Department of Chemical and Process Engineering

Nucleation of β - Lactoglobulin Clusters in Solvent-Induced Denaturation

Eva Álvarez González

Degree of Doctor of Philosophy

2011

DECLARATION OF AUTHORS RIGHTS

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed:

Date:

To my family - Mercedes, Justo, Adelina, Gregorio and Ana -
for your encouragement and unconditional support
over the last four years.

ABSTRACT

The formation of nucleated clusters of β-Lactoglobulin (β-Lg) under alcohol-induced denaturation has been studied by applying spectroscopic and scattering techniques. The concentration dependence of the cluster formation process of β-Lg in 50% (v/v) buffer-ethanol solutions was study at pH 7 and 22°C. Near UV circular dichroism results indicated that the tertiary structure of the native protein was completely lost upon addition of ethanol. Changes in the secondary structure were characterised by a drastic increase of the α-helical content (and decrease of β-sheets content) immediately after ethanol addition. The percentages of these two types of secondary structure evolved further during the 10 days of incubation at room temperature attained an equilibrium state which differs from the secondary structure found in the native protein: while dimeric β-Lg in its native state has a majority content of β-sheets, clusters presented a majority content of α-helix. Similar results on the evolution of the secondary structure of the protein upon ethanol addition were found with attenuated total reflection infrared spectroscopy (ATR-IR). ThT-binding fluorescence spectroscopy revealed that the kinetics of structural rearrangement follows a nucleation dependent growth mechanism for the lowest protein concentration studied here (0.5 mg/ml and 1 mg/ml) where a lag phase, an exponential growth, and a plateau phase could be clearly identified and quantified. For higher protein concentrations (from 4 to 40 mg/ml) the structural transition kinetics of the protein was much faster and no lag phase was observed. CD and fluorescence results suggest that the high amount of α-helix structure formed immediately after ethanol addition for the lowest concentrated samples acts as an activation barrier and slows down the nucleation process leading to a distinct lag phage. Kinetic of cluster growth was monitored by dynamic light scattering, indicating that clusters grew rapidly after ethanol addition reaching a constant hydrodynamic radius shortly after the organic solvent was present in solution. Correlation with spectroscopic results revealed that structural changes were still taking place even after the cluster size attained its equilibrium value for high concentrations of protein (10 and 40 mg/ml). DLS results were in good agreement

with those obtained by small angle X-ray scattering (SAXS) and were consistent with compact sphere-like structure of clusters. Our results indicate that clusters formed in the presence of 50% (v/v) of ethanol corresponded to a thermodynamically controlled intermediate on the β-Lg aggregation pathway. However, the internal structure of these clusters must contain suitable β-sheet arrangements with enhanced binding of the ThT dye to explain the fluorescence intensity increase found when the clusters were formed.

GLOSSARY OF ABBREVIATIONS

α-La	α-Lactalbumin
β-Lg	β-Lactoglobulin
[a.u]	Arbitrary units
ACF	Autocorrelation function
AD	Alzheimer's disease
AFM	Atomic force microscopy
ATR-IR	Attenuated total reflection infrared spectroscopy
BSA	Bovine serum albumin
CD	Circular dichroism
CR	Congo Red
Cys	Cysteine
DLS	Dynamic light scattering
DNA	Deoxyribonucleic acid
DSC	Differential scanning calorimetry
DTA	Differential thermal analysis
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunoadsorbent assays
EM	Electron microscopy
KRS-5	Bromoiodide
MRW	Mean residual weight
NMR	Nuclear magnetic resonance
NRMSD	Normalised root mean square deviation
PD	Parkinson's disease
pI	Isoelectric point
pdb	Protein data bank
PTFE	Polytetrafluoroethylene
PXRD	Powder X-ray diffraction
RNA	Ribonucleic acid
SAXS	Small angle X-ray Scattering
SDS	Sodium dodecyl sulphate
SLS	Static light scattering
TEM	Transmission electron microscopy
ThT	Thioflavin T
WPI	Whey protein isolate
TFE	Trifluoroethanol

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Jan Sefcik for his dedication and guidance during this project. Thanks also to Dr. Nadeem Javid for his help and contributions to my work.

Many thanks to Dr. Sharon Kelly from the Molecular and Systems Biology Department at the University of Glasgow for all her help and the warm welcome to her lab where I thoroughly enjoyed work.

Thanks are given to the laboratory technician Mr. S. Adams for all his technical support and for making my work in the lab more enjoyable. Thanks to Alba Turja, my friend and colleague in the Chemical and Process Engineering Department who made my second year of PhD unforgettable.

