images were recorded. The length of the bar represents 100 nm.

#### *Functional properties of heat-induced ovalbumin gels*

It has been previously shown that the visual appearance of a gel is strongly associated to the organisation of aggregates in the gel. Transparent gels primarily consist of fibrillar structures while turbid gels are built-up from branched structures (Langton & Hermansson 1992). Therefore, measuring turbidity is a convenient technique to obtain an impression of the microstructure of gels. Solutions of OVA-12 were heated at 78°C (pH 7.0) for 1 h and the turbidity ( $\tau$ ) at 500 nm was measured spectrophotometrically as a function of protein concentration. As literature suggested that a high net charge can be significantly shielded by

the addition of a salt we also studied the effect of NaCl concentration. If the effects of the charge variation are directly responsible for the observed effects on aggregate formation and morphology, shielding these differences should lead to the possibility to form similar gels for all ovalbumin variants (Figure 8).



**Figure 8** Diagram of states of heat-induced aggregates of OVA variants determined spectrophotometrically at 500 nm. OVA solutions were heated at 78°C at different protein concentrations and salt concentrations as described in the materials and methods. Symbols: ( $\Delta$ ) transparent solution or gel,  $\tau < 0.4 \text{ cm}^{-1}$ ; ( $\diamond$ ) “low” translucent system,  $0.4 \leq \tau \leq 1.04$ ; ( $\blacksquare$ ) “medium” translucent system,  $1.05 \leq \tau \leq 11.4$ ; ( $\blacktriangledown$ ) “heavy” translucent system,  $11.5 \leq \tau \leq 18.4$ ; ( $\bullet$ ) turbid system,  $\tau > 18.4$ . The boundaries between transparent and translucent and between translucent and turbid are indicated in the figures by dashed lines. a) Diagram of states of OVA-12; b) Turbidity of OVA-26 (62 mg/mL), OVA-5 (50 mg/mL) and OVA-1 (18 mg/mL) gels are incorporated in the diagram of states of OVA-12. Note that for OVA-26, OVA-5 and OVA-1 the turbidity is measured at one protein concentration.

Figure 8a shows the diagram of states of OVA-12 for which the protein concentration was varied between 10 and 100 mg/mL and the NaCl concentration was varied between 0 and 100 mM NaCl. We classified  $\tau$  in three regimes: transparent, translucent and turbid (see Figure 8 and for details materials & methods section). The individual measurements are presented by various symbols as explained in the legend. Figure 8a shows that at all protein concentrations transparent solutions or gels are formed when heated at a NaCl concentration

below 20 mM. Increasing the NaCl concentration resulted in the formation of translucent gels, also dependent on the protein concentration, e.g. up to 40 mM NaCl at 100 mg/mL and up to 80 mM NaCl at a protein concentration of 30 mg/mL. Increasing the NaCl concentration even further led to the formation of turbid solutions or gels. The turbidities of OVA-26, OVA-5 and OVA-1 gels/solutions were determined as a function of NaCl concentration at a single protein concentration of 62, 50, and 18 mg/mL, respectively (Figure 8b). These protein concentrations were chosen such that the change from transparent to turbid could be visualised. Similar heating conditions were used as for OVA-12: 1 h at 78°C and pH 7.0. As expected, the development of turbidity or transparency is significantly affected by the net charge. For example, solutions or gels obtained upon heating OVA-26 remained transparent, even at an NaCl concentration of 100 mM, while OVA-1 is already turbid at 10 mM NaCl.

**Table III** *Aggregate characteristics: contour length ( $L_c$ ) and shape of ovalbumin variants.*

Ovalbumin variant	$L_c$ (nm) <sup>a</sup>	shape <sup>a</sup>
OVA-26	20 - 200	linear rods
OVA-12	400 - 1100	curved fibrils
OVA-5	400 - 1000	curved, branched fibrils
OVA-1	n.d.	clusters of branched aggregates

<sup>a</sup>Determined by analysis of cryo-TEM images.

n.d.: Not determined.

## Discussion

### *Protein unfolding kinetics is affected by net charge*

To obtain a better understanding of the underlying mechanism of fibril formation and morphology, the effect of net charge on the kinetics of unfolding, fibril formation and morphology and their resulting networks was studied. OVA variants were produced with varying net charges at pH 7.0, ranging from −1 to −26. The unfolding kinetics of these OVA variants followed first-order kinetics and obeyed a ‘two-state’ model for irreversible unfolding with the exception of OVA-1, which has been found in the past to significantly populate thermodynamically stable intermediates (Figures 1 and 2 and Broersen et al 2005). It was found by Broersen et al 2005 that the intermediate state can also participate in the aggregation process and thus that the loss of a folded conformation is more or less

synonymous for the generation of aggregation prone particles. Also in this publication we also found no indication that the generation of an intermediate state affected the generation of particles that can participate in the aggregation process (Figure 3).

Charge modification generally led to changes in thermal stability, i.e. both  $E_a$  and  $T_t$  were affected by charge modification (Figures 1 and 2). This finding is consistent with many other reported observations on the effect of charge on protein stability for a wide range of proteins (Loladze et al 1999, Frankenberg et al 1999, Sali et al 1991). Net charge was however not linearly correlated with the activation barrier and  $T_t$  and knowledge of the net charge thus cannot be used to predict thermal stability in a straightforward manner (Table II). That the relation between charge and stability is more complex has also been reported repeatedly before (Xiao & Honig 1999, Godoy-Ruiz et al 2004, Loladze et al 1999). Upon heat-induced unfolding, all variants appeared to have similar unfolding kinetics at  $7 \pm 1^\circ\text{C}$  below their  $T_t$  (Figure 3). These observations suggest that, even though the stability of the variants is significantly affected by the modification of charge, no indications were found to suggest that the unfolding mechanism upon charge modification differed.

#### *Protein aggregation can be delayed but is not prevented by a high net charge*

The effect of net charge on the kinetics of the formation of disulphide interactions and a number of parameters related to aggregate morphology such as size and curvature of the aggregates was investigated using a combination of SDS-PAGE, SDS-agarose electrophoresis, SEC-MALLS and cryo-TEM (Figures 4, 6 and 7). We found that generally charge modification on the protein molecule does not result in significant changes in the rate of disulphide bond formation during heat-induced aggregation (Figures 4 and 5). In contrast, the OVA-26 variant showed a significant inhibition of the rate of covalent aggregate formation (Figure 5). The slow rate at which disulphide bonds were formed for OVA-26 was consistently found using SDS-PAGE and SEC-MALLS, and is likely to be due to the inability of disulphide bonds and sulfhydryl groups to approach each other close enough as a result of the high electrostatic repulsion the molecules experience in solution. This suggestion was further confirmed by the observation that it was possible to form aggregates upon increasing either the protein concentration or the salt concentration. The latter option leads to a shielding of the electrostatic repulsion (Figure 5). Apparently, the formation of disulphide bonds requires a close proximity of the reactants (disulphide bonds and sulfhydryls). Also, the process of disulphide bond formation proceeds slowly as a signature of its high activation

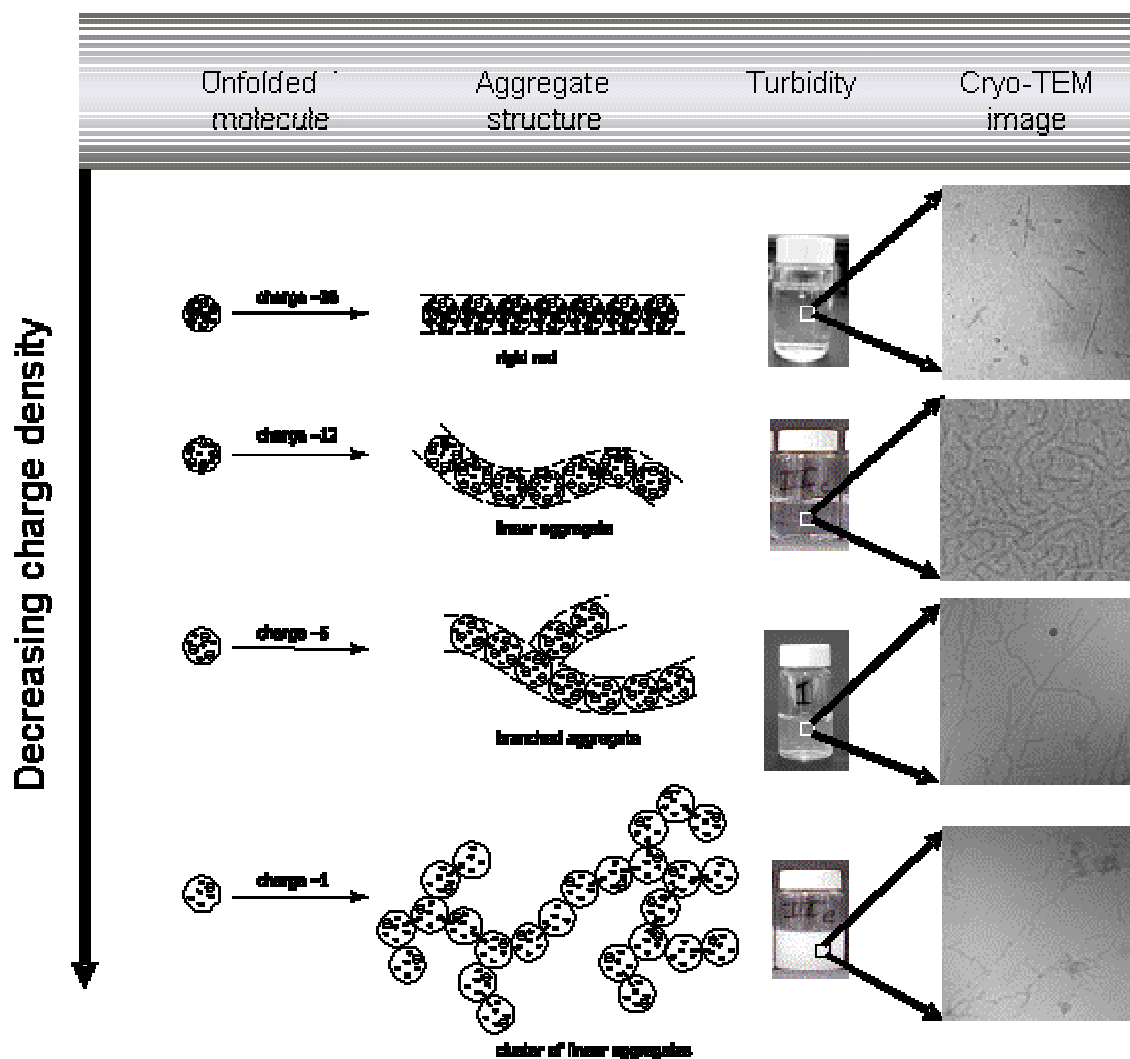
energy. As was demonstrated for the formation of nanotubes of  $\alpha$ -lactalbumin before (Graveland-Bikker et al 2004), the absence of disulphide bonds in aggregates does not automatically imply that the aggregation process is altered. As we found using cryo-TEM and SEC-MALLS, a strong electrostatic repulsion may slow down the aggregation process but does not automatically prevent it altogether (Figures 6 and 7). This has been demonstrated in the past using  $\beta$ -lactoglobulin. At pH 2, where the net charge of this protein is  $\sim +20$ ,  $\beta$ -lactoglobulin forms rigid rods as observed by cryo-TEM (Aymard et al 1999) without any evidence of the formation of intermolecular disulphide bonds. Clearly, the formation of disulphide interactions is not necessarily a prerequisite for fibril formation. An aspect of fibril formation that requires consideration is the effect of charge modification on the propensity to form secondary structure. This aspect has not been investigated in this paper, but it cannot be ruled out that the propensity to form secondary structure has been affected by the charge modification.

#### *Net charge strongly determines aggregate morphology*

SEC-MALLS (Figure 6) and cryo-TEM (Figure 7) data showed that an increasing net charge results in the formation of gradually smaller aggregates with a wider aggregate size distribution (Figure 6). Furthermore, cryo-TEM images showed that, next to size, the shape of the aggregates, such as curvature and degree of branching, was also significantly influenced by net charge. For example, heat-induced aggregation of OVA-26 resulted in the formation of a mixture of short and long rigid rods. On the other hand, heat-induced aggregates of OVA-1 appeared as clusters of heavily branched aggregates. By increasing the net charge on a protein molecule the persistence length increases. Stiff fibrils were formed, while a lower net charge results in the formation of more branched and clustered fibrils. Since OVA-1 still forms linear aggregates, we suggest that not only the net charge by itself but maybe also a factor such as the distribution of this charge on an unfolded protein molecule may play a crucial role in determining fibril forming properties. The importance of electrostatic repulsion in fibril formation is further supported by the observation that increasing the negative charge on OVA-12 fibrils *after* their formation (post-succinylation) resulted in the transformation and fragmentation of the formed aggregates into stiffer and shorter fragments closely resembling the aggregates formed by OVA-26 (unpublished results). This suggests that strong electrostatic repulsion is sufficiently energetically unfavourable to transform readily formed aggregates into shorter and stiffer fibrils.

*Net charge determines the formation and visual appearance of gels*

The ability of proteins to form fibrillar structures has been related in the past to the formation of transparent gels (Langton & Hermansson 1992). In this work we have shown that heat-induced fibril formation is strongly affected by protein net charge. This observation led to the hypothesis that the visual appearance of gels should also be affected by net charge. Heat-induced aggregate solutions and gels prepared from OVA-26 remain transparent over a wide range of salt concentrations (Figure 8). As OVA-12 is significantly less charged, within line of the expectations, transparent aggregate solutions or gels are only formed at low ionic strength (see also Pouzot et al 2005). Decreasing the net charge further (OVA-5) led to even lower ionic strengths which can be applied to form transparent gels/solutions and for OVA-1 it was not possible at all to form transparent systems either in the presence or absence of NaCl. This indicates that the characteristic dimension of the structure of these gels is in the order or larger than the order of the wavelength of visible light. Thus the ability to form transparent gels depends on the properties of the fibrils, such as contour length ( $L_c$ ), rigidity, degree of branching and spatial organisation and all these properties have been shown to be controlled by net charge.



**Figure 9** The proposed general aggregation scheme for the thermal unfolding and fibril formation of OVA with various charge densities on the surface of the protein molecule. Shadows indicate hydrophobic areas exposed by thermal unfolding; minus signs ( $\ominus$ ) and dashed symbols (—) indicate the calculated net charge at pH 7.0 and disulphide interactions, respectively. Turbidity and cryo-TEM images are illustrations of experimental results.

Summarising, the charge density on the protein molecule has a major effect on a number of parameters related to aggregate formation. First, net charge significantly affects protein stability (Broersen et al 2005). However, upon correcting the rate of unfolding for the differences in the thermal transition temperature, the rate of unfolding is not significantly affected anymore by the charge modification itself. A significant increasing of the net charge however limited the formation of disulphide interactions. The finding that the rate of aggregation is not necessarily coinciding with the formation of disulphide bonds has been

suggested before (Broersen et al 2005b). The morphological character of the formed aggregates is also affected by net charge. A high net charge leads to fibrillar structures, which are visually transparent over a wide range of ionic strengths. We suggest that a high net charge decreases the possible orientations a protein molecule can assume when organising into an aggregate. A high net charge will place a significant electrostatic constraint on a number of the orientations. A low net charge generally resulted in the formation of turbid systems, also at low ionic strength. To account for the formation of linear rods as observed in this work, the existing scheme (Koseki et al 1989) has been extended, shown in Figure 9. This scheme shows that a highly charged protein molecule results in the formation of heat-induced linear rods which appear transparent and are visualised by cryo-TEM. The other extreme is when almost all charge is eliminated from the protein. The heat-induced aggregates formed from this virtually (net) uncharged protein, are branched to a high degree and clustered due to a low electrostatic repulsion. This consequently leads to the formation of turbid systems already at low ionic strength. It can be concluded that controlling the net charge on the protein molecules, i.e. the electrostatic repulsion, allows us to control fibril formation and therefore the turbidity of OVA networks. This knowledge may provide an important tool for the mechanistic understanding of fibril formation in a variety of (biological) fields, i.e. amyloid formation, as well as on the manipulation of biotechnological processing techniques.

## **Materials and methods**

### *Reagents and Chemicals*

Ovalbumin was purified from fresh hen eggs according to a previously described purification protocol (Kosters et al 2003). A method to chemically modify the OVA variants (OVA-26, OVA-5, and OVA-1) is described in the preceding paper (Broersen et al 2005).

### *DSC*

The experimental method has been described elsewhere (Broersen et al 2005). Samples were prepared at protein concentrations of 10 mg/mL, pH 7.0 in 10 mM sodium phosphate buffer. The same solution without the protein was used in the reference cell. The temperature was scanned from 25 to 120°C at a scanning rate of 1.00°C min<sup>-1</sup>. To obtain  $C_p^{\text{eff}}$  curves, reference-reference baselines were obtained at the same scan rate and subtracted from the sample curves. Subsequently, a baseline correction was applied assuming  $\Delta C_p=0$ . The fits



have been performed, using the theory described by Sanchez-Ruiz (Sanchez-Ruiz et al 1988) for a practical two-state model.

### *HPSEC*

For the determination of the fraction folded OVA (OVA-12, OVA-26, OVA-5, and OVA-1) upon heating, different series of screw-cap vials containing ca. 5 mL of OVA solution with different initial protein concentrations were heated for various times up to 24 h at 72.0, 66.8, 72.0, and 69.4°C, respectively (corresponding to 7°C below their actual peak temperature, determined with DSC) at pH 7.0. It has been verified before that ovalbumin concentration does not affect the observed thermal transition temperatures using DSC (Weijers et al 2003). Next, the vials were cooled in ice-water and the protein solutions were diluted to a final concentration of folded protein between 0.1 to 5 mg/mL, to be within the calibration range. It has been verified before that cooling results in either refolding or aggregation of ovalbumin and that the refolded fraction closely resembles the native-like state (unpublished results). After centrifugation at 20,000g for 5 minutes at room temperature, the concentration of folded OVA in the supernatant was determined by HPSEC (Phenomenex BioSep-SEC-S2000 column, 300×7.5 mm) with UV detection at 280 nm.

### *SDS-PAGE*

SDS-PAGE was performed to follow the formation of disulphide cross-linked aggregates during heating. The experimental method is described elsewhere (Broersen et al 2005).

### *Agarose Gel Electrophoresis*

SDS-agarose gel electrophoresis (1 % (w/w) agarose) was performed to determine the differences in molecular weight of the different variants of OVA aggregates prepared just below their critical gelation concentration. The experimental method is described elsewhere (Alting et al 2000). Samples were prepared at protein concentrations just below their critical gelation concentration: OVA-26 42 mg/mL, OVA12 55 mg/mL, OVA-5 52 mg/mL, OVA-1 25 mg/mL.

### *SEC-MALLS*

SEC-MALLS experiments were performed as described before to determine the total amount of monomeric protein as well as the amount of aggregates (Weijers et al 2003). Samples were prepared at protein concentrations of 5 and 10 mg/mL at pH 7.0.

### *Cryo-TEM*

Cryo-transmission electron microscopy was carried out as described elsewhere (Weijers et al 2002), using a Philips CM12 transmission electron microscope operating at 120 kV. Images were recorded digitally by a Gatan 791 CCD camera using the Digital Micrograph software package. Samples were prepared at protein concentrations ( $C^*$ ) as mentioned in Table III, pH 7.0.

### *Turbidity measurements*

Turbidity measurements were performed at 78°C on a Cary 1E UV-vis spectrophotometer (Varian Nederland B.V.) equipped with a temperature controller. OVA samples were prepared at a range of protein concentrations (5- 100 mg/mL) and ionic strengths (3 – 500 mM) at pH 7.0. The turbidities of the OVA solutions and gels were measured in quartz cuvettes with a path length of 0.2 cm from steady-state absorbance at 500 nm (typically after 1 h heating). The diagram of states (Figure 8) was divided into three regimes: transparent, translucent and turbid. These samples were defined transparent when the turbidity was lower than  $0.4 \text{ cm}^{-1}$ , i.e. more than 92% of the light passes through the sample in a 0.2 cm sample cell. Translucent sample have turbidities between 0.4 and  $18 \text{ cm}^{-1}$ , i.e more than 2.5% and less than 92% of the light passes the sample. Samples with turbidity higher than  $18 \text{ cm}^{-1}$  were defined as turbid, i.e. less than 2.5% of the light passes through the sample.

### **Acknowledgements**

The authors acknowledge Anne van de Pijpekamp for performing the turbidity measurements and Hans Kusters for determining the sulfhydryl-disulphide exchange rate.