I would also like to acknowledge my family, especially my mother, for her enduring patience, understanding and encouragement without which this thesis would not have been possible. Very special thanks to Euan, for his support when it was most needed.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

There is a broad range of human diseases caused by the failure of a specific peptide or protein to adopt its native functional conformation state. These pathological conditions are referred as "protein misfolding (or protein conformational) diseases". The largest group of misfolding diseases is associated with the conversion of specific proteins from their soluble functional state ultimately into highly organised fibrilar aggregates. These structures are generally described as amyloid fibrils (Chiti and Dobson, 2006).

Aggregation is a general term that encompasses several types of interactions or characteristics. Aggregates of proteins may arise from several mechanisms and may be classified in numerous ways, including soluble/insoluble, covalent/non-covalent, reversible/irreversible and native/denatured. Just as there are many types of interactions that can lead to protein aggregation, there are many environmental factors that can lead to aggregation too. Solution conditions such as temperature, protein concentration, pH and ionic strength may affect the quantity and structure of aggregates observed. The present of certain ligands, including specific ions, may enhance aggregation. Stress to the protein such as freezing, exposure to air or interactions with metal surfaces may result in surface denaturation, which then leads to the formation of aggregates. Finally, mechanical stress may cause protein aggregation as well (Cromwell, Hilario, and Jacobson, 2006).

In this thesis, experimental studies have been carried out with a protein model: β-Lactoglobulin (β-Lg). This protein was chosen due to the fact it is inexpensive, its structure is already well defined and previous studies indicate that under certain conditions it forms fibrils, as it will be shown later.

So far, it has been established that many proteins form fibrils even under mild conditions, understanding the definition of fibrils as any type of aggregate with fibril-

like structure. However, many of them further aggregate until the formation of the high ordered structures known as amyloid fibrils. This type of aggregates are the final step in the degeneration of proteins causing neurodegenerative diseases; but many proteins can also lead to the formation of amyloid fibrils under the appropriate conditions, not only those related to neurodegenerative diseases (Chiti and Dobson, 2006). In the initial stage of amyloidogenesis, small fibrilar-like aggregates are formed – the so-called protofibrils. Several lines of evidence suggest that this is the highly neurotoxic stage in the formation of amyloid fibrils, being that the protofibrils are more toxic than the final and mature amyloid fibrils (Walsh and Selkoe, 2004). Most studies suggest a nucleation growth mechanism for the amyloid formation pathway started with a lag phase, followed by an exponential growth and finalised with an equilibrium phase (plateau). It would be in the lag period where the potentially toxic species would be formed. The study of this stage and the understanding of the parameters driven the formation of toxic protofibrils will be of high importance to find a successful way to stop neurodegeneration (Lasagna-Reeves, Glabe, and Kayed, 2011).

This research project is focused in the study of β-Lg aggregation behaviour, when it is subjected to solvent-induced denaturation, at neutral pH and at ambient temperature conditions for different concentrations of protein. The aim was to investigate the first stage in the aggregation mechanism before the final formation of amyloid fibrils as a model system to have a better insight in the amyloidogenesis process. Even though it was found that the system under study did not lead to the formation of amyloid fibrils within the experimental time of 10 days, very interesting results were found in the formation of β-Lg clusters immediately after adding the denaturant solvent for the fixed conditions. These results may open a new venue for further studies in the field.

1.1. PROTEINS

Proteins are essential parts of organisms and participate in every biological process within cells. Proteins make up approximately 5% of our average body mass and have various functions from structural material to functional, mainly depending on its structure.

1.1.1. Protein structure

Proteins are made of amino acids arranged in a linear chain and joined together by peptide bonds between the carboxyl and amino groups of adjacent amino acid residues. However, only 21 different L-α-amino acids build proteins. All α-amino acids possess common structural features, including a α-carbon to which an amino group, a carboxyl group, a hydrogen atom, and a variable side chain (which characterises the amino acid) are bonded. Most proteins fold into a unique three-dimensional structure. The shape into which the protein naturally folds is called the native state. Protein structure is usually divided into four distinct aspects by biochemists: primary, secondary, tertiary, and quaternary structure, which will be explained as follows (Price and Nairn, 2009, chapter 1).

1.1.1.1. Primary Structure

The primary structure of a protein is the amino acid sequence of the polypeptide chain, without regarding to spatial arrangements (apart from the configuration at the α-carbon atom). These amino acids are joined by peptide bonds. The π electrons from C=O bond delocalise across all three atoms C,O and N, giving the CO-N bond some double bond character (40%) due to that resonance. Consequently, all peptide bonds in protein structures are found to be almost planar. This rigidity of the peptide bond reduces the degrees of freedom of the polypeptide during folding (Buxbaum, 2007, pp.13-15). In Figure 1.1, a sequence of the primary structure is shown.

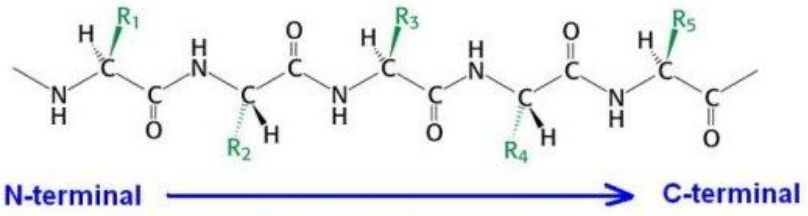


Figure 1.1 - Sketch of a protein's primary structure, created with the program ACD/ChemSketch.

1.1.1.2. Secondary structure

The secondary structure of polypeptide chain is the local spatial arrangement of its main-chain atoms without regard to the conformation of its side chains or to its relationship with other segments. Regular secondary structure occurs when all the rotational angles in the peptide bond (ϕ and ψ) are equal to each other (Prince and Nairn, 2009, chapter 1). There are three common secondary structures in proteins, namely alpha helices (α-helix), beta sheets (β-sheets) and beta turns (β-turns). That which cannot be classified as one of the standard three classes is usually grouped into a category called "random coil."

1.1.1.2.1. *α-helix*

α-helix is a right-handed coiled conformation resembling a spring in which every backbone N-H group donates a hydrogen bond to the backbone C=O group of the amino acid four residues earlier (exactly 3.6 amino-acids per rotation) (Petsko and Ringe, 2004, chapter 1). Figure 1.2 shows a molecular model of α-helix structure.

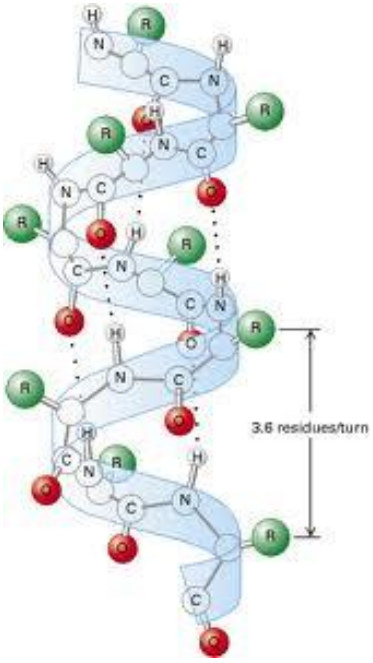


Figure 1.2 - Molecular model of the α-helix structure. Retrieved from the scientific blog of Mike Gene (author of "The Design Matrix") available online at http://designmatrix.wordpress.com - *The Rational Essence of Proteins and DNA*, posted on December 2009.

1.1.1.2.2. *β-sheets*

β-sheets consist in beta strands connected laterally by three or more hydrogen bonds, forming a generally twisted, pleated sheet. A β-strand is a stretch of amino acids, typically 5-10 amino acids long, whose peptide backbones are almost fully extended. Alternative β-strands can run in the same direction to give a parallel sheet or in opposite directions to give an antiparallel β-pleated sheet (Prince and Nairn, 2009, chapter 1). In Figure 1.3, both parallel and anti-parallel β-sheet can be seen.