### **References**

- Alting AC, Hamer RJ, de Kruif CG, Visschers RW (2000) J Agric Food Chem 48: 5001-5007
- Alting AC, Weijers M, Hoog de EHA, van de Pijpekamp AM, Cohen Stuart MA, Hamer RJ, Kruif de CG, Visschers RW (2004) J Agric Food Chem 52: 623-631
- Aymard P, Nicolai T, Durand D, Clark A (1999) Macromolecules 32: 2542

- Bartkowski R, Kitchel R, Peckham N, Margulis L (2002) *J Protein Chem* 21: 137-143
- Broersen K, Weijers M, de Groot J, Hamer RJ, de Jongh HHJ (2005) Electrostatics controls fibril formation. Part I: Charge engineering of proteins affects unfolding mechanism (Preceding paper)
- Broersen K, van Teeffelen A, Vries A, Voragen AGJ, Hamer RJ, de Jongh HHJ (2005b) Do sulfhydryl groups affect aggregation and gelation properties of ovalbumin? Manuscript in preparation.
- Donovan JW, Beardslee RA (1975) *J Biol Chem* 250: 1966-1971
- Ellis RJ, Pinheiro TJT (2002) *Nature* 416: 483-484
- Frankenberg N, Welker C, Jaenicke R (1999) *FEBS Lett* 454: 299-302
- Godoy-Ruiz R, Perez-Jimenez R, Ibarra-Molero B, Sanchez-Ruiz JM (2004) *J Mol Biol* 336: 313-318
- Gosal W S, Clark AH, Pudney PDA, Ross-Murphy SB (2002) *Langmuir* 18: 7174-7181
- Graveland-Bikker JF, Ipsen R, Otte J, de Kruif CGJ (2004) *Langmuir* 20: 6841-6846
- Kallberg Y, Gustafsson M, Persson B, Thyberg J, Johansson J (2001) *J Biol Chem* 276: 12945-12950
- Klibanov AM, Ahern TJ (1987) In *Protein Engineering*. (Oxender DL, Fox CF, eds) Alan R Liss, New York, pp 213-218
- Koo EH, Lansbury PT Jr, Kelly JW (1999) *Proc Natl Acad Sci* 96: 9989-9990
- Koseki T, Kitabatake N, Doi E (1989) *Food Hydrocolloids* 3: 123-134
- Kosters H, Broersen K, Groot de J, Simons J-WFA, Wierenga P, de Jongh HHJ (2003) *Biotech Bioeng* 84: 61-70
- Langton M, Hermansson AM (1992) *Food Hydrocol* 5: 523-539
- La Rosa C, Milardi D, Grasso DM, Verbeet MP, Canters GW, Sportelli L, Guzzi R (2002) *Eur Biophys J* 30: 229-570
- Loladze VV, Ibarra-Molero B, Sanchez-Ruiz JM, Makhatadze GI (1999) *Biochemistry* 38: 16419-16423
- Lopez De La Paz M, Goldie K, Zurdo J, Lacroix E, Dobson CM, Hoenger A, Serrano L (2002) *Proc Natl Acad Sci USA* 99: 16052-16057
- Ma C-Y, Holme J (1982) *J Food Sci* 47: 1454-1459
- Ma C-Y, Poste LM, Holme J (1986) *Can Inst Food Sci Technol J* 19: 17-22
- Mine Y (1996) *J Agric Food Chem* 44: 2086-2090
- Nakamura R, Sugiyama H, Sato Y (1978) *Agric Biol Chem* 42: 819-824
- Odijk T, Mandel M (1978) *Physica A* 93: 298-306
- Odijk T (1978) *Polymer* 19: 989-990
- Odijk T (1977) *J Polymer Sci* 15: 477-483
- Pouzot M, Nicolai T, Visschers RW, Weijers M (2005) *Food Hydrocolloids* 19: 231-238
- Raman B, Chatani E, Kihara M, Ban T, Sakai M, Hasegawa K, Naiki H, Rao CM, Goto Y (2005) *Biochemistry* 44: 1288-1299
- Sali D, Bycroft M, Fersht AR (1991) *J Mol Biol* 220: 779-788
- Sanchez-Ruiz JM, Lopez-Lacomba JL, Cortijo M, Mateo PL (1988) *Biochemistry* 27: 1648-1652
- Shah JV, Janmey PA (1997) *Rheol Acta* 36: 262-268
- Stein PE, Leslie AG, Finch JT, Carrell RW (1991) *J Mol Biol* 221: 941-959
- Surroca Y, Haverkamp J, Heck AJR (2002) *J Chrom A* 970: 275-285
- Weijers M, Visschers RW, Nicolai T (2002) *Macromolecules* 35: 4753-4762

Weijers M, Barneveld PA, Cohen Stuart MA, Visschers RW (2003) *Protein Sci* 12: 2693-2703

Xiao L, Honig B (1999) *J Mol Biol* 289: 1435-1444

Zbilut JP, Giuliani A, Colosimo A, Mitchell JC, Colafranceschi M, Marwan N, Webber CL, Uversky VN (2004) *J Proteosome Res* 3: 1243-1253

Zurdo J, Guijarro JI, Jiménez JL, Saibil HR, Dobson CM (2001) *J Mol Biol* 311: 325-340

## **PART III**

### **DISULPHIDE BONDS AND PROTEIN AGGREGATION**



# CHAPTER 8

## SULFHYDRYL GROUPS DO NOT AFFECT THE HEAT-INDUCED AGGREGATION RATE OF OVALBUMIN BUT DO AFFECT AGGREGATE MORPHOLOGY

Broersen K, van Teeffelen A, Vries A, Voragen AGJ, Hamer RJ, de Jongh HHJ

This chapter is part of a manuscript in preparation ‘Do sulfhydryl groups affect aggregation and gelation properties of ovalbumin?’

### **Abstract**

The aim of this work is to evaluate the relative impact of sulfhydryl groups on protein aggregation, which is currently unknown. Purified hen egg ovalbumin was chemically engineered to add sulfhydryl groups on the protein surface in various degrees. Even though the structure of ovalbumin was destabilised upon modification, as was clear from a lower thermal transition temperature, the secondary and tertiary structure remained intact at ambient temperature. Also, calorimetric studies showed an enthalpic transition upon heat-induced unfolding suggesting that the structure upon modification still displays stability. The rate of aggregation was not affected by the introduction of reactive sulfhydryl groups. The results also show that disulphide bond formation was preceded by physical interactions upon aggregation. Hence, disulphide interactions are not the driving force for aggregation of ovalbumin. Using transmission electron microscopy it was observed that the microstructure of the aggregates had shifted from a fibrillar, linear aggregate system with the addition of up to four sulfhydryl groups towards a curved, highly branched and random aggregate system with the addition of more than four sulfhydryl groups. The reported findings suggest that, even though the aggregation rate is not affected by the presence of more sulfhydryl groups, the resulting aggregates clearly show a different morphology.

### **Keywords:**

Ovalbumin, aggregation, sulfhydryl groups, disulphide bonds, cryo-TEM

## Introduction

Heat treatment has considerable implications for the behaviour of proteins in food systems (Koseki et al 1989). One example of the consequences of heat processing is protein aggregation, which is a major determinant for the structural properties of food products. Egg white proteins are used frequently in the food industry for this purpose (Sun & Hayakawa 2002). Ovalbumin is the most abundant protein of egg white and contributes significantly to the functional properties of egg white protein preparations (Sun & Hayakawa 2002). Ovalbumin has one disulphide bond and four sulfhydryl groups buried in the interior of the protein (Iametti et al 1998).

A first step in the aggregation process is the (partial) unfolding of protein molecules upon heating. The conformational changes that are associated to unfolding result, among other changes, in solvent exposure of hydrophobic and sulfhydryl groups. Subsequent aggregation involves non-covalent or intermolecular disulphide (covalent) interactions (Kato et al 1990). Aggregate formation is thus the result of a combination of hydrophobic and electrostatic (non-covalent) interactions and, in some cases, disulphide bonds (covalent interactions) can be present (Kato et al 1983, Koseki et al 1989, Sun & Hayakawa 2002). In the past, kinetic models have been proposed to explain for some proteins the formation of disulphide-linked aggregates through a disulphide-sulfhydryl exchange reaction (Fava et al 1957). It is suggested that the sulfhydryl group acts as an oxidising agent to the intramolecular disulphide bond leading to the reduction of the disulphide bond. This suggestion is based on the observation that, upon disulphide linked aggregation, the total number of sulfhydryl groups is not affected. The resulting liberated sulfhydryl group may then undergo disulphide bridging with a sulfhydryl group of a proximate unfolded polypeptide chain. This reaction then proceeds further in analogy with a radical polymerisation reaction (Verheul et al 1998). Clear evidence on the relative contribution of this mechanism to the aggregation process has, however, not been provided up to now as other, non-covalent forces have been identified that are potentially supporting the development of aggregates (McSwiney et al 1994).

In the past, various approaches have been employed to evaluate the importance of sulfhydryl groups in the aggregation process including the reduction or chemical blockage of sulfhydryl groups (Hoffman & van Mil 1997, Sawyer 1967, Owusu Apenten 1998, Tanaka et al 1996) as well as activation of sulfhydryl groups by unfolding of the protein (Phillips & Jenness 1967, Xiong et al 1993, Shimada & Cheftel 1989). However, as aggregation of proteins is preceded by unfolding, using (partially) unfolded molecules may influence the kinetics of aggregate formation and hence, not provide direct information on the effect of



sulfhydryl groups. Also, the presence of chemicals such as DTT and NEM are known to affect the aggregation properties other than only by means of interruption of disulphide bond formation (e.g. Alting et al 2000). Our approach involving the covalent attachment of acetylthiogroups (that can be converted into reactive sulfhydryls on command through the cleavage of the acetylgroup) to the surface of folded protein molecules enables us to investigate the relative importance of sulfhydryl groups in the formation of aggregates while retaining as much as possible other protein characteristics that potentially affect the aggregation process. This allows a direct comparison of the effect of the introduced sulfhydryl groups as opposed to the combined effect of covalent and non-covalent interactions. As such, the approach of chemical introduction of additional sulfhydryl groups is used in this work to investigate the role of these sulfhydryl groups in the process of aggregation formation and aggregation kinetics.

## Results

### *Preparation and fractionation of modified ovalbumin fractions*

To modify ovalbumin, various ratios of S-AMSA:lysine were used to produce a series of ovalbumins with different numbers of additional sulfhydryl groups (Table I).

**Table I** *Physicochemical characteristics of modified ovalbumin fractions – degree of modification (% modified lysines of the total number of lysines available), isoelectric point and thermal transition temperature*

	Number of sulfhydryl groups (Ellman assay)	Number of modified groups (OPA assay)	Degree of modification (%)	Isoelectric point (IEF)	Transition temperature (°C, DSC) SX/SH
Unmodified	0.9 ± 0.1	0	0	4.7	78.6
SX/SH1	1.8 ± 0.3	1.41 ± 0.02	7	4.4	n.d./n.d.
SX/SH4	3.7 ± 0.4	3.9 ± 0.4	18	4.3	n.d./n.d.
SX/SH6	5.4 ± 0.5	5.5 ± 0.1	26	4.2	73.9/72.7
SX/SH8	9.2 ± 0.9	8.1 ± 0.8	38	4.0	73.2/66.5
SX/SH10	10.1 ± 0.6	10.0 ± 0.3	47	3.9	n.d./n.d.

n.d. Not determined

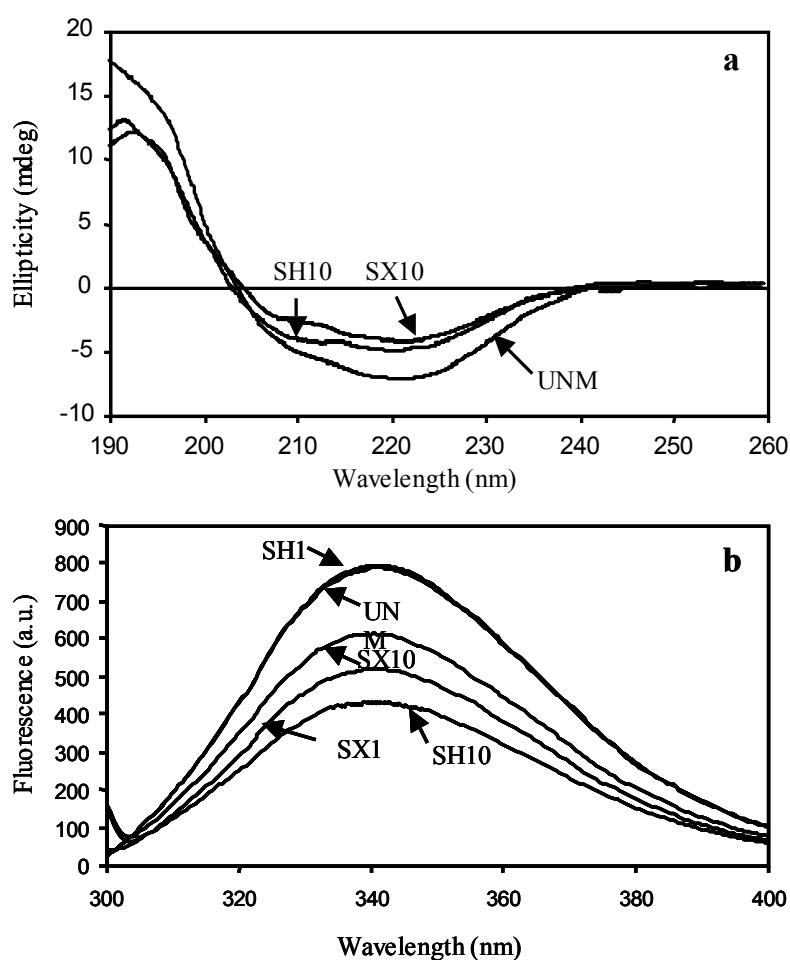
Using the OPA assay (sensitive to free amino groups) and the Ellman assay (sensitive to free thiol groups) it was found that the number of additional sulfhydryl groups ranged from zero to ten (Table I). The results obtained from Ellman's assay are in agreement with the results from the OPA assay (Table I). The degree of modification increased with increasing the S-AMSA:lysine ratio. A positive correlation between the ratio S-AMSA:lysine and the degree of modification has also been found by Strange et al 1993, Kim et al (1990) and Klotz and Heiney (1962). As a consequence of the reaction of lysine with S-acetylmercaptosuccinic anhydride (S-AMSA), acetylthiogroups are formed (Klotz & Heiney 1962). The sulphur group is blocked by an acetyl group. Reaction of S-AMSA with the lysines also results in the introduction of an additional carboxyl group. This additional carboxyl group can be used to estimate the degree of heterogeneity of the modification as well as to limit the heterogeneity of the formed proteins by fractionation using ion exchange chromatography. The batches produced with different S-AMSA:lysine ratios showed heterogeneity as was established by a significant broadening of the elution profiles obtained from ion exchange chromatography, compared to unmodified ovalbumin (results not shown). By fractionation using ion exchange chromatography we were able to isolate more homogeneous fractions ( $\pm 1$  acetylthiogroup) (results not shown). The protecting acetyl groups were cleaved off to obtain reactive sulfhydryl groups just prior to usage in order to limit the extent of auto oxidation of unprotected sulfhydryl groups as has been reported before (Kim et al 1990b). In doing so, we obtained ovalbumins within each degree of modification which contained either acetylthiogroups (SX) or (SH) sulfhydryl groups. The fractions were annotated numbers that reflect the number of introduced sulfhydryl groups per protein molecule (e.g. SX8 or SH8).

The isoelectric point decreases with increasing degree of modification from pI 4.7 for unmodified ovalbumin down to pI 3.9 for SX10/SH10 (Table I). The change in pI results from the parallel introduction of one carboxyl group with each additional sulfhydryl group and the concomitant loss of a positive charge of the reacted lysine.

#### *Structural characterisation of modified ovalbumin fractions*

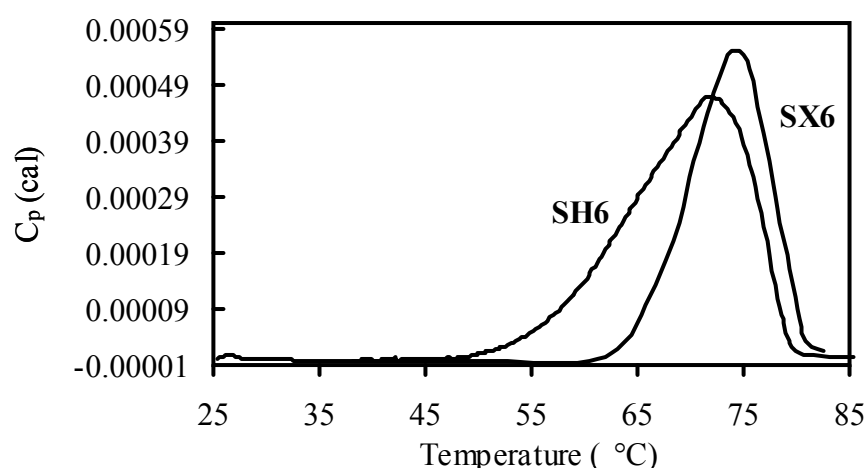
Fluorescence and far-UV circular dichroism (CD) spectra were recorded of all SX and SH fractions produced to determine the impact of the modification on the secondary and tertiary structure of ovalbumin. Figure 1a shows the far-UV CD spectra of unmodified, SX10

and SH10 ovalbumin, the latter two being the highest degrees of modification. Even though some minor changes can be observed, particularly in intensity of the spectra, it was deduced from these spectra (Figure 1a) that the secondary structure was not significantly affected by the modification procedure as the zero-crossing and shape of the spectra remained similar for all modifications. The far-UV CD spectra of the other ovalbumins also did not provide indications for structural changes compared to unmodified ovalbumin (results not shown). Analysis of the spectra using the method by de Jongh et al (1994) provided estimates of the relative contribution of random coil elements, which were found not to differ significantly from the unmodified ovalbumin for the modified fractions. For example, the percentage of random coil calculated from the spectra varied between 24 and 27% random coil for the unmodified and modified ovalbumin variants.



**Figure 1** Structural characterisation of ovalbumin fractions a) Far-UV circular dichroism spectra of unmodified, SX10 and SH10 ovalbumin. b) Intrinsic tryptophan fluorescence spectra of unmodified, SH1, SX1, SH10 and SX10 ovalbumin.

Typical fluorescence spectra of unmodified ovalbumin, SH1, SX1, SH10 and SX10 are shown in Figure 1b. The intensity of the fluorescence shows a significant difference upon modification, which appeared not to be related to the degree of modification (results not shown). It was also found that the wavelength with the maximum fluorescence intensity (340 nm) did not shift upon modification nor the shape of the spectra differed. This suggests that the tertiary structure of the ovalbumin fractions was retained. As a variation in fluorescence intensity may be the result of a different energy transfer efficiency between tryptophan and tyrosine, fluorescence data obtained at an excitation wavelength of 295 nm and 275 nm were compared. The results gave no indication that the globular structure of the ovalbumin fractions was affected by the modification (results not shown). Near-UV CD spectra of the ovalbumin fractions also indicated that the tertiary structure was not affected upon modification (results not shown). From these results it was concluded that introduction of zero to ten additional sulfhydryl groups did not result in a significant loss of structural properties of ovalbumin.



**Figure 2** *Differential scanning calorimetry of ovalbumin fractions. Heat capacity with increasing temperature of SX6 and SH6 ovalbumin.*

The thermal transition temperatures of the SH/SX fractions were determined using differential scanning calorimetry (DSC). The thermograms of SX6 and SH6 ovalbumin are shown as examples in Figure 2. As illustrated in this figure, the variants are able to undergo an unfolding transition as can be deduced from the endothermic contribution at the thermal transition temperatures. The fraction with six acetylthiogroups has a thermal transition temperature of 73.9°C while removal of the acetyl groups resulted in a lower thermal

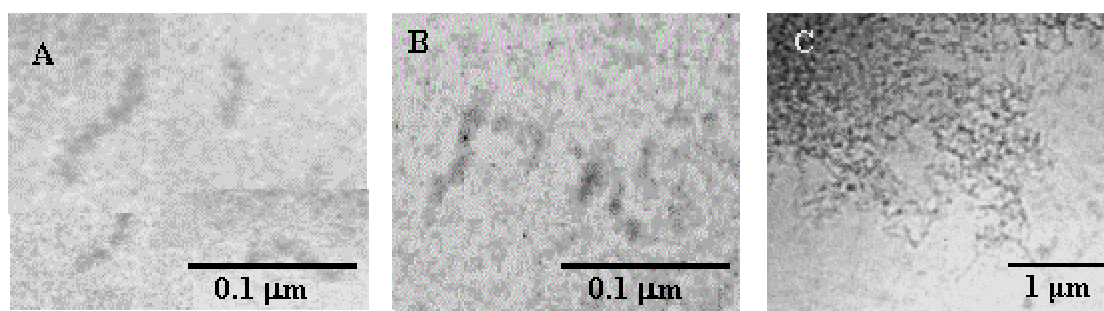
transition temperature of 72.7°C. The introduction of eight groups resulted in a further reduction of the thermal transition temperature down to 66.5°C (Table I). It has been established before that charge modification can significantly contribute to the observed thermostability effects (Broersen et al 2005).

#### *Correction for net charge to study aggregation*

As net charge can affect the unfolding kinetics (Broersen et al 2005) and aggregation properties of ovalbumin (Weijers et al 2005), the fractions with acetylthiogroups (SX) were directly compared to the reactive fractions (SH) with the same degree of modification, thus correcting for the effect of net charge. Next to this, 0.15 to 0.25 M NaCl was added and found to be sufficient to cancel out the side-effect of the introduction of carboxyl groups in addition to sulfhydryl groups based on calculations using the equations by Wu et al (1998). For this reason, most aggregation experiments, which will be discussed in the following paragraphs, have been carried out in the presence of these concentrations of NaCl.

#### *The effect of sulfhydryl groups on the microstructure of aggregates*

To study the effect of sulfhydryl groups on the morphology of the formed aggregates, transmission electron microscopy was used to obtain information concerning the organisation of aggregates. The aggregates were prepared by heating protein solutions of 0.5 mg/mL in the presence of 0.15 M NaCl for 1 h at 90°C at pH 7.0 (Figure 3). Aggregates of ovalbumin with the acetylthiogroups (Figure 3a) are characterised by short and unbranched fibrillar aggregates with a homodisperse size distribution. Aggregates produced under similar conditions but containing four reactive sulfhydryl groups are slightly longer, homogeneously distributed in size, and also unbranched (Figure 3b). Upon introducing ten additional sulfhydryl groups, curved aggregates are formed with a high degree of branching (Figure 3c). Clearly, the degree of modification has an impact on the microstructure of the aggregates.