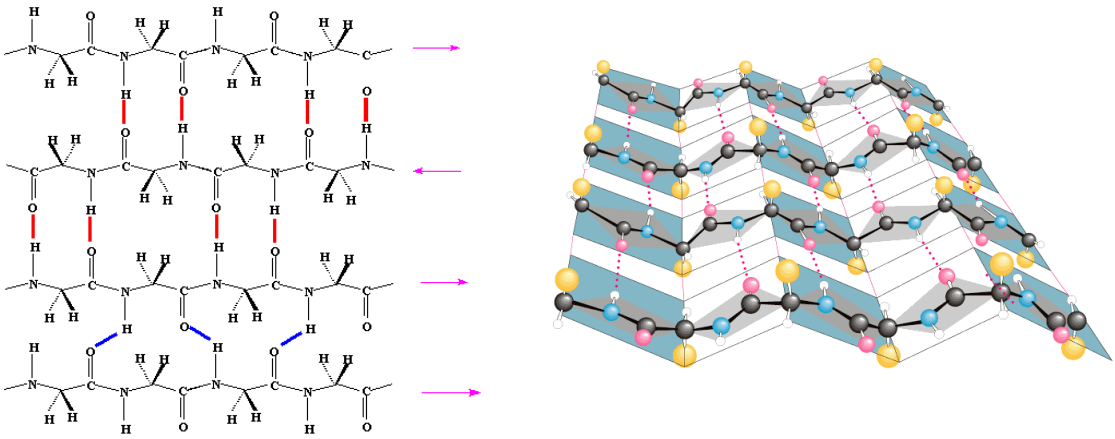


Figure 1.3 - β-sheet secondary structure. Left: linear representation created with the program ACD/ChemSketch; line between strands indicate hydrogen bond patterns; red lines indicate hydrogen bonding between antiparallel β-strands while blue lines indicate hydrogen bonding between parallel strands. Right: planar representation retrieved from the Lecture site of Rutgers University, New Jersey – *Chemical Kinetics and Mechanism* 160:563, autumn 2008, available online at http://andromeda.rutgers.edu

1.1.1.2.3. *β-turns*

β-Turns, more commonly known as turns, are the third of the three "classical" secondary structures that serves to reverse the direction of the polypeptide chain. They are located primarily on the protein surface and accordingly contain both polar and charged residues. The classic turn types (I, II and their backbone mirror images I' and II'') were first identified by Venkatachalam (1968). There is a third type of turn (III) which consists simply in a 3.10 helix (only one hydrogen bond between the first and the third amino acid) but it is not a common secondary structure in proteins. The different types of turns differ to each other depending on the preference that amino acids have to be positioned and the requirement for the formation of hydrogen bonds. Figure 1.4 shows β-turn type I and its mirror image (β-turn type I') as an example of this type of secondary structure.

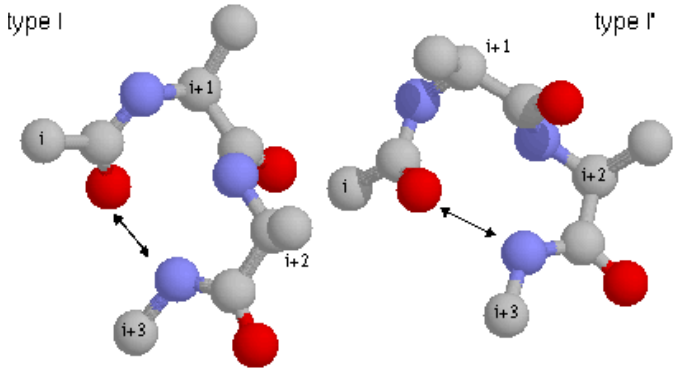


Figure 1.4 - Type I β-turn secondary structure. Retrieved from crystallography tutorials (available online at http://www.cryst.bbk.ac) of the former School of Crystallography currently part of the Department of Biological Sciences of Birkeck, University of London.

1.1.1.3. Tertiary structure

The elements of secondary structure are usually folded into a compact shape using a variety of loop and turns. Tertiary structure is mostly driven by the buried hydrophobic residues, but other interactions such as hydrogen bonding, ionic (or electrostatic) and van der Waals interactions and disulfide bonds can also stabilize the tertiary structure (Prince and Nairn, 2009, chapter 1). Proteins do not always have all these bonds/interactions present (Bull, 1998). The following paragraphs detail the types of interactions and their effect on protein stability.

(i) Electrostatic interactions

Electrostatic interactions are ionic interactions between molecules containing charged groups. Electrostatic interactions are long-range interactions occurring between positively charged amino groups and the negatively charged carboxyl group. Both of these groups are in abundance in proteins. They may also occur between the negatively charged amino acids (aspartic acid, glutamic acid) and the positively charged residues (arginine, lysine). Electrostatic interactions are pH and/or ionic strength dependent (Dill, 1990). The energy of an electrostatic interaction is typically in the rage of 42-84 kJ/mol.