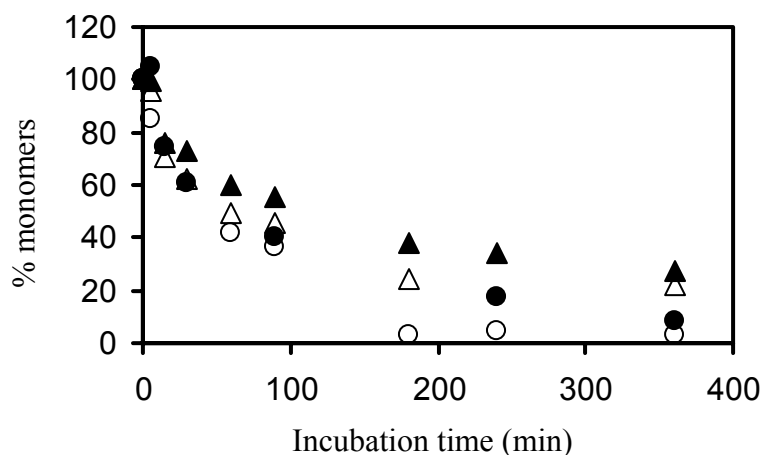


**Figure 3** Cryo-TEM micrographs of aggregates prepared from ovalbumin with additional sulfhydryl groups. a) TEM micrograph of SX1 aggregates in the presence of 0.15 M NaCl. b) TEM micrograph of SH4 aggregates in the presence of 0.15 M NaCl. c) TEM micrograph of SH10 aggregates in the presence of 0.15 M NaCl.

#### *The effect of sulfhydryl groups on the aggregation kinetics*

Figure 4 shows the aggregation kinetics for SX/SH8 and SX/SH10 ovalbumin in the presence of 0.25 M NaCl. Protein solutions were incubated at a concentration of 5 mg/mL at pH 7.0 and a temperature of 70°C for various time intervals from 0 to 380 minutes. After cooling, the samples were applied to non-reducing SDS-PAGE to determine the content of disulphide bonds formed upon heating. It was found that the fraction of monomeric proteins declined with increasing incubation time due to aggregation (Figure 4). Comparing the rates of the loss of the monomeric protein fraction for SX8 and SH8 fractions, showed a faster rate of aggregation of sulfhydryl versus acetylthiogroups for both the SH/SX10 fractions as for the SH/SX8 fractions. The other fractions tested also showed similar results for the rate of aggregation when comparing acetylthiogroups and sulfhydryl groups with the same degree of modification (results not shown). For example, 50% of the SH8 molecules aggregated after 60 minutes of incubation at elevated temperature and only after 100 minutes 50% of the SX8 molecules had aggregated.

To test if these results were affected by the difference in thermal transition temperatures, the aggregation kinetics of the SX8 and SH8 fractions were also compared at their individual thermal transition temperatures (at 66°C for SH8 ovalbumin and at 73°C for SX8 ovalbumin). It was found that under these conditions the rate of aggregation was actually slightly lower for the SH8 fraction compared to the SX8 fraction (results not shown).



**Figure 4** Aggregation kinetics of ovalbumin with different degrees of modification  
Disappearance of monomeric material as determined by SDS-PAGE with increasing incubation time at elevated temperatures of SH ovalbumin (open symbols), SX ovalbumin (closed symbols), SX/SH10 (circles), SX/SH8 (squares).

It was found that aggregates consisting of acetylthiogroups were well soluble in SDS only (in the absence of  $\beta$ -mercaptoethanol). This indicates that the aggregates were primarily constructed of non-covalent interactions. In contrast, aggregates containing reactive sulfhydryl groups were only partially dissolved in SDS. Complete dissolving required the presence of  $\beta$ -mercaptoethanol indicating a combined contribution of covalent and non-covalent interactions to the network (results not shown).

## Discussion

### *Approach*

This work aims to investigate the contribution of disulphide bonds to the aggregation process of ovalbumin. Aggregation is the result of attractive forces exceeding repulsive forces. Attraction is governed by a combination of non-covalent and covalent (disulphide bonds) interactions, and the repulsive forces are often electrostatic or steric in nature. The introduction of additional sulfhydryl groups at the surface of ovalbumin was therefore hypothesised to enhance the opportunity for covalent interactions thus affecting aggregation kinetics. To this end, ovalbumin fractions were produced containing a range from zero to ten additional sulfhydryl groups. The reactivity of the introduced sulfhydryl groups was tested (results not shown) and the modification did not significantly affect the conformational properties of the proteins or their ability to unfold at elevated temperature. The results for

ovalbumin with acetylthiogroups or sulfhydryl groups were compared as this, next to the reduction of electrostatic repulsion by the addition of salt, allowed us to draw conclusions on the contribution of sulfhydryl groups to disulphide bonds in the aggregation process while eliminating other side-effects related to the modification procedure. It was found that homogeneous ovalbumin fractions could be produced with a range of sulfhydryl groups attached to the surface of the protein without affecting the secondary and tertiary structure (Table I, Figure 1).

#### *The impact of sulfhydryl groups on aggregation*

The attachment of reactive sulfhydryl groups to the surface of ovalbumin molecules did not result in spontaneous formation of covalently linked aggregates at room temperature, even when increasing the concentration to 200 mg/mL (Wierenga et al 2005) even though the sulfhydryl groups were already exposed and available for cross-linking. This implies that other attractive forces are required to enable sulfhydryl groups to remain in close proximity of each other for a certain time scale sufficient for disulphide interactions to occur, and also that disulphide bond formation may not be a driving force for aggregation.

At elevated temperatures, reactive sulfhydryl ovalbumin showed a comparable rate of aggregation as the acetylthiogroup fractions suggesting that the rate of aggregation was not significantly affected by sulfhydryl groups (Figure 3). Apparently, sulfhydryl groups do not affect the rate of aggregation for ovalbumin and therefore it is suggested that disulphide linkage is not the driving force for aggregation. The question remains how generic this observation is for other proteins. Arntfield et al (1991) and Alting et al (2004) also showed that non-covalent aggregation precedes the disulphide cross-link reaction for ovalbumin and whey protein isolate as well as for vicilin.

#### *Conclusions*

It was concluded that disulphide bond formation was preceded by physical interactions for ovalbumin aggregation and, as such, that disulphide interactions are not the driving force for aggregation of ovalbumin. It was also found that the morphology of the formed aggregates is shifted from a fibrillar and linear aggregate system toward a highly branched and random aggregate system upon increasing the number of reactive sulfhydryl groups per protein. Even though the rate of aggregation is not affected by sulfhydryl groups, the morphology of the formed aggregates is clearly affected by their presence.



## Materials and methods

### *Protein purification and modification*

*Isolation and purification* - Ovalbumin was purified from fresh hen eggs using a procedure based on previously described purification protocols (Kosters et al 2003). The protein solutions were dialysed, freeze-dried and stored at -20°C until use. The purity was over 98% as determined using densitometric analysis of SDS-PAGE traces.

*Modification* – Acetylthiogroups were introduced on the surface of the protein by the reaction between primary amino groups of lysine and S-acetylmercaptosuccinic anhydride (S-AMSA, Sigma) according to a previously described procedure (Klotz & Heiney 1962). To vary the degree of modification, S-AMSA was added to the protein in ratios varying from 0.1 : 1 to 1.2 : 1 (S-AMSA : lysines). The acetylthiolation was performed at 25°C and pH 8.0 for 24 h and a protein concentration of 25 mg/mL. After the incubation the solutions were dialysed extensively against demineralised water at 4°C and subsequently freeze-dried and stored at -20°C until use.

*Fractionation* – To obtain series of more defined ovalbumin variants, ion-exchange chromatography (Biopilot, Amersham) was used with a Source Q 900 mL column. 200 mL of a 5 mg/mL protein solution in 10 mM sodium phosphate buffer was applied to the column at a flow rate of 15 mL/min and eluted using a gradient from 0.08 M NaCl pH 6.0 to 1.00 M NaCl pH 6.0.

*Deblocking* – The fractions obtained from ion-exchange chromatography were divided into two batches. The acetyl group was cleaved off from the modified proteins in one of the batches by the addition of 1.0 M hydroxylamine hydrochloride at pH 7.0 and 25°C. The mixture was incubated for 1 h and then dialysed extensively against demineralised water at 4°C. This batch now contains proteins with a deliberated sulfhydryl group. The other batch was left untreated and contains an acetylthiogroup.

### *Protein characterisation*

*Chromogenic OPA Assay* – The degree of modification of the primary amino groups was determined indirectly by a chromogenic assay described by Church et al (1983) based on the specific reaction between ortho-phthaldialdehyde (OPA) and free primary amino groups in proteins. All measurements were performed in duplicate and corrected for a protein-free sample. The protein concentration was determined by the adsorption at 280 nm using an extinction coefficient of 0.712 ml\*mg<sup>-1</sup>\*cm<sup>-1</sup>.

*Determination of free sulfhydryl groups* – Ellman's reagent was used to determine the number of free thiol groups. The reagent, 5,5'-dithiobis(2-nitrobenzoic acid), DTNB, reacts with free thiol groups on the protein (Ellman 1958). Cysteine was used as a calibration standard. 4.8 mg DTNB was dissolved in 1.0 mL Tris-HCl buffer (pH 8.0 and 2% SDS), 2.5 mL Tris-HCl buffer. 50  $\mu$ L DTNB solution was added to 250  $\mu$ L 5 mg/mL protein solution. The solution was mixed and incubated for 20 minutes at 25°C. Then the absorbance was measured at 412 nm. The calibration curve, obtained by using various dilutions of a 0.50 mM cysteine solution, provided an extinction coefficient of sulfhydryl groups of 13425 M<sup>-1</sup> cm<sup>-1</sup>. Previously, it has been verified using the Sulfhydryl-disulphide Exchange Index (SEI index) described by Owusu-Apenten et al (2003) that the sulfhydryl groups introduced were chemically reactive (unpublished data).

*Circular dichroism spectroscopy* – Far-UV circular dichroism (CD) spectra of 0.1 mg/mL protein solutions in 10 mM phosphate buffer pH 7.0 were recorded at 25°C in the spectral range from 190 to 260 nm with a Jasco J-715 spectropolarimeter (Jasco Corporation Japan) using a quartz cuvette with an optical path of 0.1 cm. The spectral resolution was 0.5 nm, the scan speed was 100 nm/min and the response time was 0.125 sec with a bandwidth of 1 nm. 16 scans were accumulated and averaged. The spectra were corrected for the corresponding protein-free sample. Near-UV CD spectra were obtained of 1.0 mg/mL protein solutions in 10 mM phosphate buffer (pH 7.0) at 25°C in the spectral range of 250-350 nm in quartz cuvettes with an optical path of 1 cm. A spectral resolution of 0.5 nm was used and a scan speed of 100 nm/min, a response time of 0.125 sec and a bandwidth of 1 nm. 16 scans were accumulated and averaged and the spectra were corrected for a protein-free sample.

*Fluorescence spectroscopy* – Emission spectra of 5  $\mu$ g/mL protein in a 10 mM phosphate buffer (pH 7.0) were analysed using a Perkin Elmer Luminescence Spectrometer (LS50B) at 20°C. Quartz cuvettes with an optical path of 1 cm were used. The excitation and emission slit widths were 5 nm and a scan speed of 120 nm/min was used. Spectra were obtained from 300 – 400 nm at excitation wavelengths of 295 and 274 nm. All spectra were performed in duplicate and corrected for the corresponding protein-free sample.

*Isoelectric focusing (IEF)* – The apparent isoelectric points of the ovalbumin variants were determined using the Phast System (Pharmacia). 4  $\mu$ L of 1.0 mg/mL protein solutions were applied to IEF gels with a pH gradient ranging from 2.5 to 6.5 (Pharmacia) and from 3 to 10 (Pharmacia). The gels were fixed with 20% trichloric acid, stained using Coomassie Brilliant blue (R-250) and destained in 30% methanol/10% acetic acid.

*Differential Scanning Calorimetry (DSC)* – 4.0 mg/mL protein solutions in 10 mM phosphate buffer pH 7.0 were sealed in the cell of a VP-DSC MicroCalorimeter (MicroCal Inc., Northampton USA). A 10 mM phosphate buffer was used as reference sample. The heat flow was recorded in separate duplicates from 25 to 110 °C at a heating rate of 1°C/min. The data were analysed using the MicroCal Origin software.

#### *Aggregation kinetics*

The aggregation kinetics of the ovalbumin variants were tested, either by incubation at 2°C below the gel temperature (defined as the temperature at which G' significantly deviates from G'' in rheology measurements), or by incubation at a fixed temperature of 70°C. Unmodified, SX8 (ovalbumin with eight acetylthiogroups) and SH8 (ovalbumin with eight additional sulfhydryl groups) ovalbumin were incubated at 72°C, 76°C and 64°C, respectively, at a concentration of 3 mg/mL in 50 mM phosphate buffer at pH 7.0. 50 µL aliquots were withdrawn at various time intervals from 0 to 6 h. Alternatively, all ovalbumin variants were incubated at 70°C at a concentration of 5 mg/mL in 50 mM phosphate buffer pH 7.0 and 0.25 M NaCl and aliquots were withdrawn at a series of time intervals up to 400 min. After sampling, the solutions were diluted directly into equal volumes of SDS PAGE sample buffer containing 2% SDS and 30 mM NEM and cooled on ice. After cooling, the samples were applied to an SDS-PAGE gel as described below.

*SDS-PAGE* – Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to a protocol described by Laemmli (1970). A 15% (w/v) acrylamide separating gel and a 4% (w/v) acrylamide stacking gel containing 0.1% SDS (w/v) were run using a Mini-PROTEAN II Electrophoresis Cell (Biorad). Samples of 0.1 mM protein were prepared in sample buffer containing 10% SDS (w/v) and 1.25% (v/v) β-mercaptoethanol. Gels were stained with Coomassie Brilliant Blue R-250 and destained with 30% methanol/10% (v/v) acetic acid in water. The fractions of monomeric and polymeric material were estimated using densitometric analysis. The accuracy of this analysis was estimated 5% derived from triplicate measurements and subsequent densitometric analysis.

#### *Aggregate morphology*

*Transmission electron microscopy* – 20 mg/mL protein solutions in 10 mM phosphate buffer pH 7.0 in the absence and presence of 0.15 M NaCl were prepared, heated for 60 minutes at 90°C and subsequently cooled to 20°C. The solutions were diluted to 5 mg/mL and

placed on a perforated carbon film, supported by a 200 mesh copper grid and subsequently blotted. The liquid film was vitrified by rapidly plunging the grid into liquid propane at a temperature of -170°C. The specimens were stored in liquid nitrogen until transfer into the cryo-holder. The images were recorded digitally using a Philips CM12 transmission electron microscope operating at 80 kV with a Gatan 791 CCD camera using Digital Micrograph software.

### Acknowledgements

The authors acknowledge Hans Kosters for assistance with the modification procedure and Ellmann's assay and Johan Hazekamp for support with cryo-TEM analysis. The authors also acknowledge Ton van Vliet for useful discussions.

### References

- Alting AC, Hamer RJ, de Kruif CG, Visschers RW (2000) *J Agric Food Chem* 48: 5001-5007
- Alting AC, Weijers M, de Hoog EH, van de Pijpekamp AM, Cohen Stuart MA, Hamer RJ, de Kruif CG, Visschers RW (2004) *J Agric Food Chem* 52: 623-631
- Arntfield SD, Murray ED, Ismond MAH (1991) *J Agric Food Chem* 39: 1378-1385
- Broersen K, Weijers M, de Groot J, Hamer R, de Jongh H (2005) Electrostatics controls fibril formation – Part I: Charge engineering of proteins affects unfolding. Manuscript in preparation.
- Church FC, Swaisgood HE, Porter DH, Catignani GL (1983) *J Dairy Sci* 66: 1219-1227
- Ellman GL (1958) *Arch Biochem Biophys* 82: 70-77
- Fava A, Iliceto A, Camera E (1957) *J Am Chem Soc* 79: 833-838
- Hoffman MAM, van Mil PJJM (1997) *J Agric Food Chem* 45: 2942-2948
- Iametti S, Donnizelli E, Vecchio G, Rovere PP, Gola S, Bonomi F (1998) *J Agric Food Chem* 46: 3521-3527
- de Jongh HHJ, Goormaghtigh E, Killian JA (1994) *Biochemistry* 33: 14521-14528
- Kato A, Nagase Y, Matsudomi N, Kobayashi K (1983) *Agric Biol Chem* 47: 1829-1834
- Kato A, Ibrahim H, Takagi T, Kobayashi K (1990) *J Agric Food Chem* 38: 1868-1872
- Kim SC, Olson NF, Richardson T (1990). *Milchwissenschaft* 45: 627-631
- Kim SC, Olson NF, Richardson T (1990b) *Milchwissenschaft* 45: 580-583
- Klotz IM, Heiney RE (1962) *Arch Biochem Biophys* 96: 605-612
- Koseki T, Kitabatake N, Doi E (1989) *Food Hydrocolloids* 3: 123-134
- Kosters HA, Broersen K, de Groot J, Simons JFA, Wierenga P, de Jongh HHJ (2003) *Biotech Bioeng* 84: 61-70
- Laemmli UK (1970) *Nature* 227: 680-685
- McSwiney M, Singh H, Campanella OH (1994) *Food Hydrocolloids* 8: 441-453
- Owusu Apenten RK (1998) *Int J Biol Macromolecules* 23: 9-25
- Owusu-Apenten RK, Chee C, Hwee OP (2003) *Food Chem* 83: 541-545
- Phillips NI, Jenness R (1967) *Arch Biochem Biophys* 120: 192-197
- Sawyer WH (1967) *J Dairy Sci* 51: 323-329

- Shimada K, Cheftel JC (1989) *J Agric Food Chem* 37: 161-168
- Strange ED, Holsinger VH, Kleyn DH (1993) *J Agric Food Chem* 41: 30-36
- Sun Y, Hayakawa S (2002) *J Agric Food Chem* 50: 1636-1642
- Tanaka N, Tsurui Y, Kobayashi I, Kunugi S (1996) *Int J Biol Macromolecules* 19: 63-68
- Verheul M, Roefs SPFM, de Kruif CG (1998) *J Agric Food Chem* 46: 896-903
- Weijers M, Broersen K, Barneveld P, Cohen Stuart M, Hamer R, de Jongh H, Visschers R (2005) Electrostatics controls fibril formation. Part II – Elucidation of the molecular mechanism of aggregation. Manuscript in preparation.
- Wierenga PA, Kusters H, Egmond MR, Voragen AGJ, de Jongh HHJ (2005) Disulfide bridges do not affect rheological properties of interfacial protein films. Unpublished work.
- Wu J, Bratko D, Prausnitz JM (1998) *Proc Natl Acad Sci USA* 95: 15169-15172
- Xiong YL, Dawson KA, Wan L (1993) *J Dairy Sci* 76: 70-77



# CHAPTER 9

## DISCUSSION

### DRIVING FORCES FOR PROTEIN AGGREGATION IN PERSPECTIVE – THE AGGREGATION PROPENSITY (AGPRO) MODEL

Broersen K, de Jongh HHJ, Voragen AGJ, Hamer RJ

Manuscript in preparation

#### **Introduction**

Despite numerous efforts undertaken to effectively solve the ‘protein-folding problem’, still a number of unresolved issues surround the phenomenon of irreversible folding of polypeptides into a kinetically trapped self-associated state. The extensive interest in this topic can be rationalised by the finding that this kinetically trapped state is generally considered a dysfunctional state of proteins, which, on top of that, has a proven track record of associated, debilitating human diseases. Examples of these diseases have been previously reported by Bucciantini et al (2002). The list includes Alzheimer’s, Huntington’s and Parkinson’s disease. At the same time, the food industry has adopted the topic of protein aggregation from a very different perspective, as the property of aggregation has been related to the formation of texture in food products (e.g. Doi 1993). Protein aggregation can result in changes in viscosity of food products (such as dairy products) or gel formation, the latter of which is a process generally regarded as a follow-up process of aggregation. Both fields of interest can benefit from a fundamental knowledge of molecular forces that potentially drive the aggregation process. A better understanding of the processes driving the induction of aggregation will provide potential means to develop treatment or preventive measures for the treatment of disease. In the food sciences, this understanding will improve ways to adapt processing conditions of food products or to exchange the functionality of food proteins as functional ingredients.

*The processes of folding/misfolding and aggregation*

The processes of protein unfolding and refolding have been extensively studied. The pioneering work by Anfinsen et al (1973) showed that the unfolding step of ribonuclease is actually a reversible step and the primary sequence contains all necessary information to refold the unstructured polypeptide chain on a relatively short timescale into a fully functional protein. Later, it was recognised that unfolded proteins can undergo an irreversible self-assembly reaction, also called aggregation (see f.e. Zettlmeissl et al 1979). The aggregates are trapped in a minimum of conformational energy, which is protected by a high activation energy of dissociation (Rudolph et al 1979). Once the protein conformation has reached a certain state of unfolding, allowing the exposure of hydrophobic or disulphide/sulphydryl groups, aggregation becomes inevitable and is merely a matter of time and concentration. The protein-folding scheme fitting this model has first been defined by Ferry in 1948 and was later modified by Lumry and Eyring (1952) and describes the ‘point-of-no-return’ as follows by the induction of a conformational change of the secondary structure:

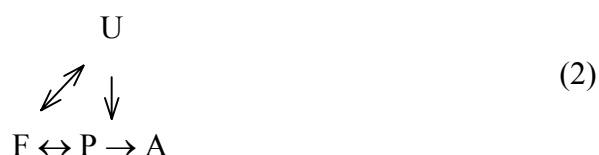


where F,  $U_{3^{\circ}}$ ,  $U_{2^{\circ}}$  and A represent the folded state, the unfolded tertiary structure, the unfolded secondary structure and an aggregated state, respectively. According to this scheme, only unfolding of the secondary structure of a protein leads to irreversible folding while destruction of the tertiary structure allows a reversible transition to the folded state.