(ii) Hydrogen bonds

Hydrogen bonds are attractive interactions between a partial positively charged hydrogen atom (the donor) and an electronegative atom such as oxygen, nitrogen, sulphur or fluorine (the acceptor) coming from another molecule or chemical group. The hydrogen atom must be covalently bonded to another electronegative atom. Therefore, hydrogen bonding refers to the sharing of a hydrogen atom that is attached to an electronegative atom (nitrogen, oxygen, or sulphur) with another electronegative atom or chemical group (oxygen of carboxyl group). The energy of a hydrogen bond depends on the group donor-acceptor so there is a continuous and wide energy scale covered by this type of interaction (from approximately 1 kJ/mol to nearly 167 kJ/mol). Due to this fact, hydrogen bonds are often classified in weak and strong depending on its energy (Desiraju and Steiner, 199, pp. 11-13).

(iii) Hydrophobic interactions

Non-polar side chains present in proteins have a low affinity for water. On folding into their native structures, proteins tend to remove these non-polar groups from the aqueous phase and form an inner hydrophobic core. The formation of this hydrophobic core causes the non-polar groups to adhere to each other and create hydrophobic interactions between these groups. Most of the studies in protein folding have led to the conclusion that hydrophobic interactions are the dominant force in protein folding (Dill, 1999). However, recent studies suggest that the packing of the protein in such a way that it optimises van der Waals contact (detailed below) and minimises cavities could be of equal importance energetically (Chen and Stites, 2001).

(iv) Van der Waals forces

This non-covalent stabilising force in protein structure is attributed to attractive and repulsive forces between molecules or different parts of the same molecule. This term encompass forces between a permanent and an induced dipole (Debye force) or between two instantaneously induced dipoles (London dispersion force). Van der Waals forces are non-specific, short-range forces (Roth, Neal, and Lenhoff, 1996).

When the polypeptide chain folds, van der Waals interactions come to a maximum due to the short distances between atoms, contributing to the strength of the hydrophobic effect, since non-polar atoms are favoured in this type of interaction. Even though the energy of such interaction is small if compared to other forces involving protein stability (4-13 kJ/ml (Bull, 1998)) it is known from crystallography studies that the core of the tertiary structure of the protein has a packing density similar to solids. In this environment, the van der Waals forces would sum up to a value which would have significant influence in protein stability (Lins and Brasseur, 1995).

(v) Disulphide bonds

Disulphide bonds are formed through the oxidation of two sulphydryl groups on cysteines amino acid residues. The presence of these covalent bonds in a folded protein provides an additional measure of stability to the native conformation, e.g., by stabilising the packing of a local cluster of hydrophobic residues (Wedemeyer, Welker, Narayan, and Scheraga, 2000) or by lowering the entropy of the unfolded state (Abkevich and Shahnovich, 2000). In multi-cysteines containing proteins, more than one arrangement of disulfide bonds may be possible. The energy of a disulphide bond is typically 126-420 kJ/mol (Bull, 1998).

In accordance with their degree of folding (tertiary structure), proteins are usually classified as globular or fibrous. Globular proteins are globe-like proteins in which hydrophobic tails groups are bounded towards the molecule's interior whereas the polar hydrophilic heads are bound outwards, allowing dipole-dipole interactions with the solvent. This explains the molecule's solubility: they form colloidal dispersions in water. Fibrilar proteins form long protein filaments, rod- or wire-like shapes. They are usually inert structural or storage proteins. These proteins are generally water-insoluble (difference with globular proteins) and are found as aggregates due to hydrophobic groups that stick out of the molecule. Special mention deserves the membrane proteins, attached to, or associated, to the lipid bilayer surrounding the cells. This type of proteins have different functions within the cell, from structural

proteins to transport, cell adhesion, enzymes and receptor proteins; therefore is difficult to place them within the general classification of globular and fibrous proteins. Membrane proteins are classified in integral and peripherical proteins, depending if they are permanently or temporarily bounding to the lipid bilayer respectively. One of the most important functions of membranes proteins is to allow the communication between the cytosol and the exterior of the cell (Lodish et al., 2000, section 3.4).

1.1.1.4. Quaternary structure

The quaternary structure is due to interactions between several individual monomeric chains that are called subunits. Proteins that are not functional as monomers usually have quaternary structure and they assembly into dimers, trimers and so forth (Petsko and Ringe, 2007, chapter 1). The same range of interactions as the tertiary structure stabilizes this structure.

1.1.2. Protein conformation

A protein can only be functional when it adopts a specific conformation, i.e. its native state. This conformation at certain conditions is given by the amino acid sequence and the thermodynamics of the various interactions aforementioned that stabilises the structure (Dill, 1990; Lins and Brasseur, 1995; Cooper, 1999). Interactions with the solvent are also of high importance in protein conformation. The highest stability is achieved when most hydrophobic groups are tightly packed in the interior of the structure (and the hydrophilic groups sticking out the protein surface). In this conformation, the free energy is minimal and so the protein is most stable. This conformation is called the native state of the protein as said above.

1.1.3. Protein denaturation

A delicate balance of non-covalent forces holds the native state (N) of a protein together: hydrophobic, ionic and van der Waals interactions and hydrogen bonds explained before. Upon exposure to certain concentrations of denaturing agents or to adverse environmental conditions (specified below), these non-covalent forces are

weakened, thereby causing the protein to unfold at least partially if not entirely. This transition to a less ordered conformation causes the inactivation (denaturation) of the protein. This unfolding event is usually reversible. In fact, the native state is often only marginally more stable than the fully unfolded state (U) (Creighton, 1990). Irreversible events are specific for individual proteins and individual cases of inactivation. Environmental conditions and denaturant agents that can lead to protein denaturation are detailed next.

(i) Change of pH

Denaturation and modification of proteins due to change of pH is a common, well-documented event (e.g. Rajendran and Prakash, 1988; Finsk, Calciano, Goto, Kurotsu, and Palleros, 1994). Furthermore, the degree of inactivation can range from minor conformational changes to irreversible inactivation depending on incubation conditions. The important factor at either pH extremes is that, once far away from a protein's isoelectric point, electrostatic interactions between like charges within the protein molecule result in a tendency to unfold.

(ii) Oxidation

Oxidation of amino acids, especially those with aromatic side-chains, can occur with a variety of oxidants: hydroxyl ($OH\cdot$) and superoxide (O_2^-) radicals, hydrogen peroxide (H_2O_2) and hypochlorite ion (OCl^-). Special interest is the oxidation of surfaced-exposed methionine and cysteine, which is reversible unlike the oxidation of the rest of amino acids. Oxidation of protein residues can lead to the cleavage of the polypeptide chain and to the subsequent formation of cross-linked protein aggregates (Stadtman and Levine, 2003).

(iii) Surfactants

Surfactants are interesting agents that can both induced denaturation or stabilise the native state of the protein depending on the nature of the surfactant. Binding of ionic surfactant to proteins lead to its gradual denaturation thanks to the great attraction between the hydrophobic tails of the surfactant and the non-polar residues of the

protein. This denaturation of globular proteins can be irreversible (Holmberg and Jönsson, 2002). A typical strong ionic denaturant surfactant is sodium dodecyl sulphate (SDS). On the other hand, non-ionic surfactants do not alter the tertiary structure of the protein. In fact, they often prevent interface-induced aggregation by limiting the adsorption of the protein to the surface (Chi, Krishnan, Randolph, and Carpenter, 2003).

(iv) Denaturant agents

(iv.i) Denaturants: Greene and Pace (1973) already reported that urea and guanidine hydrochloride (GdnHCl) were the most commonly used protein denaturants due to the fact that they generally lead to a greater level of unfolding. However, there is not generally accepted mechanism for their mode of action (Makhatadze and Privalov, 1992; Monera, Kay, and Robert, 1994). It is clear though that these denaturants diminish the hydrophobic interactions that stabilises the tertiary structure of the protein.

(iv.ii) High salt concentration: at high concentrations, salts exert effects on proteins that depend on both concentrations and nature of the salt. These effects can be either stabilizing or denaturing; they correlate with the Hofmeister lyotropic series:

$$(CH_3)_4N^+ > NH_4^+ > K^+ > Na^+ > Mg^{2+} > Ca^{2+} > Ba^{2+}$$

$$SO_4^{2-} > Cl^- > Br^- > NO_3^- > ClO_4^- > SCN^-$$

Anions and cations in the left part of the series stabilize protein. They are called kosmotropic ions and have a ‘salting out’ effect on the protein, increasing the strength of the hydrophobic interactions and therefore increasing its stability. In this way, the hydrophobic protein-protein interactions are stronger than solvent-protein interaction and the protein can precipitate. Salting out is also used as a technique for protein purification (Coen, Blanch, and Prausnitz, 1995). On the other hand, ions in the right hand side of the series are called chaotropic and they destabilised the protein by reducing the strength of hydrophobic interactions within the protein. This is known as the ‘salting in’ effect (Herberhold, Royer, and Winter, 2004).

(iv.iii) Chelating agents: metal ion-binding reagents such as EDTA can inactivate metalo-enzymes. On the other hand, proteins that do not require metal ion to be active are generally stabilised by chelating agents, because they remove possible harmful metal ions for the protein.