Soon after, Zettlmeissl et al (1979) recognised that the formation of a correctly folded conformer competes with the formation of misfolded states or aggregates during the process of protein folding and that the refolding yield is not necessarily related to the unfolding of the secondary structure. This finding underlines the fact that the aggregation process does not necessarily take place as a consequence of the complete unfolding of the polypeptide chain. Instead, self-association probably requires similar stabilising forces as the folding process. This also raises the possibility that protein molecules perhaps do not have to be completely unfolded in order to aggregate. This has also been observed within our research group for  $\beta$ -lactoglobulin. Partial unfolding is often sufficient to initiate the aggregation process (Broersen et al 2005d). Also Zettlmeissl et al (1979) have shown that lactic dehydrogenase in the aggregated state resembles the unfolded enzyme with respect to their fluorescence properties



while far-UV circular dichroism suggests that the secondary structure of the enzyme in the aggregated state is partially restored. The aggregation process was therefore restated in the following manner as first proposed by Minton et al (1982) and as modified by our research group.



where F, P, U and A represent the folded state, a (partially) unfolded state prone to aggregation, the completely unfolded state and an aggregated state, respectively (Minton et al 1982). The states in this scheme are distinguished by the energy barriers or kinetics of the various processes involved. The scheme illustrates that refolding and aggregation processes can compete with each other. The ratio of refolded vs aggregated proteins depends on a number of factors (Jaenicke & Rudolph 1977, Teipel & Koshland 1971). The aggregation reaction obeys more or less first-order kinetics (various values for this reaction have been reported, see f.e. Koseki et al 1989a, 1989b). Protein concentration thus plays a crucial role in protein aggregation. To illustrate this, Ellis and Minton (2003) reported that the concentration of proteins and other macromolecules inside cells can be up to 400 g/L. At such high protein concentrations, self-association seems inevitable. For example, the occurrence of cataract in the eye lens has been related to insolubility of  $\gamma$ -crystallin (Delaye & Tardieu 1983). Still, more frequently, the body manages to maintain such high protein concentrations in solution without significant self-association over many years. Apparently, the body must have built in some type of preventive mechanism to escape from this disease-related pathway. Indeed it has been recognised that the endoplasmic reticulum houses a so-called quality control system that encompasses molecular chaperones and glycan modification enzymes. This system provides a measure for misfolded proteins to be degraded or to acquire a correct and functional conformation after all (Chevet et al 2001). Chevet reviewed the routes by which the endoplasmic reticulum regulates the release of correctly folded proteins.

### *Defining states*

As is illustrated in the previous paragraph, it cannot unambiguously be defined which conformational state is associated with a high aggregation propensity. This largely depends on

the way the various conformations a protein can assume, are defined. In this paragraph, we will discuss the available definitions and identify possible routes to identify these states and, using data from literature, relate these to aggregation propensity.

Native (and denatured) states – The biologically active state of a protein has frequently been referred to as the native state (e.g. Johansson et al 2005, Sahin et al 2004). Biologically active proteins are often well-structured at a secondary, tertiary and quaternary level and are usually compact, but there is still mobility allowed around many bonds in the molecules. This flexibility then allows for the biological activity to take place. For example, phenol hydroxylase enables substrate binding through a conformational change of its flavine cofactor (Enroth et al 1998). Another more distinct example is cytochrome c, which partially unfolded form has been reported to have considerably more peroxidase activity as compared to the fully folded form (Diederix et al 2002). The conformational changes of these types of enzymes functioning through a so-called ‘induced-fit-mechanism’ have been reviewed by Koshland (2003). Another example is provided by plasminogen that can be activated by a large conformational change converting the inactive ‘closed’ conformation into an ‘open’ conformation with plasmin activity (Wang et al 2000). Defining the native state by biologic activity thus does not account for the, sometimes significant, conformational changes required for full functionality. The loss of the native state is therefore usually defined as the loss of enough structure to render the enzyme inactive and is synonymous for denaturation. Considering the examples provided in this paragraph, a loss of activity sometimes occurs only after significant changes in structure are observed. This definition has some serious problems: for example, in the case where the biological activity is unknown, or in the case when the biological activity depends on solvent conditions. The latter aspect has been reviewed for enzyme function by Gupta (1992). For example,  $\beta$ -lactoglobulin belongs to the lipocalin family of proteins and can bind small hydrophobic molecules within its central cavity, such as retinol (Papiz et al 1986) but up to today the biological activity of this protein has not been confirmed (Collini et al 2000). The term ‘native’ is in these cases often referred to as the folded state, which, however, as has been discussed before with phenol hydroxylase, is not always synonymous of the native state. From these examples it appears that it is difficult to unambiguously define the ‘native’ and ‘denatured’ states and the use of these terms should therefore be carefully considered within the context they are used.

Folded and unfolded state – Folded proteins have been defined in the past as protein molecules with intact structural properties (e.g. Hamada et al 2005). Folded proteins have a compact conformation and exhibit secondary, tertiary and often quaternary structure. Whether a protein is structurally intact is an ambiguous discussion as for some systems information on the *in vivo* folded state cannot be obtained. We consider it therefore more reasonable to assume that the folded state of a protein is a state of which it can be expected that an energy minimum has been achieved by some globular fold under physiological conditions (in the absence of heat, denaturant, high pressure etc.). This definition will immediately raise the concern that a number of proteins are known to be partly unfolded under physiological conditions. It is clearly difficult to propose an all-inclusive definition for this state. We therefore recommend that publications considering protein conformation should refer to the criteria used to define these states. A completely unfolded protein can be defined as an extended chain in which the peptide bonds and disulphide bonds, if present, are the only intramolecular bonds that restrain the spatial arrangement of the polypeptide. In between the two extremes of being totally folded or totally unfolded, many intermediately folded states have been found to populate under various conditions. For example, the ‘molten globule’ which has been first defined for  $\alpha$ -lactalbumin by Kuwajima et al in 1975. These partially folded states are characterised by their loss of compactness while some interactions still exist.

Aggregated state – Aggregation has been defined by Walstra (2003) as ‘a state in which the particles stay close together for a much longer time than they would in the absence of attractive colloidal forces’. This definition also includes proteins with a folded and functional multimeric quaternary structure such as tubulin and actin which form filaments that play a role in the cytoskeleton of cells, cell division and muscle function (Guan et al 2005, Nogales et al 1998, Roychowdhury et al 2000). Next to this, dysfunctional pathological aggregates can be distinguished such as amyloid fibrils or the pathological aggregation of hemoglobin in sickle cell anaemia (Poillon et al 1998). We therefore propose to distinguish between proteins with a multimeric quaternary structure required to perform a specific function as ‘native’ aggregates such as bovine  $\beta$ -lactoglobulin or fibrin and dysfunctional aggregates formed upon (partial) unfolding of the non-functional polypeptide chains resulting in interactions as ‘non-native’ aggregates.

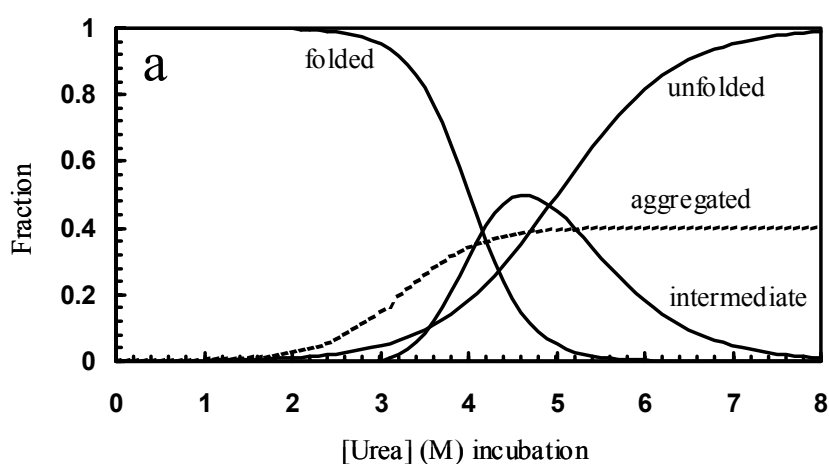
### *Methodology*

The use of the definitions will improve the discussion on protein refolding and aggregation since that what constitutes the defined states is largely dependent upon the method applied to observe the protein molecule. Some methods can detect for example the smallest changes in structure or function while others require rather large conformational alterations or function before changes are observed. For example, one of the oldest methods used to follow the course of unfolding is to measure changes in solubility (Schaefer et al 1960). However, the loss of solubility often only reflects the very late stages in a series of changes in structure. This method is thus a rather crude measure of protein unfolding. Recently, a number of reviews on the available methodology to determine the states of proteins have been published (Brockwell et al 2000, Clarke & Waltho 1997, Eftink 1994, Ladokhin 2000, Price 2000). In the next paragraph we will discuss the most commonly used methods and their role in detecting aggregation-prone conformations.

Fluorescence techniques – Fluorescence relies on the emission of radiation from the excited electronic state of tryptophan or tyrosine side residues that have been populated by absorption of light. Both the intensity and wavelength of maximum emission are sensitive measures of the polarity of the local tryptophan environment, and hence provide useful information regarding the tertiary structure (Campbell & Dwek 1984, Schmid 1997). Fluorescence has been used both in equilibrium unfolding studies (e.g. Navea et al 2002, Onda et al 1997) as in time-resolved mode (e.g. Shastry et al 1998) and has provided valuable information on the processes of folding, folding kinetics and folding pathways particularly for proteins exerting two-state unfolding behaviour (Luo & Baldwin 2001, Pandya et al 1999). Nevertheless, this method heavily relies on the presence of tryptophan residues in the primary sequence of the polypeptide studied. Also, the polarity of the tryptophan residues is only representative of the local state of conformation around these residues and does not necessarily reflect the global state of folding, especially in the case of large multi-domain proteins. Next to this, signals obtained from intrinsic tryptophan fluorescence could be complex by the presence of multiple tryptophans in the primary sequence of many proteins.

Circular Dichroism - The method of circular dichroism relies on the interaction of circularly polarised light with optically active compounds (Price 2000). The use of this technique to investigate protein structure has been reviewed by Kelly and Price (2000). The various regions in the spectrum are sensitive to either secondary structure elements or

aromatic side chains of the aromatic amino acid residues (Fasman 1996, Kelly & Price 1997, 2000). This method, like fluorescence, has been used extensively to investigate the kinetics (stopped-flow circular dichroism) as well as equilibrium status of conformation and is sensitive to heat-induced as well as denaturant-induced unfolding (e.g. Kosinski-Collins & King 2003, Navea et al 2002, Onda et al 1997, Raschke et al 1999). An important limitation of circular dichroism is that this method provides relative low-resolution structural information and is unable to pinpoint which particular regions of the sequence are of a particular secondary structure (Kelly & Price 2000). This is especially relevant to aggregation where detailed structural information is needed. This last limitation can be more or less overcome when the experimental results are used along with algorithms developed to predict secondary structure of particular regions of the polypeptide chain (Kelly & Price 2000, Ribas de Pouplana et al 1991). Comparable to intrinsic fluorescence techniques, far-UV circular dichroism has been tested in our research group for the ability to detect the minute alterations in the secondary structure of bovine  $\beta$ -lactoglobulin required for aggregation (Broersen et al 2005d). Figure 1 shows that the augmentation of the unfolded fraction detected by far-UV circular dichroism does not superimpose with the observation of denaturant-induced aggregate formation through gel permeation chromatography.



**Figure 1** Model for the equilibrium between the folded, intermediate, unfolded and aggregated fractions of bovine  $\beta$ -lactoglobulin at various urea-concentrations. The model is based on data obtained from intrinsic tryptophan fluorescence, gel permeation chromatography and far-UV circular dichroism and has been published before in Broersen et al (2005d). In this model it is assumed that aggregates constitute unfolded molecules.

NMR and X-ray crystallography – Both fluorescence and far-UV circular dichroism sometimes appear insufficiently sensitive to the minute conformational changes required to initiate the aggregation process. Such changes should be readily observed using methods sensitive to conformational changes at an atomic level. Both NMR and X-ray crystallography techniques obey this criterion. The majority of the structures deposited in the Protein Data Bank (<http://www.pdb.bnl.gov>) have been determined using X-ray crystallographic or NMR techniques illustrating their detailed structural, high-resolution information obtained. However, even though X-ray crystallography provides high-resolution information at an atomic level, it also provides a static image of the protein structure without giving insight into the dynamics. Moreover, X-ray crystallography is rather time-consuming as the preparation of crystals which diffract to high resolution can take a long time and sometimes even proves impossible (Wess 1997). NMR overcomes these problems as with this technique it is possible to obtain high-resolution information on the dynamics of proteins. This method however requires high concentrations of sample, has difficulties to deal with heterogeneity and is an expensive way to obtain high-throughput structural information of a large range of different protein molecules (James & Oppenheimer 1994).

Recommendations for studying conformational aspects related to aggregation - Bhattacharjee and Das (2000) found complete reversibility of the unfolding process of  $\beta$ -lactoglobulin using intrinsic tryptophan fluorescence while extrinsic ANS fluorescence showed a significant difference in the fluorescence spectra of the folded and refolded protein conformation. Also the finding that intermediate states can form upon protein folding having pronounced secondary structure but lacking specific tertiary structure as determined by intrinsic fluorescence (Kuwanjima 1989, 1996, Ptitsyn 1995) raises the possibility that perhaps a combination of techniques to study the folding process need to be explored. Arai et al (1998) for example used small angle X-ray scattering, circular dichroism, absorption spectroscopy and fluorescence spectroscopy to obtain comprehensive information of the kinetic refolding process of  $\beta$ -lactoglobulin, providing a more powerful approach to obtain a detailed description of the conformational state of protein. Also within our research group, it was observed that even though no unfolding is detected through intrinsic fluorescence, the protein apparently attained a slightly unfolded conformation able to promote the aggregation process (Broersen et al 2005d) (see Figure 1 in which it is shown that aggregation already occurs before changes in the fluorescence spectrum become apparent). These observations

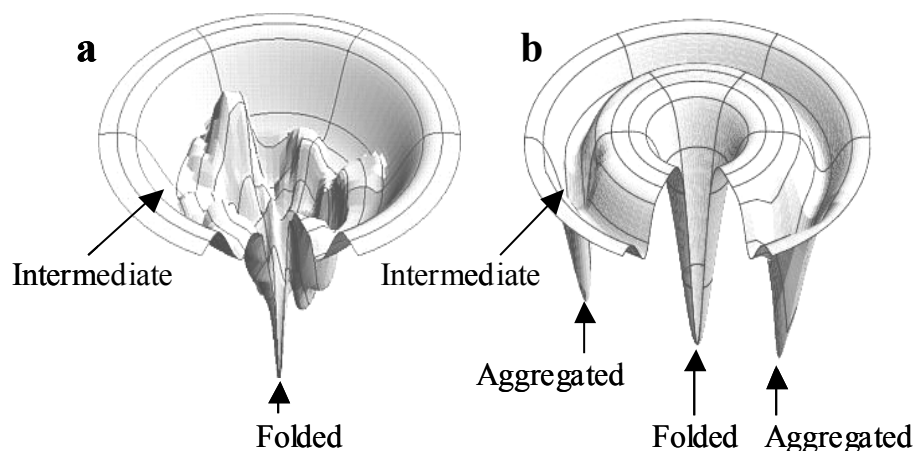
indicate in the first place that rather little unfolding would already be sufficient to induce aggregation of some proteins and secondly, that using a single method to detect the unfolding process, the amount of unfolding required for aggregation may not always be detected by a single technique.

### *The onset of aggregation*

The onset of protein aggregation originates from a kinetic (fast formation) or thermodynamic preference of the polypeptide chain to reside in an aggregated state rather than in an unfolded or folded conformation. Energy landscapes or folding funnels provide a (for folding funnels three-dimensional) representation into the energy household of a selected polypeptide chain during the transition from the unfolded to the folded state (Bryngelson et al 1995, Dill & Chan 1997). An example of a folding funnel is shown in Figure 2 and presents the various valleys that correspond to many intermediate folded states. Traversing down to the folded state, the entropy of the ensemble of conformations decreases due to the development of an increasing number of native contacts. This means that the folding landscape is funnel-like. As more and more local sections fall into place, the protein molecules achieve lower energy, with the final folded structure at the bottom of the funnel. Figure 2a represents the folding funnel of a protein that folds, via a number of intermediate states, into the correct (fully folded) conformation. That the folding funnel shown in Figure 2a does not comprise the folding process of all proteins has been illustrated by many reported experimental data in which an energetically stable aggregated state populates with increasing protein concentration (e.g. Chiti et al 2002, Hamada & Dobson 2002).

What determines whether a protein will fold correctly or misfold into an aggregated state? Figure 2b presumes that aggregate formation depends on a significant population of partly (un)folded proteins and hence, thermodynamically, the intermediate folded state needs to be marginally stable. Also, the activation barrier to transfer from the intermediate state to the aggregated state should not exceed the activation energy for correct folding to a large extent. A low activation energy for aggregation can be achieved for example when the intermediate state exposes sufficient hydrophobicity and low electrostatic repulsion. In other words, the kinetic energy barrier is in that case negligible, while the gain in thermodynamic energy upon aggregation is significant. Moreover, the activation energy required for correct refolding is most likely dictated by intramolecular reorganisation of the polypeptide in order to proceed into a properly folded state. Summarising, there are many routes by which proteins can become potentially aggregating species. As Chiti et al (2002) stated, the possibility of

forming aggregates is intrinsic to any protein and these intrinsic factors will essentially describe the thermodynamic and kinetic barriers in the energy landscape to the correct and functional conformation.



**Figure 2** *Folding funnels of protein folding (reproduced from Chan & Dill 1998, Dill & Chan 1997) representing a) the folding funnel of proteins that fold without population of aggregation prone intermediates and b) the folding funnel of proteins that fold with the population of aggregation prone intermediates resulting in a second energetic minimum with a high activation energy of dissociation. The width of the funnel represents the entropy of the system and the length scale of the funnel represents the energy.*

#### *Strategy for studying parameters determining the aggregation propensity*

In the previous paragraphs it has been consolidated that the aggregation propensity of proteins has been difficult to predict due to a number of issues. First, the definitions handled in this area are frequently applied in a proliferated way. Second, the detection of aggregation prone conformers has been proven difficult in the past. Next to this, a clear difficulty exists in determining which (combination of) factors contribute to the aggregation propensity of a protein. In the following paragraphs, the molecular details of proteins, which determine the probability of proteins to aggregate, will be discussed. The strategy we employed to evaluate the relative contribution of a number of selected parameters to the aggregation propensity includes the critical evaluation of relevant existing literature with results obtained in our laboratory and the merging of these results into a model obtained from multiple linear regression analysis, the AGgregation PROensity (AGPRO) model.



### *Factors related to the kinetics of aggregation*

The observation that proper protein folding can be complicated by misfolding and aggregation, is a well-established fact today (see f.e. Dobson 2004). The origins behind the varying irreversibility upon refolding for different proteins are diverse. Stability (Chiti et al 2000, Hurle et al 1994, Ramirez-Albarado et al 2000, Siepen & Westhead 2002, Quintas et al 1999) quite clearly has been shown to determine the exposed hydrophobicity (Chiti et al 2002, Otzen et al 2000, Schwartz et al 2001) and the feasibility of thiol-disulphide exchange (Yagi et al 2003). Also electrostatic repulsion has been reported to contribute to the selection of states (Chiti et al 2002b, de la Paz et al 2002, Konno 2001, Tjernberg et al 2002). The contribution of these factors to aggregation has been firmly established in the past by various authors. In the next paragraphs we will discuss the various factors identified in the past and by us using the approach of chemical engineering in more detail, relate their significance to the reported results of the very early stages of the aggregation process and provide a general mechanism to predict the aggregation propensity (AGPRO).

### *The effect of kinetic and thermodynamic stability on the aggregation propensity*

Partial unfolding of proteins may be required to form aggregates (Calamai et al 2003, Dobson 1999, Kelly 1996, Rochet & Lansbury 2000). If mutations in the protein sequence lead to a destabilisation of the native state this will increase the population of partially folded molecules and enhance the possibilities for aggregation. This theory was experimentally supported by Hamada and Dobson (2002) who showed that efficient fibril formation of  $\beta$ -lactoglobulin requires a balance between destabilising actions aimed to populate unfolded or partially unfolded proteins and attractive forces driving aggregation. Furthermore, Orsini and Goldberg (1978) and Zettlmeissl et al (1979) proposed that a kinetic competition between the folding route and the aggregation route may be responsible for the observation that the yield of renatured protein is dependent on protein concentration. Kinetic and thermodynamic stability of the folded conformation have also been inversely associated with the deposition of proteins in the form of aggregates by other authors (Chiti et al 2000, Hurle et al 1994, Kelly 1998, Ramirez-Alvarado et al 2000, Quintas et al 1999, Siepen & Westhead 2002). However, it has been shown that in a number of cases the destabilisation of a protein may not always result in enhanced aggregation. For instance, amyloidogenic mutations of prion proteins do not affect the instability of the folded state in a significant way compared to the wild type prion proteins (Liemann & Glockshuber 1999). Also, other authors have recently found

indications that protein instability is not fully correlated with the aggregation propensity of proteins in general (DuBay et al 2004, Hammarstrom et al 2003). Our results showed that, whereas glucosylation leads to a significant destabilisation at room temperature, the extent of aggregation is actually reduced (Broersen et al 2005d). Thus, it seems that, even though (partial) protein unfolding is an absolute requirement to expose reactive groups for aggregate formation, destabilisation of the folded state does not necessarily predict the aggregation propensity of proteins in a straightforward fashion.