(iv.iv) Organic solvents: proteins function both in aqueous solutions and when suspended in non-polar solvents. However, total or partial inactivation of the protein is observed when mixtures of the two are formed upon addition of water-miscible organic solvents to aqueous protein solutions. This project is focused in this type of protein denaturation.

(v) Heavy metals and thiol reagents

Inactivation of proteins via heavy metal poisoning is well documented for different specific proteins (e.g. Giese, Levin, Bertch, and Kretzshmar, 2004). Heavy metal cations such as mercury (Hg^{2+}), cadmium (Cd^{2+}), copper (Cu^{2+}) and lead (Pb^{2+}) are known to react with proteins sulphydryl groups and denaturing the protein by cross-linking. On the other hand, this effect is used to advantage in some clinical treatments, such as the treatment of patients suffering from heavy metal poisoning (Smith and Wood, 1992, p. 61).

(vi) Temperature

Thermal denaturation is perhaps the most frequently encountered and most investigated. In fact, there is a vast amount of literature related to this type of protein denaturation very often together with another denaturant effect such as the presence of denaturant agents in solution or mechanical forces (see below). Inactivation due to heating is usually a two-step process: the reversible thermal unfolding of a protein exposes reactive groups and hydrophobic areas that can subsequently react to result in an irreversible inactivation. At high temperatures, the protein loses its regular non-covalent interactions, but upon return to ambient conditions, non-native, non-covalent interactions “trap” the protein molecule in a scrambled structure (Bull, 1998).

(vii) Mechanical forces

Proteins are susceptible to denaturation by mechanical forces such as pressure, shearing, shaking and ultrasound (e.g. Maa and Hsu, 1997; Hummer, Garde, García, Paulaitis and Pratt, 1998; Tian, Wan, Wang and Kang, 2004). Although in principle denaturation should be reversible, this is difficult to test because it is often accompanied by either aggregation or covalent reactions that cause irreversible loss of activity.

In several previous studies, it has been concluded that globular proteins need to unfold, at least partially, to aggregate into amyloid fibrils. A large body of experimental data supports this hypothesis. The identification of all conformational states during the amyloidogenic pathway is crucial to elucidate the whole process (Chiti and Dobson, 2006). Special interest has been focused in the so-called molten globule state, which is detailed below.

1.1.4. Unfolded proteins: Molten globule state

The attempt to characterize protein-folding pathways has been often based in the finding of folding intermediates. Secondary structure has a crucial importance in the final three-dimensional conformation of the protein; however, its formation is very fast and usually transient. This fact makes the study of intermediates species difficult. On the other hand, it is well known that certain proteins do present intermediate species in its unfolding pathway between the native (N) and total unfolded (U) states. This condensed intermediate is still a rather dynamic conformation, but sufficiently stable to be analysed in some detail. This intermediate is known with the name of ‘molten globule’ and it has contributed significantly to the understanding of protein folding (Pain, 1994, chap. 3). Wong and Tanford (1973) found that the equilibrium transition between the native and unfolded state of the protein not always followed a simple two state model, with solely these two species significantly populated. In this contest, it was proposed a model of folding which involves an “intermediate compact structure different from the native structure and whose formation is determined mainly by non-specific interactions of amino acid

residues with their environment (water and the hydrophobic core of a forming globule)" (Ptitsyn, Pain, Semisotnov, Zerovnik, and Razgulyaez, 1990). The characteristics of this folding intermediate have been established as:

(i) It is condensed in a globular form.
(ii) It contains native-like secondary structure.
(iii) It is stabilised mainly by non-specific hydrophobic interactions with non-polar residues concentrated in the interior.
(iv) It has a topology close to that of the native protein structure.
(v) It is "molten" in character.

The term "molten" implies that unlike the usual co-operative native-to-unfolded state transition, the lack of rigid tertiary structure would result in a non-co-operative phase transition between this intermediate and the fully unfolded state. Between the ordered state and the unordered one, a continuum of molten globule states could exit (Masson and Clery, 1996). One single protein could therefore, give rise to a multitude of intermediate molten globule states, with the potential of having different functional properties since the structure of the protein determines its functionality (Aouzelleg, 2004).