*The effect of thiol-disulphide exchange on the aggregation propensity*

Thiol-disulphide exchange was first observed by Lechner already in 1920 and later the kinetics of this reaction was investigated by Fava et al in 1957. The relevance of this reaction for intermolecular protein-protein interactions has since been shown for example by Yagi et al (2003) for heat-induced unfolding of  $\beta$ -lactoglobulin, Kraut and Sagone (1984) for granulocyte aggregation and Essex et al (2001) for platelet activation, processes which strongly rely on polymerisation rather than general aggregation processes. Yagi et al (2003) observed irreversible refolding of bovine  $\beta$ -lactoglobulin upon cooling from a heat- or denaturant-induced unfolded state and found that the irreversibility originated from a thiol-disulphide exchange reaction between a free thiol and a disulphide bond. They provided further evidence for their case by constructing and expressing  $\beta$ -lactoglobulin mutants in which the free thiol group was replaced by other residues. The mutants showed complete reversibility from the unfolded state. From this finding, and other reports (Essex et al 2001, Kraut & Sagone 1984) it has been postulated that the presence of sulfhydryl groups on proteins are important determinants for the formation of stable aggregates but it was not clear whether the onset of the reaction (or rate of aggregation) was also affected in this way. Common approaches to investigate this include chemical blockage of sulfhydryl groups present on the cysteine residue of a range of proteins using (*N*-ethylmaleimide) NEM or competitors in the sulfhydryl/disulphide exchange process such as DTT (Hoffman & van Mil 1997, Owusu Apenten 1998, Sawyer 1967, Tanaka et al 1996). Also activation of the sulfhydryl groups by unfolding has been used in the past to investigate the role of disulphide bonds in aggregation (Phillips & Jenness 1967, Shimada & Cheftel 1989, Xiong et al 1993). This last approach possibly leads to enhanced aggregation due to accumulation of unfolded proteins resulting in ambiguous interpretation of the obtained results. Furthermore, the use of NEM and DTT has been reported to affect the aggregation properties other than only by

means of the interruption of disulphide bond formation (Alting et al 2000). Another approach to investigate the relevance of sulfhydryl groups to the aggregation process involves chemical modification by using the reaction between primary amino groups of lysine and S-acetylmercaptosuccinic anhydride according to a procedure that was previously described by Klotz and Heiney (1962). In our research group, we found that the rate of aggregation of ovalbumin was not affected by this reaction, whereas the morphology and stabilisation of the formed aggregates can actually be ascribed to thiol-disulphide exchange (Broersen et al 2005b). This finding is consistent with the reports by Alting et al (2004) and Arntfield et al (1991) who found that disulphide linkage is not a driving force for aggregation but rather a stabilising interaction once the aggregates are formed. The fact that the length of a disulphide bond is limited to 2.05 Å in addition to the time required to enable the thiol-disulphide exchange reaction to take place (Fava et al 1957) provides a reasonable explanation why this presumably stabilising reaction is preceded by fast, non-specific and longer-range non-covalent interactions. It must be noted that even though the protein aggregation process is usually determined by non-covalent interactions, for polymerisation reactions disulphide interactions often provide the only type of link to induce intermolecular association and as such provide the driving force for the association. In this chapter we will not further explore polymerisation processes.

#### *The effect of glycosylation on the aggregation propensity*

Unwanted self-association of proteins in the cell upon synthesis can be prevented by the interactions of molecular chaperones with glycans attached to the protein (Helenius et al 1997, Land & Braakman 2001). Many other examples in literature confirm that the aggregation of proteins is reduced by the presence of covalently linked sugar moieties (Marquardt & Helenius 1992, Song et al 2001). The aggregation inhibiting effect of glycosylation has been related to the kinetic partitioning between folding and aggregation. Numerous experimental data confirmed that the folding rate of proteins is enhanced with the covalent attachment of glycan chains (see f.e. Fenouillet & Jones 1995). This effect of protein glycosylation has been firmly associated to the chaperone system present in the endoplasmic reticulum, which functions through specific cleavage of parts of the glycan chains attached to the nascent polypeptide chain (Helenius et al 1997). Later it was recognised that, also *in vitro*, in the absence of the chaperone system, the folding rate of glycosylated proteins is faster compared to their non-glycosylated counterparts (Broersen et al 2005d, Wang et al 1996). Apparently, the chaperone system is not the only mechanism by which sugar chains can

enhance protein folding. In our laboratory we investigated the possible molecular mechanism behind this observation (Broersen et al 2004, 2005a, 2005d). It was first suggested that the deceleration of the aggregation process upon glycosylation was related to destabilisation of the folded conformation (Hurle et al 1994). We indeed found a reduced stability ( $\Delta G$ ) at room temperature (Broersen et al 2005d) but we also found an increased thermal transition temperature (Broersen et al 2004). Further studies (van Teeffelen et al 2005) revealed that the heat capacity change upon unfolding was significantly affected by glycosylation. If glycosylation indirectly controls the aggregation propensity of proteins through stability, it can be expected that the temperature at which the aggregation rate is observed is affected accordingly. Surprisingly, at a low temperature (by urea-induced denaturation) the rate of aggregation was lower (Broersen et al 2005a, Wang et al 1996) compared to non-glycosylated protein, whereas at a high temperature the aggregation rate of the glycosylated protein even exceeded that of the non-glycosylated protein (Broersen et al 2005a). This apparent controversy can be explained as follows: Upon urea-induced aggregation the aggregation was brought about by dilution of the solution with buffer resulting in a kinetic partitioning between refolding and aggregation (Chiti et al 2002). Upon heat-induced aggregation, the refolding step is practically eliminated and the unfolded molecules are directed towards the aggregation pathway (Najbar et al 2003, Tani et al 2004). This indicated that not the rate of aggregation, but the rate of refolding was affected by the glycosylation, after an electrostatic or hydrophobic contribution to the observed effects was ruled out. Using time-resolved fluorescence techniques, we found that the refolding of glycosylated protein was significantly faster, and mostly occurred during the dead-time of the experiment, compared to non-glycosylated protein (Broersen et al 2005d). This is consistent with many other reports on the effects of glycosylation on the rate of folding (Fenouillet & Jones 1995). Glycosylation thus results in fast refolding kinetics leading to a preference of the refolded state over the aggregated state. It can therefore be suggested that glycosylation indirectly paved the way for fast refolding and that also other factors, which affect the refolding kinetics of  $\beta$ -lactoglobulin, supposedly modify the kinetic partitioning between aggregation and refolding. The reasons postulated for this effect have not been confirmed yet experimentally but may include a promoting effect of the sugar chain on the rapid orientation of the polypeptide segment containing the attached sugar group toward the surface of the domain (Broersen et al 2005d, Imperiali & O'Connor 1999, Kern et al 1993).

*The effect of hydrophobicity on the aggregation propensity*

The relevance of hydrophobicity to aggregation is firmly established as an important determinant for the onset of aggregation (Calamai et al 2003, Chiti et al 2002, 2003, Cottingham et al 2002, DuBay et al 2004, Fernandez-Escamilla et al 2004, Valerio et al 2005, Zbilut et al 2003, 2004). First, there is evidence that in some cases non-native interactions are transiently formed to bury highly aggregation prone regions of the unfolded proteins such as exposed hydrophobic areas (Capaldi et al 2002, Hore et al 1997). Second, inclusion of a hydrophobicity term into an algorithm to predict mutational effects on peptide aggregation (TANGO) results in a better prediction of aggregation (Fernandez-Escamilla et al 2004). Third, Zbilut et al (2004) found that calculated amino acid hydrophobicity values and the subsequent derivation of theoretical hydrophobic patches significantly correlated with experimentally derived aggregation propensities obtained from literature. And last, comparison of the aggregation rate of two homologous proteins was largely explained by the difference in hydrophobicity (Calamai et al 2003). These, and other results (Chiti et al 2002, Valerio et al 2005, Zbilut et al 2003) indicate that hydrophobicity invariably plays a significant role in the aggregation propensity of proteins. The hydrophobic effect displayed by proteins upon folding is dictated by the environmental entropy (Huang & Chandler 2000). Hydration of extended hydrophobic stretches of an unfolded molecule relies on a high entropic cost to the free energy of solvation of the system and aggregation or folding are therefore driven to minimise the system energy in this way by protecting exposed hydrophobic patches through hydrophobic interactions.

*The effect of electrostatics on the aggregation propensity*

The contribution of electrostatics to aggregation has been recognised by a large number of authors in the past (e.g. de la Paz et al 2002). The electrostatic contribution to aggregation reactions can be investigated through a number of approaches. For example, variation of net charge can be brought about by mutation (see f.e. Campos et al 2004, Garcia-Mayoral et al 2003) or random modification (see f.e. Broersen et al 2005c, Weijers et al 2005). Both types of modification have been observed to effectively introduce differences in the aggregation probability of proteins. For example, Calamai et al (2003) compared the aggregation rate of two homologous proteins that were previously shown to be able to form fibrils *in vitro*. The rate of the aggregation process was evaluated upon normalising the net charge contributions to each of the homologous proteins, namely the N-terminal domain of HypF and human muscle acylphosphatase. The difference in net charge was estimated to

account for 20-25% of the difference in the aggregation rates (Calamai et al 2003). Also Chiti et al (2002b, 2003) and DuBay et al (2004) reported a significant, and linear dependence of charge to the aggregation rate. These observations can be explained on the basis of simple physical laws: when the net charge of a protein is high, the approach and interaction between distinct protein molecules can be hindered by an overall effect of electrostatic repulsion. A decrease in net charge, for example due to mutations or charge modification, leads to a reduction in the extent of such repulsions, contributing to an acceleration of the aggregation process. In our research group we used chemical charge engineering of ovalbumin. The observed differences in the aggregation rate upon modification could be explained fully by the destabilising effect of the modification (Weijers et al 2005). To support this, we first determined the thermal transition temperature of the modified species using DSC. Subsequently, the aggregation kinetics of the ovalbumin variants were subjected to incubation temperatures at a fixed temperature away from the experimentally observed thermal transition temperature. It was found that the differences in aggregation rates first observed when incubating at a fixed temperature, had now disappeared even though the morphology of the formed aggregates differed significantly (Weijers et al 2005). Together with the reported results by Calamai et al (2003), Chiti et al (2002b, 2003) and DuBay et al (2004), this led us to suggest that the charge effect is probably more complex than can be explained by the simple inclusion of net charge in the predictive models. Perhaps the sequestering of charged patches along the sequence may importantly contribute to the inferred electrostatic repulsive effect. This suggestion has in fact been explored by Zbilut et al (2004) who also verified that the charge effect displays a more complex mechanism than that of general repulsion between charged molecules.

#### *A combination of factors determines the aggregation propensity of proteins*

An understanding of the parameters involved in aggregation is fundamentally important, not only for elucidating the mechanism of aggregation but also for correctly predicting the effect of environmental factors on the aggregation propensity as well as to predict the effect of mutations or modifications. For example, Uversky et al (2000) found that folded proteins probably are able to avoid aggregation as a consequence of their low hydrophobicity and high net charge. As discussed in the previous paragraphs, a number of factors have been implied to predict protein aggregation. It has also been found that none of these effects alone is able to fully explain the observed aggregation rates and that perhaps a combination of these factors is required to model the aggregation process. Combining these

factors into a single model could possibly enable the prediction of aggregation kinetics of a divergence of polypeptide sequences. Calamai et al (2003) found a combined effect of hydrophobicity and net charge on aggregation of importance to correctly predict the aggregation properties of proteins. They calculated that the maximum change in aggregation rate that can be achieved with a single mutation that only changes charge and not hydrophobicity is 22% of that obtained by changing charge and hydrophobicity together. However, the contribution of charge to the probability of aggregation of proteins must not be underestimated as very strong electrostatic repulsion may prevent collision and the formation of subsequent stabilising hydrophobic interactions altogether (e.g. Weijers et al 2005). Bosques and Imperiali (2003) also observed that the contribution of the independent factors is not always a simple summation of factors. They found that the N-linked glycan of the human prion protein significantly reduces the rate of aggregation by promoting intermolecular disulphide formation via a cysteine residue. The stabilisation that this cross-linking incurs, has been postulated as the possible route by which aggregation is prevented through glycosylation.

#### *Existing multifactor models for predicting protein aggregation*

As it seems clear that the propensity of a given polypeptide chain to aggregate under specific conditions varies dramatically with a combination of intrinsic physico-chemical factors of the polypeptide chain, it appears useful to derive an approach to estimate the relative importance of the parameters identified to aggregation kinetics. A number of models have appeared in the past to fulfil such requirement. One of the first attempts to model aggregation propensity of polypeptide chains upon mutations has been reported by Chiti et al (2003). They provided an equation to predict the change of the aggregation rate of human muscle acylphosphatase upon mutation. The mutations targeted included a change in hydrophobicity of the polypeptide chain resulting from mutation, the propensity to convert from  $\alpha$ -helical to  $\beta$ -sheet structure and the overall charge ( $r = 0.85$ ).

In 2004, the group of DuBay et al (2004) reported a model correlating changes in aggregation rates due to mutation relative to the wild type protein of human muscle acylphosphatase with a correlation coefficient of  $r = 0.92$ . This equation includes the intrinsic properties of hydrophobicity, charge and protein concentration. The aggregation rates of other polypeptide sequences ranging from 8 to 127 residues upon single mutations were also well predictable. In addition to this, environmental factors such as pH and ionic strength were

taken into account. As this equation predicts the absolute values of the aggregation rates from the unfolded states of the proteins *in vitro*, conformational stability as a parameter is excluded.

Another model to predict aggregation propensity includes the development of the statistical mechanics algorithm TANGO by the group of Serrano (Fernandez-Escamilla et al 2004). This model appeared successful in predicting any differences in aggregation rates of a wide range of experimentally tested sequences, in the  $\beta$ -sheet aggregation kinetics of wild type primary sequences and their mutants. The algorithm is based on the identification of  $\beta$ -aggregating regions of a protein sequence. Peptides are considered as having aggregation tendency when they possess segments of at least five consecutive residues populating the  $\beta$ -aggregated state. The principle of the method considers different competing conformations:  $\beta$ -turn,  $\alpha$ -helix,  $\beta$ -sheet, the folded state and  $\beta$ -aggregates and takes into account hydrophobicity and solvation energetics, electrostatic interactions and hydrogen bonding. The evaluation of this method, as described in the paper, suggests that this method is very useful when determining the effect of single mutations on the aggregation propensity of a peptide or full-length protein. The method, however, cannot be used to compare the aggregation propensity of proteins or peptides differing widely in sequence (Fernandez-Escamilla 2004).

#### *Background and rationale for developing the AGPRO model*

The models discussed (Chiti et al 2003, DuBay et al 2004, Fernandez-Escamilla et al 2004) can lead to inconsistencies when predicting the random heat-induced aggregation behaviour of widely varying sequences and structures of proteins used in the food industry. The models already developed have been shown, however, to describe very well the ordered assembly of peptides and proteins into a  $\beta$ -sheet aggregate (Fernandez-Escamilla et al 2004) or amyloid conformation (Chiti et al 2003). DuBay et al (2004) use the aggregation propensity of a wild type protein sequence as a starting point and, on the basis of this knowledge, predict the effect of single mutations on the difference in aggregation propensity. The TANGO algorithm developed by the research group of Serrano (Fernandez-Escamilla et al 2004) uses a similar approach and is unable to compare the aggregation propensity of proteins that differ widely in sequence or structural properties. Chiti et al (2003) do not include a term describing the protein stability, which has been repeatedly shown, even though not exclusively, to significantly affect the aggregation propensity of proteins (Chiti et al 2000,



Hurle et al 1994, Kelly 1998, Quintas et al 1999, Ramirez-Alvarado et al 2000, Siepen & Westhead 2002).

In our laboratory we make use of a very diverse range of proteins varying widely in for example primary sequence, number of residues, stability, function, hydrophobicity and electrostatics and the aggregates formed are often the product of non-specific interactions driven through the balance of attracting and repulsive forces. Moreover, in food industry and biotechnological applications, heating is a commonly applied procedure to extend the shelf life of food products and this treatment often involves the generation of aggregates. Thermal stability is therefore an important parameter to take into account. It would thus be useful to be able to predict the aggregation propensities of this wide range of proteins beforehand. Even though it has been shown before that probably all polypeptide sequences can be converted into an aggregated state, the propensity to do so varies substantially. In order to compare and predict the aggregation propensities of widely varying sequences, we developed a model called AGgregation PROpensity model or AGPRO model. According to this model the aggregation rate ( $r = 0.93$ ) can be calculated by the following additive equation of parameters:

$$\text{Log } k \text{ aggregation rate (min}^{-1}\text{)} = \alpha \text{Seq}_{\text{S-S}}^{\text{SeqS-S}} + \alpha \text{Seq}_{\text{Hydr}}^{\text{SeqHydr}} + \alpha T_t^{\text{Tt}} (\text{°C}) + \alpha \text{glyc res}^{\text{glyc res}} + \alpha 0 \quad (3)$$

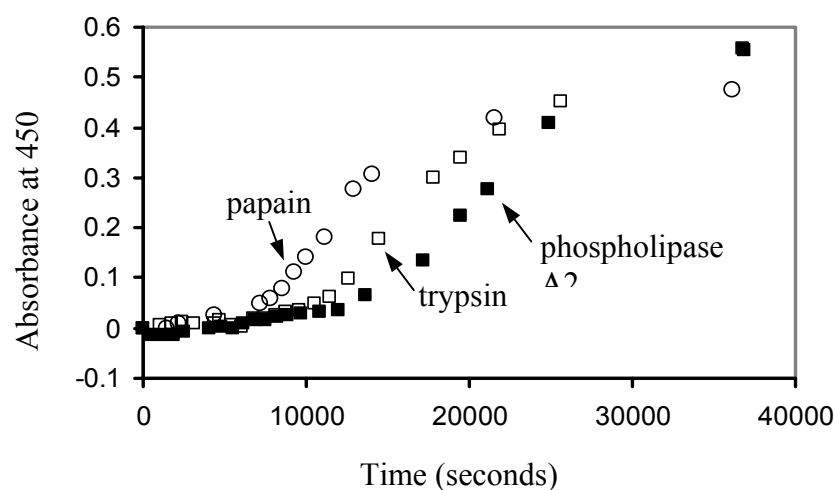
where  $\text{Seq}_{\text{S-S}}$  is the number of disulphide bonds present in the primary sequence of the polypeptide chain,  $\text{Seq}_{\text{Hydr}}$  is the hydrophobicity of the primary sequence calculated using the Eisenberg-Weiss hydrophobicity scale and  $T_t$  is the thermal transition temperature observed from experimental values using differential scanning calorimetry,  $\text{glyc res}$  is the number of glycosylated residues in the primary sequence. In the following paragraphs the validity of this model will be discussed. The model proposed by DuBay et al (2004) served as a starting point and the various parameters related to random heat-induced aggregation were included or excluded as their significance to the aggregation process was proven or rejected, respectively.

To develop and validate the method, a number of well-known proteins which significantly differ in the terms described relevant to the aggregation process have been selected and evaluated for the above-mentioned properties on the basis of their primary sequence and experimental observations. The model was based on experimental work conducted in our laboratory focused on the identification of factors responsible for aggregate

formation. The parameters involved in the equation are all intrinsic to the polypeptide sequence and include hydrophobicity, glycosylation, electrostatics, stability and number of sulfhydryl groups. The choice of these parameters originated from both our research (Broersen et al 2004, 2005a, 2005b, 2005c, 2005d) as well as previously obtained results in literature. These parameters were determined for thirteen proteins, which are i) frequently applied in food industry and/or ii) have well-known thermodynamic behaviour. At present environmental or extrinsic factors such as solution properties (pH, ionic strength and protein concentration) are excluded from the study. As it was however recognised that colloidal aspects, such as electrostatic repulsion, must play a significant role in the aggregation process, the influence of these factors on the aggregation rate could be subject to further studies. As the work performed described here is focused primarily on the prediction of random aggregation the model also does not take into account the propensity to form various types of secondary structural elements, a factor that has been well studied at present by other research groups (Chiti et al 2003, Fernandez-Escamilla 2004, Linding et al 2004).

In Table II of the appendix the properties of the selected systems are summarised and in Table III the parameters are summarised that have been related to the aggregation propensity of proteins before (see for references paragraphs on individual terms). The aggregation propensity of these proteins was experimentally tested using heat-induced aggregation (details of the method can be found in Figure 3) and the aggregation kinetics was spectroscopically evaluated by the development of turbidity in time at 450 nm. This paragraph will discuss the results of this study and evaluate the importance of the identified factors on the basis of the AGPRO model. Examples of the observed aggregation kinetics can be found in Figure 3 for papain, trypsin and phospholipase A2. The kinetic behaviour has been calculated without including the lag phase in the aggregation, which is a common feature in the study of aggregation kinetics (Pignatelli et al 1997, Silver & Trelstad 1979). After the lag-phase, single exponential behaviour was observed for all sequences tested. The origins of the lag-phase are currently unknown although a nucleation and growth mechanism has been proposed to explain this behaviour (Zurdo et al 2001). Cottingham et al (2002) investigated the rate of aggregation of acetylcholinesterase using a diversity of techniques: thioflavin-T binding assay and detection of  $\beta$ -sheet secondary structure development with CD as well as turbidimetry. They observe a disparity between the rate of aggregation observed by these techniques and ascribe this difference to the inability of turbidimetric analysis to detect particles with a hydrodynamic radius smaller than the wavelength of incident light (450 nm).

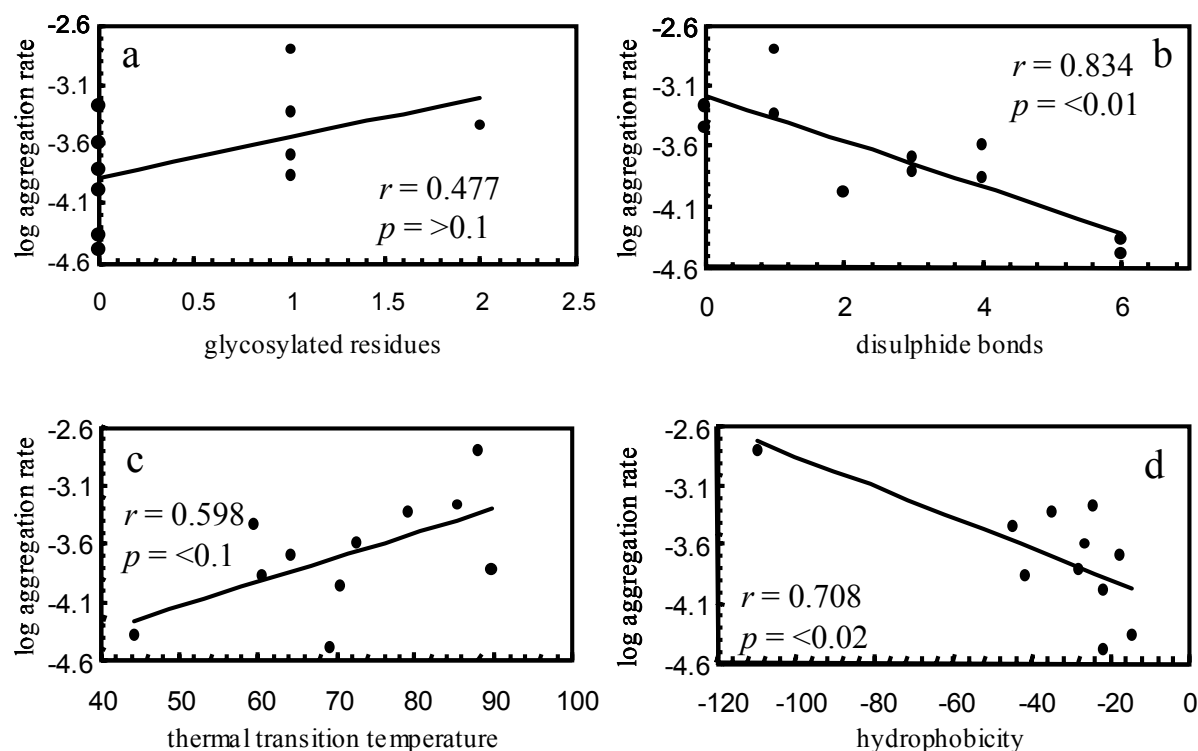
This method is thus insensitive to smaller particles and does not provide information on the initial rate of aggregation. As our analysis concerns the detection of random aggregates however, which are not driven by secondary structure formation or the specific thioflavin-T binding, turbidimetry was found a valuable method to detect differences in the aggregation rates of the selected polypeptides.



**Figure 3** The aggregation kinetics of papain, trypsin and phospholipase A2 determined using turbidimetry absorption at 450 nm.

#### *The contribution of individual components*

First, the individual contribution of the intrinsic factors selected for evaluation to the data set obtained for aggregation rates will be discussed. This discussion may, however, be of phenomenological nature as intrinsically the unfolding of a protein may be related to the exposure of sulfhydryl/disulphide groups and exposed hydrophobicity. As the aggregation reaction proceeded at a fixed temperature, it is possible to express the relationship between the logarithm of the rate constant versus the parameter varied, according to the Arrhenius equation.



**Figure 4** Correlation of the observed aggregation rate using turbidity measurements with intrinsic protein parameters. The logarithm of the observed aggregation rate is plotted against a) Number of glycosylated residues in the primary sequence, b) Number of disulphide bonds present in the primary sequence, c) Thermal transition temperature of the proteins observed using DSC, d) Hydrophobicity calculated from the primary sequence using the Eisenberg-Weiss hydrophobicity scale.

### Hydrophobicity

The hydrophobicity of the polypeptide sequences selected has been evaluated in two ways. First, the hydrophobicity of the intact proteins (in non-denaturing conditions) were evaluated using hydrophobic interaction chromatography. It has been reported that, in many cases, proteins are only required to be partially unfolded in order to contribute to the aggregation process (Calamai et al 2003, Fink 1998). The term responsible for aggregation of partially unfolded proteins has been associated to the hydrophobicity in the partially unfolded state and therefore it was hypothesised that the hydrophobicity of the folded state as well as the unfolded state can account for the observed aggregation rate. Therefore, the hydrophobicity of all sequences were also calculated using the Eisenberg-Weiss Hydrophobicity scale which assigns a positive or negative value to each individual amino acid

comprising the primary sequence of the protein (Eisenberg et al 1982). Theoretically, this value could reflect the hydrophobicity of the fully extended (unfolded) polypeptide chain. No significant correlation was found between the hydrophobicity observed using hydrophobic interaction chromatography ( $r = 0.06$ ) and the hydrophobicity of the polypeptide chains calculated using the Eisenberg-Weiss hydrophobicity scale suggesting that protein folds are unique and most likely related to the topographical location of the proteins *in vivo*.

It was found that the exposed surface hydrophobicity of the folded protein molecule did not account for the observed aggregation rates ( $r = 0.07$ ) whereas the hydrophobicity calculated using the Eisenberg-Weiss hydrophobicity scale significantly correlated to the aggregation rate at a correlation coefficient of  $r = 0.71$  (Figure 4d). It should be noted from Figure 4d that the found correlation coefficient is based for a significant share on the experimental rate for aggregation found for glycinin and its very high hydrophobicity. When not taking the data of glycinin into account, the correlation coefficient found however, still accounts  $r = 0.51$ . For the smaller proteins tested (e.g. cytochrome c, phospholipase A2) it can be assumed that a destabilisation of the folded structure may lead to the exposure of the vast majority of the non-polar residues. For larger multi-domain proteins (e.g.  $\alpha$ -amylase or ovalbumin) the calculated hydrophobicity may be overestimated in this way.

### Charge

Recently DuBay et al (2004) reported a strong negative correlation for the charge contribution with the aggregation rate. We found that charge influences the stability of the protein conformation and thus affects the rate of aggregation in an indirect manner, even though the colloidal aspect of charge repulsion significantly influences the obtained aggregate morphology upon varying charge (Broersen et al 2005c, Weijers et al 2005). To involve the charge term into our equation, the combination of values derived from  $\zeta$ -potentials, isoelectric point as well as the calculated charge at pH 7.0, were found to contain information relevant to determine the rate of aggregation. The correlation coefficients were low and not significant for all the charge parameters tested:  $r = 0.28$  for  $\zeta$ -potential,  $r = 0.31$  for isoelectric point and  $r = 0.08$  for net charge at pH 7.0. For this reason, the electrostatic term was not incorporated in the AGPRO model. In this experiment a high ionic strength (100 mM) was used to induce turbid aggregate solutions to enable the evaluation of the aggregation rate using spectrophotometry. It is very likely that this high ionic strength shields the electrostatic

repulsion exerted by the protein molecules. The effect of net charge thus does not emerge to the full extent.

#### Glycosylation

Glycosylation has been reported to affect the aggregation rate of proteins indirectly through a direct effect related on the refolding behaviour and stability parameters (Broersen et al 2005d). The SwissProt Tools provided information on the number of glycosylated residues on each sequence tested in this study and the number of glycosylated residues was correlated directly and in a, although not strongly, significant manner to the rate of heat-induced aggregation ( $r = 0.48$ , Figure 4a). The number of glycosylated residues has not been corrected for the total number of amino acid residues in each sequence, nor has the known effect of glycosylation on charge, hydrophobicity or stability parameters been included (Broersen et al 2004, 2005d).

#### Sulphydryl/disulphide groups

The number of sulphydryl and disulphide groups of the selected sequences has been evaluated by calculating the number of cysteine residues. It was found that the number of disulphide bridges shows a very strong positive correlation with the rate of aggregation ( $r = 0.83$ , Figure 4b) and, as such, this term has been incorporated in the model. This result seems rather surprising at first as earlier studies clearly indicate that disulphide groups provide solely a stabilising means of the formed protein-protein interactions and are not a driving force as such. This result can, however, be rationalised by the method used to determine the rate of aggregate formation. The formed aggregates are stabilised through disulphide interactions such that a redissociation of the formed aggregates is not possible.

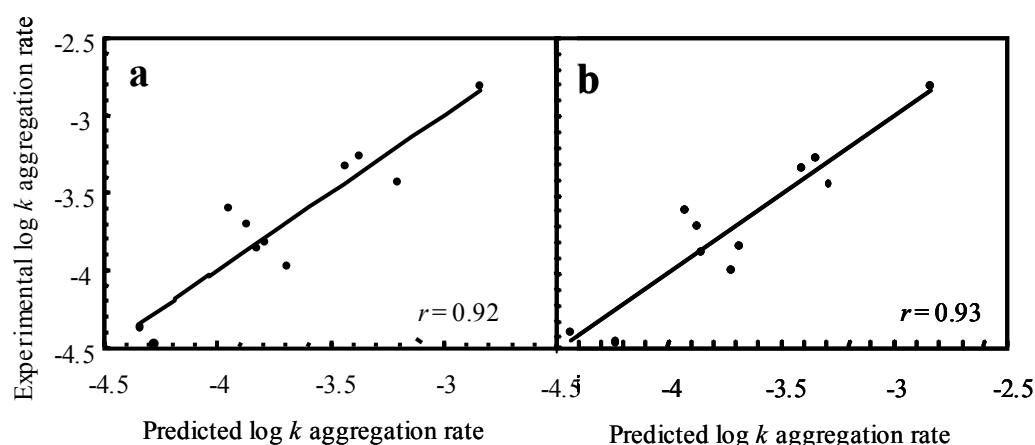
#### Thermal transition temperature

The thermal transition temperature of a protein is a measure of the thermal resistance of the polypeptide chain to unfold. The thermal transition temperatures of the individual proteins selected for this study have been investigated either using differential scanning calorimetry or the values have been derived from literature, taking into account that environmental conditions can significantly affect the thermal transition temperature observed. It was found that the thermal transition temperature showed a correlation of  $r = 0.60$  with the experimentally observed aggregation rates which was found to be significant (Figure 4c). It was also found that this factor cross correlates with other factors such as the number of

disulphide bonds in the sequence ( $r = 0.528$ ) and hydrophobicity determined using hydrophobic interaction chromatography ( $r = 0.426$ ), however, not in a significant manner ( $p > 0.05$ ). This cross-correlation can result in an apparent weaker contribution to the aggregation rate when considering a multicomponent analysis.

#### Contribution of the combination of components

As was observed from the correlation coefficients of the individual terms to predict the aggregation rate of the selected sequences, none of the terms was able to fully predict the observed aggregation rates. This finding is consistent with earlier attempts to create models predicting the ordered aggregation behaviour of proteins (Chiti et al 2003, DuBay et al 2004, Fernandez-Escamilla et al 2004, Linding et al 2004) into ordered  $\beta$ -sheet aggregation or amyloid formation. For this reason, a combination of terms is expected to predict the aggregation behaviour better. We have tested which combination of terms resulted in the highest significant correlation with observed aggregation rates and found that, starting from a correlation coefficient for the number of disulphide bonds present in the sequence of  $r = 0.83$  (Figure 5a), the addition of the sequence hydrophobicity term increases the predictability of the aggregation behaviour to  $r = 0.91$ . Addition of the thermal transition temperature further increased the correlation coefficient to  $r = 0.92$ . This apparently very small contribution was in part explained by the cross-correlation this factor shows with hydrophobicity and the number of disulphide bonds. As a large number of proteins are known in nature which do not contain disulphide bonds, the thermal stability factor may in those cases become a strong determinant of the aggregation rate, justifying its inclusion in the model. Even though the number of glycosylated residues in isolation showed a significant correlation with the observed aggregation rates the addition of this term does not increase the correlation coefficient for aggregation any further and is therefore excluded from the model.



**Figure 5** *Logarithm of the aggregation rate plotted against a) the number of disulphide bonds present in the primary sequence and hydrophobicity and b) against the combined effect of the number of disulphide bonds present in the primary sequence, the hydrophobicity calculated from the Eisenberg-Weiss hydrophobicity scale, the thermal transition temperature obtained from DSC and the number of glycosylated residues in the primary sequence.*

**Table I** *Multiple linear regression analysis of the data set of the aggregation studies involving glycosylated residues, disulphide bridges, transition temperature and hydrophobicity (calculated from sequence Eisenberg-Weiss hydrophobicity scale). Data obtained from Table III. Analysis were performed using the Analyse-It add-in for Excel and The Unscrambler.*

	$\alpha$ (significance level)	$p$ -value
Intercept	$-4.0191 \pm 0.7669$	0.0019
Glycosylated residues	$0.0347 \pm 0.1830$	0.8557
Disulphide bridges	$-0.1281 \pm 0.0564$	0.0636
Transition temperature	$0.0058 \pm 0.0089$	0.5410
Hydrophobicity	$-0.0070 \pm 0.0038$	0.1159

#### *Boundaries, demerits and limitations*

It must be noted that the primary sequence of the protein selected must be known and the thermal transition temperature of the protein, under the conditions for which the aggregation rate is predicted, is known or can be determined experimentally. Next to that, this



empirical model does not include external factors such as pH, ionic strength or incubation temperature. It has been recognised that colloidal factors, such as electrostatic repulsion, can significantly affect the rate of aggregation while the model presented only predicts the aggregation rate from an intrinsic perspective. The model can be further extended by including these colloidal and external factors as well as to vary the incubation conditions of the aggregates. To accommodate this last comment, it seems likely that other incubation conditions such as varying the pH or ionic strength effectively change the electrostatic contribution of the aggregating species and may provide a significant correlation at different conditions. It must also be stressed that the predictive value of this model is limited to random aggregation and does not include or pretend to predict ordered aggregation such as amyloid fibril formation or  $\beta$ -sheet aggregation. The prediction of the rates of these aggregate types includes a secondary structure propensity parameter in the equation as has been shown by Fernandez-Escamilla et al (2004).

Another aspect that deserves more attention is the finding in this research thesis that, even though the correlation coefficient to predict the aggregation rate by this model is  $r = 0.92$ , some proteins clearly deviate from the predicted aggregation rate, for example lysozyme and  $\beta$ -lactoglobulin. The experimental work described in this thesis clearly shows that more subtle aspects of protein stability, degree of unfolding, shielding of charge and the rate at which aggregation prone particles are generated are also significant parameters. Their contribution to the aggregation process should therefore be included in further extensions of this model.

### *Conclusions*

The approach described here involving the combination of readily described data and our own results involving chemical engineering of proteins enables us to identify the factors that affect the intrinsic propensities of polypeptide chains to show heat-induced aggregation. This model cannot, as such, be used to predict the morphology of the formed aggregates. As additional data become available it will undoubtedly be possible to further refine the proposed model, and also to consider additional parameters that influence the aggregation propensities of peptides and proteins. However, we have shown that inclusion of four determinants: hydrophobicity, disulphide bonds, thermal stability and glycosylation into a simple model can predict to a remarkable degree ( $r = 0.93$ ) the effects of intrinsic physico-chemical properties on the aggregation rates of a wide range of peptides and proteins. Its use should help us

substantially in efforts to design more rationally the aggregation properties of proteins or peptides. At present their utilisation in research programmes, for medical applications or for biotechnological purposes can be seriously limited.

It was also found that there are common principles in the fundamental determinants of aggregation of polypeptide chains despite evident differences in detail (Bucciantini et al 2002, Dobson 1999). As DuBay et al (2004) have suggested recently, the generic process of protein aggregation can be explained by the dominant role the polypeptide backbone plays in the process and that the side chains of the individual amino acid residues control the aggregation propensity of the entire chain. However, there should be a 'point-of-no-return' in the folding pathway beyond which the protein becomes 'committed' to the fully folded state. Thus there should be a stage beyond which the activation energy for reversal of the folding pathway into aggregation is too high and further traversal of the partly folded molecule on the folding pathway prevails over the aggregation pathway. Indeed it has been shown before that, once refolding is completed, the folded protein solution can be concentrated to a large extent without aggregation phenomena to take place. Similarly, there must be a step on the folding pathway where the protein becomes 'committed' to become aggregated. This point becomes even more interesting with the recent finding that, in pathological aggregation of proteins *in vivo*, it has been found that the primers of aggregation are likely to be the toxic components whereas the aggregated state actually is regarded now as an accumulated non-toxic state (Bucciantini et al 2002, Kaye et al 2003, Olofsson et al 2002, Walsh et al 2002). The results discussed in this paper indicate that the point-of-no-return strongly depends on a number of intrinsic protein factors as single sequence mutations or, for example glycosylation, can significantly alter the kinetic partitioning behaviour between folding and refolding. This point seems to be associated with the balance between attractive and repulsive forces at a single time during the folding state. That is, as long as the partly unfolded molecule exerts extensive hydrophobic exposure for example upon refolding (thus before the hydrophobic collapse has taken place), the aggregation pathway is still a likely pathway for the molecule to pave. Once the hydrophobic core is well buried and only small shifts in side residues take place, it is not likely that the protein will be prone to pave the aggregation pathway anymore. The rate of unfolding versus the rate of (effective) collision can be a major determining factor in this process.

The rapid progress in the development of new methods to study protein conformation at an atomic level, such as the development of synchrotron radiation sources and sensitive photon detectors to allow fast collection of data related to the dynamics of protein structures,

the development of time-resolved NMR and crystallography techniques pave the way for promising techniques to enable challenging visualisation of the minute conformational changes of proteins required to allow aggregation.

## Acknowledgements

The author acknowledges Marcel Meinders, René de Wijk and Carla Dullemeijer for support with statistical analysis.

## References

- Alting AC, Hamer RJ, de Kruif CG, Visschers RW (2000) *J Agric Food Chem* 48: 5001-5007
- Alting AC, Weijers M, de Hoog EH, van de Pijpekamp AM, Cohen Stuart MA, Hamer RJ, de Kruif CG, Visschers RW (2004) *J Agric Food Chem* 52: 623-631
- Anfinsen CB (1973) *Science* 181: 223-230
- Arai M, Ikura T, Semisotnov GV, Kihara H, Amemiya Y, Kuwajima K (1998) *J Mol Biol* 275: 149-162
- Arntfield SD, Murray ED, Ismond MAH (1991) *J Agric Food Chem* 39: 1378-1385
- Bagel'ova J, Antalík M, Tomori Z (1997) *Biochem Mol Biol Int* 43: 891-900
- Bhattacharjee C, Das KP (2000) *Eur J Biochem* 267: 3957-3964
- Bosques CJ, Imperiali B (2003) *Proc Natl Acad Sci* 100: 7593-7598
- Brockwell DJ, Smith DA, Radford SE (2000) *Curr Opin Struct Biol* 10: 16-25
- Broersen K, Elshof M, Hamer RJ, de Jongh HHJ (2005a) Pinpointing the origins of the effect of glucosylation on the aggregation mechanism of proteins. Manuscript in preparation
- Broersen K, Voragen AGJ, Hamer RJ, de Jongh HHJ (2004) *Biotech Bioeng* 86: 78-87
- Broersen K, van Teeffelen AMM, Vries A, de Jongh HHJ (2005b) The importance of disulphide interactions for protein aggregation formation and aggregate stability. Manuscript in preparation.
- Broersen K, Weijers M, de Groot J, Hamer RJ, de Jongh HHJ (2005c) Electrostatics controls fibril formation – part I: Charge engineering of proteins affects unfolding mechanism. Manuscript in preparation.
- Broersen K, Meinders M, Hamer RJ, Voragen AGJ, de Jongh HHJ (2005d) Protein glycosylation changes folding/refolding and the formation of aggregation-prone intermediates by affecting folding mechanism. Manuscript in preparation.
- Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG (1995) *Proteins* 21: 167-195
- Bucciantini M, Giannoni E, Chiti F, Baroni F, Formigli L, Zurdo J, Taddei N, Ramponi G, Dobson CM, Stefani M (2002) *Nature* 416: 507-511
- Calamai M, Taddei N, Stefani M, Ramponi G, Chiti F (2003) *Biochemistry* 42: 15078-15083
- Campbell ID, Dwek RA (1984) *Biological Spectroscopy*, Benjamin/Cummings, Menlo Park CA: 91-125
- Campos LA, Garcia-Mira MM, Godoy-Ruiz R, Sanchez-Ruiz JM, Sancho J (2004) *J Mol Biol* 344: 223-237
- Capaldi AP, Kleanthous C, Radford SE (2002) *Nat Struct Biol* 9: 209-216
- Chan HS, Dill KA (1998) *Proteins* 30: 2-33
- Chevet E, Cameron PH, Pelletier MF, Thomas DY, Bergeron JJM (2001) *Curr Opin Struct Biol* 11: 120-124
- Chiti F, Taddei N, Bucciantini M, White P, Ramponi G, Dobson CM (2000) *EMBO J* 19: 1441-1449

- Chiti F, Taddei N, Baroni F, Capanni C, Stefani M, Ramponi G, Dobson CM (2002) *Nat Struct Biol* 9: 137-143
- Chiti F, Stefani M, Taddei N, Ramponi G, Dobson CM (2003) *Nature* 424: 805-808
- Chiti F, Calamai M, Taddei N, Stefani M, Ramponi G, Dobson CM (2002b) *Proc Natl Acad Sci USA* 99: 16419-16426
- Clarke AR, Waltho JP (1997) *Curr Opin Struct Biotechnol* 8: 400-410
- Collini M, d'Alfonso L, Baldini G (2000) *Protein Sci* 9: 1968-1974
- Cottingham MG, Hollinshead MS, Vaux DJT (2002) *Biochemistry* 41: 13539-13547
- de la Paz ML, Goldie K, Zurdo J, Lacroix E, Dobson CM, Hoenger A, Serrano L (2002) *Proc Natl Acad Sci USA* 99: 16052-16057
- Delaye MT, Tardieu A (1983) *Nature* 302: 415-417
- Diederix REM, Ubbink M, Canters GW (2002) *Biochemistry* 41: 13067-13077
- Dill KA, Chan HS (1997) *Nat Struct Biol* 4: 10-19
- Dobson CM (1999) *Trends Biochem Sci* 24: 329-332
- Dobson CM (2004) *Sem Cell Dev Biol* 15: 3-16
- Doi E (1993) *Trends Food Sci Technol* 4: 1-5
- DuBay KF, Pawar AP, Chiti F, Zurdo J, Dobson CM, Vendrusculo M (2004) *J Mol Biol* 341: 1317-1326
- Eftink MR (1994) *Biophys J* 66: 482-501
- Eisenberg D, Weiss RM, Terwilliger TC, Wilcox W (1982) *Faraday Symp Chem Soc* 17: 109-120
- Ellis JR, Minton AP (2003) *Nature* 425: 27-28
- Enroth C, Neujahr H, Schneider G, Lindqvist Y (1998) *Structure* 6: 605-617
- Essex DW, Li M, Miller A, Feinman RD (2001) *Biochemistry* 40: 6070-6075
- Fasman GD (ed.) (1996) *Circular dichroism and the conformational analysis of biomolecules*. Plenum Press, New York
- Fava A, Iliceto A, Camera E (1957) *J Am Chem Soc* 79: 833-838
- Fenouillet E, Jones IM (1995) *J Gen Virol* 76: 1509-1514
- Fernandez-Escamilla A-M, Rousseau F, Schymkowitz J, Serrano L (2004) *Nat Biotech* 22: 1302-1306
- Ferry JD (1948) *Adv Protein Chem* 4:2-78
- Fink AL (1998) *Fold Des* 3: R9-R23
- Garcia-Mayoral MF, Perez-Canadillas JM, Santoro J, Ibarra-Molero B, Sanchez-Ruiz JM, Lacadena J, del Pozo LM, Gavilanes JG, Rico M, Bruix M (2003) *Biochemistry* 42: 13122-13133
- Guan JQ, Takamoto K, Almo SC, Reisler E, Chance MR (2005) *Biochemistry* 44: 3166-3175
- Gupta MN (1992) *Eur J Biochem* 203: 25-32
- Hamada D, Kato T, Ikegami T, Suzuki KN, Hayashi M, Murooka Y, Honda T, Yanagihara I (2005) *FEBS J* 272: 756-768
- Hamada D, Dobson CM (2002) *Protein Sci* 11: 2417-2426
- Hammarstrom P, Wiseman RL, Powers ET, Kelly JW (2003) *Science* 299: 713-716
- Helenius A, Trombetta ES, Hebert DN, Simons JF (1997) *Trends Cell Biol* 7, 193-200
- Hoffman MAM, van Mil PJJM (1997) *J Agric Food Chem* 45: 2942-2948
- Hore PJ, Winder SL, Roberts CH, Dobson CM (1997) *J Am Chem Soc* 119: 5049-5050
- Huang DM, Chandler D (2000) *Proc Natl Acad Sci USA* 97: 8324-8327

- Hurle MR, Helms LR, Li L, Chan WN, Wetzel R (1994) *Proc Natl Acad Sci USA* 91: 5446-5450
- Imperiali B, O'Connor SE (1999). *Curr Opin Chem Biol* 3: 643-649
- Jaenicke R, Rudolph R (1977) *FEBS Symp* 49: 351-367
- James TL, Oppenheimer NJ (eds.) (1994) *Methods Enzymol* 239, 813 p, Academic Press, San Diego
- Johansson JS, Manderson GA, Ramoni R, Grolli S, Eckenhoff RG (2005) *FEBS J* 272: 573-581
- Kayed R, Head E, Thompson JL, McIntire TM, Milton SC, Cotman CW, Glabe CG (2003) *Science* 300: 486-489
- Kelly SM, Price NC (1997) *Biochim Biophys Acta* 1338: 161-185
- Kelly SM, Price NC (2000) *Curr Prot Pep Sci* 1: 349-384
- Kelly JW (1998) *Curr Opin Struct Biol* 8: 101-106
- Kelly JW (1996) *Curr Opin Struct Biol* 6: 11-17
- Kern G, Kern D, Jaenicke R, Seckler R (1993). *Protein Sci* 2, 1862-1868
- Klotz IM, Heiney RE (1962) *Arch Biochem Biophys* 96: 605-612
- Konno T (2001) *Biochemistry* 40: 2148-2154
- Koseki T, Kitabatake N, Doi E (1989a) *Food Hydrocolloids* 3: 123-134
- Koseki T, Fukuda T, Kitabatake N, Doi E (1989b) *Food Hydrocolloids* 3: 135-148
- Koshland Jr DE (2003) *Angew Chem Int* 33: 2375 - 2378
- Kosinski-Collins MS, King J (2003) *Prot Sci* 12: 480-490
- Kraut EH, Sagone AL Jr (1984) *Blood* 63: 1056-1059
- Kuwajima K (1989) *Proteins* 6: 87-103
- Kuwajima K (1996) *FASEB J* 10: 102-109
- Kuwajima K, Nitta K, Sugai S (1975) *J Biochem (Tokyo)* 78: 205-211
- Ladokhin AS (2000) In: *Encyclopedia of Analytical Chemistry* (Meyers RA Ed.), John Wiley & Sons Ltd, Chichester: 5762-5779
- Land A, Braakman I (2001) *Biochimie* 83: 783-790
- Lechner H *Ber.*, 53B, 591 (1920)
- Liemann S, Glockshuber R (1999) *Biochemistry* 38: 3258-3267
- Linding R, Schymkowitz J, Rousseau F, Diella F, Serrano L (2004) *J Mol Biol* 342: 345-353
- Lumry R, Eyring H (1952) *J Phys Chem* 58: 110-120
- Luo YZ, Baldwin RL (2001) *Biochemistry* 40: 5283-5289
- Marquardt T, Helenius A (1992) *J Cell Biol* 117: 505-513
- Minton KW, Karman P, Hahn GM, Minton AP (1982) *Proc Natl Acad Sci USA* 79: 7107-7111
- Najbar LV, Considine RF, Drummond CJ (2003) *Langmuir* 19: 2880-2887
- Navea S, de Juan A, Tauler R (2002) *Anal Chem* 74: 6031-6039
- Nogales E, Wolf SG, Downing KH (1998) *Nature* 391: 199-203
- Olofsson A, Ostman J, Lundgren E (2002) *Clin Chem Lab Med* 40: 1266-1270
- Onda M, Tatsumi E, Takahashi N, Hirose M (1997) *J Biol Chem* 272: 3973-3979
- Orsini G, Goldberg ME (1978) *J Biol Chem* 253: 3453-3458
- Otzen DE, Kirstensen O, Oliveberg M (2000) *Proc Natl Acad Sci USA* 97: 9907-9912
- Owusu Apenten RK (1998) *Int J Biol Macromolecules* 23: 9-25
- Pandya MJ, Williams PB, Dempsey CE, Shewry PR, Clarke AR (1999) *J Biol Chem* 274: 26828-26837

- Papiz MZ, Sawyer L, Eliopoulos EE, North ACT, Findlay JBC, Sivaprasadarao R, Jones TA, Newcomer ME, Kraulis PJ (1986) *Nature* 324: 383-385
- Phillips NI, Jenness R (1967) *Arch Biochem Biophys* 120: 192-197
- Pignatelli P, Ferroni P, Ciatti F, Celestini A, Gazzaniga PP, Pulcinelli FM (1997) *Thromb Haemos PS327-PS327*
- Poillon WN, Kim BC, Castro O (1998) *Blood* 91: 1777-1783
- Price NC (2000) *Biotechnol Appl Biochem* 31: 29-40
- Ptitsyn OB (1995) *Advan Protein Chem* 47: 83-229
- Quintas A, Saraiva MJM, Brito RMM (1999) *J Biol Chem* 274: 32943-32949
- Ramirez-Alvarado M, Merkel JS, Regan L (2000) *Proc Natl Acad Sci USA* 97: 8979-8984
- Raschke TM, Kho J, Marqusee S (1999) *Nat Struct Biol* 6: 825-831
- Renkema JMS, van Vliet T (2002) *J Agric Food Chem* 50: 1569-1573
- Ribas de Pouplana L, Atrian S, Gonzalez-Duarte R, Fothergill-Gilmore LA, Kelly SM, Price NC (1991) *Biochem J* 276: 433-438
- Rochet JC, Lansbury PT Jr (2000) *Curr Opin Struct Biol* 10: 60-68
- Roychowdhury M, Sarkar N, Manna T, Bhattacharyya S, Sarkar T, BasuSarkar P, Roy S, Bhattacharyya B (2000) *Eur J Biochem* 267: 3469-3476
- Rudolph R, Zettlmeissl G, Jaenicke R (1979) *Biochemistry* 18: 5572-5575
- Sahin A, Lemercier G, Tetaud E, Espiau B, Myler P, Stuart K, Bakalara N, Merlin G (2004) *Exp Parasit* 108: 126-133
- Sawyer WH (1967) *J Dairy Sci* 51: 323-329
- Schaefer WC, Wilham CA, Jones RW, Dimler RJ, Senti FR (1960) *Cereal Chem* 37: 411-412
- Schmid FX (1997) In: *Protein Structure, a Practical Approach*, 2<sup>nd</sup> edition, Ch 11 (Creighton TE ed.) Oxford University Press, Oxford: 261-297
- Schwartz R, Istrail S, King J (2001) *Protein Sci* 10: 1023-1031
- Shastri MCR, Sauder JM, Roder H (1998) *Acc Chem Res* 31: 717-725
- Shimada K, Cheftel JC (1989) *J Agric Food Chem* 37: 161-168
- Siepen JA, Westhead DR (2002) *Protein Sci* 11: 1862-1866
- Silver FH, Trelstad RL (1979) *J Theor Biol* 81: 515-526
- Song Y, Azakami H, Hamasu M, Kato A (2001) *FEBS Lett* 491: 63-66
- Tani F, Shirai N, Nakanishi Y, Yasumoto K, Kitabatake N (2004) *Biosci Biotech Biochem* 68: 2466-2476
- Tanaka N, Tsurui Y, Kobayashi I, Kunugi S (1996) *Int J Biol Macromolecules* 19: 63-68
- Teipel JW, Koshland DE (1971) *Biochemistry* 10: 792-805
- Tjernberg L, Hosia W, Bark N, Thybreg J, Johansson J (2002) *J Biol Chem* 277: 43243-43246
- Uversky VN, Gillespie JR, Fink AL (2000) *Proteins* 41: 415-427
- Valerio M, Colosimo A, Conti F, Giuliani A, Grottesi A, Manetti C, Zbilut JP (2005) *Proteins* 58: 110-118
- van Koningsveld GA, Gruppen H, de Jongh HHJ, Wijngaards G, van Boekel MAJS, Walstra P, Voragen AGJ (2001) *J Agric Food Chem* 49: 4889-4897
- van Teeffelen AMM, Broersen K, de Jongh HHJ (2005) Glucosylation of  $\beta$ -lactoglobulin lowers the heat capacity change of unfolding: a unique way to affect protein thermodynamics. Accepted for publication in *Protein Sci*

- Walsh DM, Klyubin I, Fadeeva JV, Cullen WK, Anwyl R, Wolfe MS, Rowan MJ, Selkoe DJ (2002) *Nature* 416: 535-539
- Walstra P, van Vliet T (2003) In: *Progress in Biotechnology* (eds. Aalbersberg WY, Hamer RJ, Jasperse P, de Jongh HHJ, de Kruif CG, Walstra P, de Wolf FA), Elsevier 23: 9-30
- Wang C, Eufemi M, Turano C, Giartosio A (1996) *Biochemistry* 35, 7299-7307
- Wang X, Terzyan S, Tang J, Loy JA, Lin X, Zhang XC (2000) *J Mol Biol* 295: 903-914
- Weijers M, Broersen K, Barneveld PA, Cohen Stuart MA, Hamer RJ, de Jongh HHJ, Visschers RW (2005) Electrostatics controls fibril formation – part II: Elucidation of the molecular mechanism of aggregation. Manuscript in preparation.
- Wess TJ (1997) *Biotechnol Appl Biochem* 26: 127-142
- Xiong YL, Dawson KA, Wan L (1993) *J Dairy Sci* 76: 70-77
- Yagi M, Sakurai K, Kalidas C, Batt CA, Goto Y (2003) *J Biol Chem* 278: 47009-47015
- Zbilut JP, Colosimo A, Conti F, Colafranceschi M, Manetti C, Valerio MC, Webber CL, Giuliani A (2003) *Biophys J* 85: 3544-3557
- Zbilut JP, Mitchell JC, Giuliani A, Colosimo A, Marwan N, Webber CL (2004) *Physica A* 343: 348-358
- Zettlmeissl G, Rudolph R, Jaenicke R (1979) *Biochemistry* 18: 5567-5571
- Zurdo J, Guijarro JI, Jiménez JL, Saibil HR, Dobson CM (2001) *J Mol Biol* 311: 325-340

**Table II***Sample parameters of the protein systems tested*

	Source	Mw (Da)	PDB-code	glycosylation sites
Phospholipase A2	Porcine pancreas	13994	1A3D	-
Cytochrome C	Horse heart	11702	1CRC	-
Glycinin	Soy	32647	1IPJ	Asn351
Lysozyme	Hen egg white	14313	193L	-
Ovalbumin	Hen egg white	42750	1OVA	Asn292
$\alpha$ -amylase	Asp. Oryzae	52489	6TAA	Asn297
Trypsin	Bovine pancreas	23865	5PTP	-
$\beta$ -lactoglobulin	Bovine milk	18281	1BEB	-
Papaine	Carica papaya	23428	9PAP	-
Patatin	Potato	40009	1OXW	Asn92, 301
Bromelain	Pineapple stem	22831	1YAL	Asn117



**Table III** *Properties of proteins related to aggregation propensity*

	Hydrophobicity		Charge		Cysteines <sup>d</sup>		Tt <sup>e</sup>	Aggregation kinetics (log <i>k</i> sec <sup>-1</sup> ) <sup>f</sup>	
	HIC <sup>a</sup>	sequence <sup>b</sup>	ζ-potential <sup>c</sup>	PI	Charge at pH 7.0	No. cysteines			No. disulphide bridges
Phospholipase A2	1.00	-21.72	-8.4±0.2	7.4	-3	14	6	69.2	-4.48
Cytochrome C	1.00	-24.7	+3.7±0.3	10.1	+6	2	0	85.4 <sup>1</sup>	-3.27
Glycinin	0.85	-109.59	-20.4±0.2	4.5	-6	7	1	88 <sup>2</sup>	-2.81
Lysozyme	0.79	-26.14	+3.8±0.2	10.7	+8	8	4	72.7	-3.60
Ovalbumin	0.75	-34.76	-12.1±0.5	4.5	-12	6	1	79.1	-3.33
α-amylase	0.72	-41.44	-14.2±1.5	4.5	-23	9	4	60.5	-3.86
Trypsin	0.69	-14.42	-0.9±4.3	10.3	+6	12	6	44.2	-4.37
β-lactoglobulin	0.62	-21.41	-14±0.4	5.3	-8	5	2	70.6	-3.97
Papaine	0.64	-27.63	+1.3±2.8	8.9	+6	7	3	86.6	-3.81
Patatin	0.01	-45.01	-18.7±0.5	5.3	-10	0	0	59.4 <sup>3</sup>	-3.43
Bromelain	0.34	-17.31	-2.4±2.3	8.6	+4	7	3	64.1	-3.70

<sup>1</sup>Bagel'ova et al 1997, <sup>2</sup>Renkema & van Vliet 2002, <sup>3</sup>van Koningsveld et al 2001

<sup>a</sup>Hydrophobic Interaction Chromatography. A 5 mL Phenyl FF (high sub) (Amersham, Biosciences) was equilibrated with buffer (0.85M or 1M ammonium sulphate, 10 mM BisTris HCl, pH 7.0). 500  $\mu$ L of 10 mg/mL of the proteins were applied to the column in the folded state and subsequently eluted using 10 mM BisTris HCl (pH 7.0). The number reflects the M ammonium sulphate required to elute the protein.

<sup>b</sup>Hydrophobicity derived from the primary sequence. The hydrophobicity was calculated based on the Eisenberg-Weiss Hydrophobicity scale (Eisenberg et al 1982). The values in the column reflect the net hydrophobicity calculated from each individual amino acid comprising the primary sequence of the protein. Hydrophobicity and value in this column are positively correlated. Theoretically, this value reflects the hydrophobicity of the fully extended (unfolded) polypeptide chain.

<sup>c</sup> $\zeta$ -potential of the protein derived from experimental values using a Zeta-Sizer 2000 (Malvern Instruments Ltd.) in 10 mM BisTris HCl, pH 7.0 at a concentration of 10 mg/mL, centrifuged for 5 min at 2000 rpm and 5 mL injected into the flow-through cell. Measurements were performed in triplicate.

<sup>d</sup>The number of cysteines is calculated on the basis of the primary sequences provided by the Protein Data Bank.

<sup>e</sup>The thermal transition temperature is determined using a VP-DSC (MicroCal Inc. USA) Microcalorimeter at a protein concentration of 3 mg/mL in 100 mM sodium phosphate buffer (pH 7.0), scanning rate of 1°C/min from 20°C to 100°C. The thermal transition temperature is defined as the endothermic peak.

<sup>f</sup>The aggregation kinetics is determined at a protein concentration of 10 mg/mL in a 100 mM sodium phosphatebuffer at pH 7.0, heating at 70°C and following the changes in absorbance at 450 nm using a spectrophotometer.

# SUMMARY

## **Introduction**

Proteins are widely used in the food industry as a texturiser. This is based on the ability of proteins to aggregate. The need to understand the aggregation process of proteins is clear from the fact that different proteins show very different aggregation behaviour. This currently limits the predictability of use and exchangeability. Next to this food-related context, protein aggregation has also been of interest in the area of pathogenesis. This shows that a fundamental understanding of the origins of the protein aggregation process significantly contributes to the development of, on the one hand, well-controlled structural properties of food products and the application of proteins as functional ingredients from a wider range of sources.

On the other hand, more recently a firm relation between protein aggregation and certain diseases has been established. This raises the opportunity to consider possible routes for prevention or treatment of these diseases through an extended understanding of protein aggregation. A large number of studies have been reported over the years, which show a relation between various protein parameters and the probability of aggregation. For many of the factors identified a firm relation with the probability of protein aggregation has been shown. However, it now seems that not a single component but rather a combination of factors is responsible for the occurrence of aggregation. Also the mechanism by which protein aggregation can be either inhibited or promoted is unclear.

## **This thesis**

This thesis describes the fundamentals of unfolding, refolding and aggregation of food proteins. The parameters investigated include glycosylation (chapters 2 to 5), electrostatics (chapters 6 and 7), sulfhydryl groups (chapter 8) as well as stability parameters (chapters 3, 4 and 6). The approach we used was the chemical modification of some well-known proteins using different approaches and target functional groups. These proteins were isolated and then modified using the Maillard reaction to obtain glucosylated proteins, succinylation or methylation to affect the charge interactions and thiolation to introduce sulfhydryl groups into the protein. This approach bears the advantage that a large quantity of proteins can be obtained, the protein is already folded upon modification and that the impact of the

modification on the protein structure is limited (chapters 2 to 8). A drawback of using this method is that the modification occurs more or less random. Hence, the resulting modified proteins bear a certain degree of heterogeneity, which has to be accounted for in the interpretation of data.

Chapters 2 to 5 describe that the structural properties upon unfolding in combination with the time the protein molecule resides in an aggregation prone state largely determine the propensity to aggregate. It was found that this residence time can be limited through the introduction of sugar chains to the protein. The mechanism behind this effect was elucidated in two ways. First, as was deduced from a decreased  $\Delta C_p$  upon glucosylation (chapter 4), the hydrophobic exposure of the unfolded protein is decreased with glycosylation. It was suggested that the sugar chain somehow orientates itself toward the hydrophobic regions of the polypeptide chain resulting in a reduced hydrophobic exposure thus limiting the possibilities for aggregation. Second, it was found in chapters 3 and 5 that glycosylation increases the refolding efficiency of proteins through an increased refolding rate. This was mechanistically explained by the specific tagging of regions of the protein to guide their location in the folded state, as has been proposed earlier in literature. These observations and subsequent mechanisms suggest that sugar chains have an indirect effect on the aggregation process by affecting the kinetic behaviour and hydrophobicity of the unfolded state.

A high net charge has been generally regarded as a repulsive force, which inhibits the aggregation process. Chapters 6 and 7 explore the effect of electrostatics on the aggregation behaviour of proteins. It was found that charge has a profound effect on protein stability, presumably by a disruption of ion pairs, as was postulated before (chapter 6). Chapter 6 further describes that upon correcting for the observed changes in stability, the rate of the generation of unfolded proteins was not directly affected by the changed electrostatics. Further, the existence of thermodynamically stable intermediates was suggested to be induced by the elimination of charge but it was also observed that this intermediate was well able to participate in the aggregation process. Chapter 7 describes the effect of charge on the subsequent aggregation process. It was found that increasing net charge can significantly affect the morphology of the formed aggregates.

Chapter 8 reports on the role of sulfhydryl groups on the rate of aggregation. It was found that, even though literature often suggests a causal relation between disulphide bond formation and the rate of aggregation, no such connection exists in the cases studied here. We suggest that hydrophobic interactions play a dominant role as a driving force for aggregation

while disulphide bonds merely stabilise the formed aggregates and plays a role in determining the aggregate morphology.

The discussion places the factors identified to affect the aggregation propensity into perspective. It was found, by comparing the aggregation rates of a number of proteins, that there are common principles that predict the aggregation of polypeptide chains despite evident differences in detail. It was also suggested that there should be a 'point-of-no-return' in the folding pathway beyond which the protein becomes 'committed' to the fully folded state. Thus there should be a stage beyond which the activation energy for reversal of the folding pathway into aggregation is too high and further traversal of the partly folded molecule on the folding pathway prevails over the aggregation pathway. This suggestion relates to the impact of folding rate on the aggregation propensity or refolding yield reported in chapters 3 and 5. This point-of-no-return apparently depends on a number of intrinsic protein factors and can significantly alter the kinetic partitioning behaviour between folding and refolding (chapter 5). As long as the (partly) unfolded molecule exerts extensive hydrophobic exposure, the aggregation pathway is still preferred. In the discussion the relative importance of all the intrinsic factors is discussed and a predictive model for aggregate probability is presented.

## **Conclusions**

This thesis has clearly shown that both structural (e.g. glycosylation, hydrophobicity) and thermodynamic parameters directly or indirectly control the aggregation propensity of proteins. It was also found that protein aggregation seems to be a generically defined process over a wide range of proteins. We have shown that factors such as hydrophobicity and the presence of disulphide bonds rationalise and predict to a remarkable degree the aggregation rates of a range of proteins. Within a biotechnological perspective, this understanding will support the preparation of proteins with for example a specific thermostability or aggregation behaviour. In the food industry, insight in the aggregation properties of proteins provides the option to exchange proteins from various sources to obtain desired textural properties in food products. Within the area of disease prevention, a significant understanding of protein aggregation provides a template for suggesting new routes for the treatment of disease. The use of this model will thus help us substantially in efforts to understand the fundamentals or modify more rationally the aggregation properties of proteins with applications in various fields such as food science, biotechnological applications and prevention of disease.

# SAMENVATTING

## Introductie

De noodzaak om het aggregatieproces van eiwit te beschrijven, is in het verleden benadrukt door de gevestigde associatie ervan met de functionaliteit van eiwit in levensmiddelen. Daarnaast geniet het fenomeen eiwitaggregatie ook aandacht binnen de ziekteleer. Dit belang toont aan dat een fundamenteel begrip van de oorsprong van het eiwitaggregatieproces significant bijdraagt aan de ontwikkeling van de mogelijkheid om de structuur van levensmiddelen te beïnvloeden. Ook resulteert deze kennis in de toepassing van een grotere reeks van eiwitten als functionele ingrediënten uit een grotere diversiteit van bronnen. Daarnaast rationaliseert de relatie tussen eiwitaggregatie en het ontstaan van diverse ziektebeelden de mogelijkheid van preventie of behandeling van deze ziekten. Door de jaren heen is een groot aantal publicaties verschenen waarin de invloed van diverse eiwit gerelateerde parameters op de mogelijkheid tot aggregatie werden belicht. Voor veel van de factoren die in deze publicaties zijn geïdentificeerd, is ontdekt dat ze in belangrijke mate bijdragen aan het aggregatieproces. Echter, uit de resultaten van een groot aantal van deze publicaties blijkt het mogelijk te zijn dat de combinatie van de eerder geïdentificeerde factoren een grotere rol speelt bij het bepalen van de aggregatie, dan de individuele factoren op zich. Bovendien is het belang van deze factoren voor wat betreft het moleculaire mechanisme van het aggregatieproces vooralsnog onduidelijk.

## Dit proefschrift

Dit proefschrift beschrijft de grondbeginselen van ontvouwen, hervouwen en aggregatie van eiwitten die voorkomen in voedingsmiddelen. De grondbeginselen die zijn onderzocht zijn eigenschappen van eiwitten zoals glycosylering (hoofdstukken 2 tot en met 5), electrostatica (hoofdstukken 6 en 7), sulfhydrylgroepen (hoofdstuk 8) en stabiliteitsparameters (hoofdstukken 3, 4 en 6). De aanpak die hiervoor is gebruikt, is die van chemische modificatie van een aantal bekende eiwitten. Na een isolatieprocedure van het eiwit uit de natuurlijke bron, werd de Maillard reactie toegepast om het eiwit te glycosyleren. Succinylering of methylering is gebruikt om de lading te modificeren en thiolisering is toegepast om sulfhydrylgroepen op het eiwit te introduceren. Deze benadering heeft het voordeel dat een grote hoeveelheid gemodificeerd eiwit wordt verkregen en dat het eiwit in beginsel een

gevouwen conformatie heeft waarna de modificatieprocedure wordt uitgevoerd. Daarnaast bleek uit publicaties dat de invloed van deze modificatieprocedure nauwelijks consequenties heeft voor de eiwitstructuur (hoofdstukken 2 tot en met 8). Een nadeel van chemische modificatie is dat de modificatie min of meer aselekt plaatsvindt op het oppervlak van het gevouwen eiwit en dat het eindproduct enigszins heterogeen is. Deze heterogeniteit moet in acht worden genomen bij de interpretatie van de data.

Uit de resultaten van hoofdstukken 2 tot en met 5 kan worden opgemaakt dat de structurele eigenschappen van eiwitten, ten gevolge van ontvouwing en het tijdframe waarin het eiwitmolecuul zich in een situatie bevindt waarin het vatbaar is voor aggregatie, in belangrijke mate de mogelijkheid tot aggregatie bepalen. Glycosylering van een eiwit blijkt de residentietijd in deze voor aggregatie vatbare fase te verkorten. Het mechanisme hier achter kan op twee manieren worden uitgelegd. Ten eerste kan uit de verlaagde  $\Delta C_p$  na glycosylering (hoofdstuk 4) worden afgeleid dat de blootstelling van hydrofobe delen aan de omgeving in de ontvouwen toestand verminderd is door het glycosyleren. Deze bevinding duidt op een mogelijke oriëntatie van de covalent gebonden suikerketen richting de hydrofobe regio's van de polypeptide keten wat resulteert in een verminderde schijnbare hydrofobiciteit. Daarnaast wordt in de hoofdstukken 3 en 5 gerapporteerd dat glycosylering de efficiëntie van het hervouwen van eiwitten verbeterd door een verhoogde snelheid van hervouwing. Deze observatie kan worden verklaard aan de hand van het labelen van geglycosyleerde regio's die in de gevouwen toestand aan de omgeving geëxposeerd moeten zijn en zodoende de weg wijzen om een correcte conformatie te kunnen realiseren. Deze observaties hebben tot de aanname geleid dat suiker-ketens het aggregatieproces op een min of meer indirecte manier beïnvloeden. Ten eerste door in te werken op het kinetisch gedrag tijdens vouwing en ten tweede door de hydrofobiciteit van het ontvouwen eiwit te verlagen.

Over het algemeen wordt aangenomen dat een sterke netto lading het aggregatieproces remt of zelfs onmogelijk maakt. De hoofdstukken 6 en 7 beschrijven het effect van electrostatica op het aggregatiegedrag van eiwit. De resultaten tonen aan dat de netto lading een verregaand effect heeft op de eiwitstabiliteit (hoofdstuk 6). Verder wordt in hoofdstuk 6 beschreven dat het corrigeren van de ontvouwkinetiek voor de waargenomen veranderingen in stabiliteit, resulteert in een overeenkomstige snelheid van de vorming van ontvouwen eiwit. Electrostatica lijkt dus niet direct invloed te hebben op de snelheid van de vorming van ontvouwen eiwit. Daarnaast wordt blijkt dat de ladingsverandering van het eiwit kan leiden tot het sterk bevolken van een thermodynamisch stabiele intermediaire toestand die slechts

gedeeltelijk gevouwen is. Eiwitten die deze toestand bevolken zijn ook in staat om deel te nemen aan het aggregatieproces en volledige ontvouwing is dus geen voorwaarde hiervoor. Hoofdstuk 7 beschrijft het effect van lading op het aggregatieproces als gevolg van (gedeeltelijke) ontvouwing. De netto lading blijkt een belangrijk remmend effect te hebben op de snelheid van aggregatie door de toenemende electrostatische repulsie maar zelfs bij een extreem negatieve lading (-26) vindt aggregatie plaats. De morfologie van de gevormde aggregaten wordt echter sterk beïnvloed door de electrostatische contributie.

In hoofdstuk 8 wordt de rol van sulfhydrylgroepen op de snelheid van disulfidebrugvorming als drijvende kracht voor aggregatie in twijfel getrokken. Hoewel in de literatuur vaak een causaal verband wordt beschreven tussen de aanwezigheid van sulfhydrylgroepen en de snelheid van aggregatie, hebben wij geen aanwijzingen gevonden dat deze relatie bestaat. Gebaseerd op de uitgevoerde experimenten werd verondersteld dat hydrofobe interacties een rol spelen als drijvende kracht voor het aggregatieproces terwijl disulfide bruggen voornamelijk de gevormde aggregaten in belangrijke mate kunnen stabiliseren. Dit wordt ondersteund door bevindingen in de literatuur waarin beschreven staat dat de kinetiek van het vormen van hydrofobe interacties sneller is dan die van het vormen van disulfide bruggen.

De onderlinge relaties tussen de parameters die in de hoofdstukken 2 tot en met 8 zijn geïdentificeerd als van belang zijnde voor het aggregatieproces, worden in de discussie gemodelleerd. Door de aggregatiesnelheden van een aantal eiwitten te vergelijken, is gebleken dat er gemeenschappelijke principes bestaan die het aggregatiegedrag van eiwitten voorspellen. Ook het feit dat er kleine bijdragen aan dit proces zijn van een aantal minder belangrijke factoren kwam aan het licht. Tevens wordt een ‘point-of-no-return’ principe geïntroduceerd waarbij wordt aangenomen dat, als het eiwit tijdens het vouwproces een kritisch punt passeert, het volledig toegewijd zal zijn aan vouwing. Na dit punt wordt reversie van het proces richting aggregatie belemmerd door een zeer hoge activeringsenergie en zal het eiwit volledig vouwen in zijn originele conformatie. Deze suggestie relateert aan de eerdere bevinding dat de vouwsnelheid in belangrijke mate de opbrengst van hervouwen eiwitten bepaald (hoofdstukken 3 en 5). Dit ‘point-of-no-return’ wordt gereguleerd door een aantal factoren die inherent zijn aan het eiwit en deze kunnen de kinetische verdeling van eiwitten tussen hervouwing en aggregatie bepalen (hoofdstuk 5). Zolang het (gedeeltelijk of geheel) ontvouwen eiwit nog een aanzienlijke hydrofobiciteit heeft, is aggregatie zeer waarschijnlijk de route die het zal volgen. Naarmate de hydrofobe regio's meer in de kern van het eiwit worden verborgen, wordt het steeds onwaarschijnlijker dat het eiwit zal aggregeren. In de discussie wordt het relatieve belang van de intrinsieke factoren, die in voorgaande



hoofdstukken zijn geïdentificeerd, bediscussieerd in het kader van dit ‘point-of-no-return’ principe.

### **Conclusies**

Dit proefschrift laat zien dat er een aantal parameters bestaat, zowel structureel als thermodynamisch, die direct of indirect controle kunnen uitoefenen op de aggregatiemogelijkheden van eiwitten. Daarnaast wordt aangetoond dat de mogelijkheid tot aggregatie een algemene eigenschap is voor een grote reeks aan eiwitten. Een aantal factoren zijn geïdentificeerd die de aggregatiesnelheid rationaliseren en een grote voorspellende waarde hebben. Deze factoren zijn kinetiek, hydrofobiciteit en disulfidebruggen. Het gebruik van het ontwikkelde model waarin deze factoren worden gebruikt om de aggregatiesnelheid te voorspellen is van substantieel belang voor het rationeel modificeren of begrijpen van de grondbeginselen van aggregatie eigenschappen van eiwitten binnen diverse disciplines zoals levensmiddelenwetenschappen, biotechnologische toepassingen en ziekteleer.



# DANKWOORD

Een groot aantal personen hebben meegeholpen aan dit proefschrift zoals het nu voor u ligt.

Allereerst wil ik mijn dagelijkse begeleider Harmen de Jongh bedanken. Voor jouw grenzeloze vertrouwen in mijn capaciteiten als eiwit onderzoekende ‘pitbull’ en de kans die je me gegeven hebt als ‘ex-voedingsmiep’ ben ik je bijzonder dankbaar. Jouw enorm positieve instelling en kennis maakten het mogelijk dat ik mij onbezorgd als eiwit onderzoeker kon ontwikkelen.

Mijn zeker zo grote dank gaat uit naar mijn twee promotoren Rob Hamer en Fons Voragen. Rob, bedankt voor je niet aflatende enthousiasme om je in alles te verdiepen dat mijn project in hield. Jouw verfrissende inzichten en ideeën zijn van onmisbare waarde geweest voor de kwaliteit van het onderzoek. Fons, door jouw visie als koolhydraat specialist op mijn hoofdstukken was ik in staat om te zien wat de algemene waarde was van het werk en het zo te formuleren dat het leesbaar werd voor een breder publiek.

Julia en Peter, kantoorgenootjes, met jullie heb ik enorm plezier gehad. Wij waren het enige kantoor met een eigen huisdier, Harmen de Japanse Vechtviss en opvolgers en een discobal. Peter’s bijzondere 25<sup>ste</sup> verjaardag waarin we op zoek gingen naar een akela (een nutteloze actie bleek later), het elke week bedenken van de ‘employee of the week’ en andere grappen maakten het er naar dat we al snel de naam ‘kippenhok’ verworven.

Tijdens mijn AIO-tijd heb ik gewerkt voor het WCFS, team B009, berucht en befaamd omdat we bijna elk jaar 100% van de teambonus in de wacht sleepten, vooral door de ‘cementix’ van Harmen natuurlijk. De uitjes die we ons daardoor konden veroorloven naar de Ardennen, Moskou en Valencia waren erg bijzonder. Een aantal van jullie wil ik persoonlijk bedanken. Ten eerste Marcel Meinders, alias hardcore Willie Wortel als het ging om fysica. Je hebt me echt fantastisch geholpen met het berekenen van activerings- en vrije energieën en het oplossen van Arrhenius vergelijkingen. Verder natuurlijk Jolan de Groot en Hans Kusters. Jullie beiden hebben me veel geleerd over een hele serie biochemische eiwit metingen. Peter, naast ‘roomie’ was je ook altijd bereid tot discussies over eiwit gedrag (en de wondere wereld van qigong trouwens). Eiwitten werden altijd ware persoonlijkheden tijdens deze gesprekken, van ‘lui’ tot ‘agressief’, en ‘ronduit vervelend’, we maakten er zelfs limericks van!

‘Mijn’ studenten: Anneke, Annemarie en Marijke, bedankt voor jullie tomeloze inzet tijdens jullie afstudeervakken. Ik heb veel van jullie geleerd en ik hoop jullie ook van mij. Anneke, ik was elke dag bang dat je tweelingzus op jouw plek zou zitten en ik het niet zou merken. Annemarie, je grote liefde ontdekken tijdens je afstudeervak is ook niet alles. Marijke, eindelijk eens iemand die ook een volgens Harmen ‘levensgevaarlijke’ sport deed, rugby! Daarnaast alle mensen van Levensmiddelenchemie, bedankt dat jullie er waren voor meer dan vier jaar en voor de fantastische

Japan reis. Ik zal vast nooit meer zo kunnen lachen om Japanners die ‘hoofd, schouders, knie en teen’ doen!

Naast de mensen van onze eigen vakgroep zijn er veel mensen geweest van wie ik veel heb geleerd: Willem Norde, Willem van Berkel, Yves Bollen, Ton van Vliet en Maarten Engel hebben allen op diverse manieren bijgedragen tot wat het proefschrift is geworden. Johan Hazekamp, bedankt voor het leveren van prachtige TEM plaatjes voor de publicaties. René de Wijk, fijn dat je me hielp met statistische bewerkingen van mijn data. Het WCFS AIO Top Forum, bedankt voor de gezelligheid en samenwerking. Nu ben ik rijp voor de gemeenteraad!

Zonder vriendinnen ... geen proefschrift! Bedankt meiden voor jullie gezelligheid, etentjes, sportuitstapjes ... vrijgezellenfeestje...! Meike: Vielen Dank, je bent een topvriendin, ik ga onze dinsdag avond mensa hap, slappe lach om niks, kletsavonden met cocktails en je grappige ‘Deutsche’ accent missen in Cambridge! Anna: muchas gracias, Eleni: ευχαριστω πολυ!, Hilke, Mas en Mirei: bedankt! Mirei, volgens mij is het voor niemand een verrassing is dat jij mijn paranimf bent. Mountainbiken tijdens de ‘dark hours’ van het ovalbumine zuiveren, 24-uur Grenoble shifts die vooral te harden waren dankzij jouw aanwezigheid, een fles wijn, een rode fiets met fietsmand en een digitale camera! En natuurlijk twee gezamenlijk hoofdstukken in dit proefschrift. Maar nee, je mag nog steeds niet mijn hele huis in Cambridge mozaïeken ;-)! Also ευχαριστω και φιλια to my other paranimf and friend Eleni. I got to know you during our United Biscuits time in the UK and your fantastic positive vision on the world and all the hours of fun you, Anna and I had together visiting about half of the European capitals.

Graag wil ik ook mijn dank uitspreken naar Sydney Leijenhorst Sensei. Ook al begreep een hele vakgroep er niks van wat er leuk aan is om bont en blauw rond te lopen, ik weet zeker dat ik zonder jouw training persoonlijk niet zo gegroeid zou zijn als nu. Arigato gozaimashita! Alle karatemaatjes: bedankt voor de gezellige Hamburg gasshuku’s, de in mijn geheugen gegrifte Japan ervaring (over 4 jaar zeker weer!) en de vele leuke trainingsuren. Hier kon ik menig frustrerend promotie uur op teren.

Ook wil ik mijn ouders, Zita en Gerben bedanken. Jullie waren een rustpunt tijdens het promotie onderzoek.

En dan natuurlijk Arjen. Jij met je grappige en interessante karakter zult me altijd blijven boeien. Je hebt me aangehoord tijdens frustratie, blijdschap en andere zaken en toch wilde je met me trouwen! Je steunt me door en door, staat achter me en bent een super partner! Bedankt!

Als laatste wil ik mezelf bedanken. Zonder mezelf had dit boekje er uiteindelijk toch ook niet gelegen. Dus... Kerensa: Bedankt man! (schouderklopje).

# CURRICULUM VITAE

Kerensa Broersen werd geboren op 8 juni 1975 te Hoorn. In 1991 behaalde zij haar MAVO diploma en in 1993 haar HAVO diploma aan de Copernicus te Hoorn. In datzelfde jaar begon zij aan de opleiding Voeding & Diëtetiek aan de Hogeschool van Amsterdam. Haar stages heeft ze gelopen bij het Academisch Medisch Centrum in Amsterdam en Campina Melkunie te Woerden. Haar afstudeerstage heeft zij uitgevoerd in samenwerking met het Academisch Medisch Centrum te Amsterdam met als specialisatie 'Micronutriënten bij HIV'. Na haar afstuderen in 1997 begon zij in datzelfde jaar de opleiding Nutrition & Food Management aan Huddersfield University in Engeland. Zij studeerde af in 1998. Vanaf september 1998 tot en met april 2000 was zij werkzaam als levensmiddelentechnoloog bij United Biscuits te High Wycombe, Engeland. Van april 2000 tot en met december 2000 heeft zij ervaring opgedaan in het gebruik van transmissie elektronen microscopie van polymeren gelen bij York University in Engeland. Vanaf december 2000 tot en met januari 2005 was zij als Assistent in Opleiding in dienst van Wageningen Universiteit bij de leerstoelgroep Levensmiddelenchemie. Gedurende deze vier jaar was zij gedetacheerd bij het Wageningen Centre for Food Sciences en de resultaten van het promotie onderzoek die staan beschreven in dit proefschrift zijn uitgevoerd bij de vakgroep Levensmiddelenchemie van Wageningen Universiteit. Vanaf 1 oktober 2005 zal zij als postdoc gaan werken op het Medical Research Centre (MRC), Laboratorium voor Moleculaire Biologie in Cambridge, Engeland, in de onderzoeksgroep van Bazbek Davletov aan  $\alpha$ -synucleïne aggregatie.

## Trainingsactiviteiten

Workshop 'Structure and Function of Biomolecules'	Poland	2004
Studiegroep Eiwitonderzoek Lunteren	The Netherlands	2001-2004
PhD study trip to Japan	Japan	2004
17 <sup>th</sup> Symposium of the Protein Society in Boston	USA	2003
Symposium on Industrial Proteins	The Netherlands	2001, 2003
Protein aggregation, an interdisciplinary school	France	2002
PhD study trip to UK	UK	2002
Course on Reaction Kinetics in Food Science	The Netherlands	2001
Course on Protein Engineering in Agro- and Food Biotechnology	The Netherlands	2001

# LIST OF PUBLICATIONS

Broersen K, de Jongh HHJ Potential of  $\beta$ -lactoglobulin glycosylation for industrial applications. *Industrial Proteins* 2003, **11**: 6-8

Broersen K, Voragen AGJ, Hamer RJ, de Jongh HHJ Glycoforms of  $\beta$ -lactoglobulin with improved thermostability and preserved structural packing. *Biotechnology and Bioengineering* 2004, **86**: 78-87

Broersen K, Meinders M, Hamer RJ, de Jongh HHJ Protein glycosylation changes folding/refolding and the formation of aggregation-prone intermediates by affecting folding mechanism. Submitted

Broersen K, Elshof M, Hamer RJ, de Jongh HHJ The aggregation mechanism of proteins is affected by glycosylation. Submitted

Broersen K, Weijers M, de Groot J, Hamer RJ, de Jongh HHJ Electrostatics controls fibril formation – part I: Charge engineering of proteins affects unfolding mechanism. Submitted

Broersen K, van Teeffelen AMM, Vries A, de Jongh HHJ The importance of disulphide interactions for protein aggregation formation and aggregate stability. Submitted

Broersen K, Meinders M, de Jongh HHJ, Hamer RJ Factors determining protein stability and aggregation. Manuscript in preparation

Kosters HA, Broersen K, de Groot J, Simons J-WFA, Wierenga P, de Jongh HHJ Chemical processing as a tool to generate ovalbumin variants with changed stability. *Biotechnology and Bioengineering* 2003, **84**, 61-70

van Teeffelen AMM, Broersen K, de Jongh HHJ Glucosylation of  $\beta$ -lactoglobulin lowers the heat capacity change of unfolding: a unique way to affect protein thermodynamics. *Protein Science* 2005, **14**: 2187-2194

Weijers M, Broersen K, Barneveld PA, Cohen Stuart MA, Hamer RJ, de Jongh HHJ, Visschers RW Electrostatics controls fibril formation – part II: Elucidation of the molecular mechanism of aggregation. Submitted