1.2. PROTEIN AGGREGATION AND NEURODEGENERATIVE DISEASES

One of the most important medical issues in our time are neurodegenerative diseases. Despite all the research throughout the last century, there is not yet a successful treatment for them. Neurodegeneration is a complex and multifaceted process leading to many chronic disease states. A conventional definition implies a progressive neuronal death, which usually affects a specific population of nerve cells, the vulnerability of which determines the clinical manifestations of a particular neurodegenerative disease. A classification can be made based on the principal neuropathological changes:

(i) Neurodegenerative disorders characterised by the presence of extensive protein deposits. These abnormal components accumulate in the brain causing a loss of neurons in an age-dependent manner leading to diseases such as Parkinson's disease (PD) (α-synuclein, Lewy bodies), Alzheimer's disease (AD) (Aβ-amyloid neurotic plaques and neurofibrilary tangles), and prion disease (amyloid plaque core surrounded by "petals" of sponge-like tissue, spongiosis).

(ii) Neurodegenerative disorders resulting from dysfunction/degeneration of motor neurons (Gaeta and Hider, 2005).

The first classification of neurodegenerative diseases is the most important one in the context of this work, since it is related to intracellular inclusions made largely of aggregates of certain proteins (Kokalj et al., 2004). Some examples are:

Alzheimer's diseases → tau, A42β peptide

Parkinson's disease → α-synuclein

Lewy bodies dementia → α-synuclein

Strong evidence that protein aggregation is central to neurodegenerative diseases has been shown from transgenic animal models that reproduce the general pathology, as well as post mortem studies carried out in patients who have suffered from the disease (Kokalj et al., 2004).

1.2.1. Protein aggregates

Polypeptide chains can adopt a multitude of conformational states and interconvert between them on a wide range of timescales. A general network of equilibrium is schematically illustrated in Figure 1.5 (Chiti and Dobson, 2006).

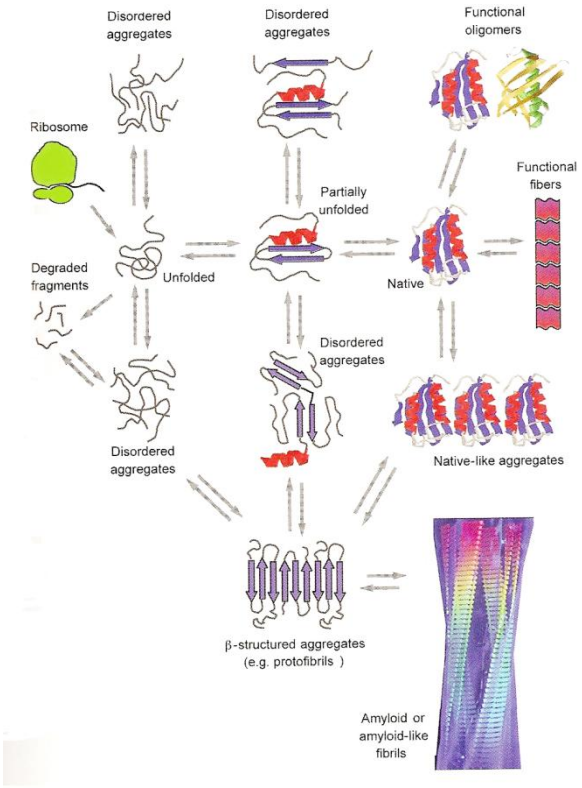


Figure 1.5 - Possible conformational changes for a polypeptide chain. Adapted from Chiti and Dobson (2006).

When a protein is synthesized in the ribosome (a sub cellular component formed by RNA and proteins where the assembly of amino acids to form the polypeptide chain takes place), it is initially unfolded. It can then populate a wide distribution of conformations, each of which contains little persistent structure (as in the case of natively unfolded proteins) or fold to a unique compact structure, often through one or more partly folded intermediates (Arnstein and Cox, 1992). The generic nature of the aggregation process has enabled extensive studies of the transition between soluble precursors and insoluble amyloid fibrils in vitro. Here the native conformation of globular proteins must be destabilised (e.g. upon addition of denaturants, low pH, high temperature) to allow protein aggregation to occur on a biochemically feasible timescale (Jahn and Radford, 2008). Chiti and Dobson (2006) mention three factors that seem to be of high importance in amyloid formation:

hydrophobicity of the side chain, electrostatic charge, and finally a low propensity of the protein to form α-helical structure and a high propensity to form β-sheet structure. Amyloid formation has been studied with a very wide range of proteins (e.g.: lyzozyme, Aβ peptide, α-La, β-Lg) but in all cases amyloid fibrils have approximately the same structure. It has been difficult to obtain high-resolution structures of amyloid fibrils, as these species are insoluble and non-crystalline. However, recent advances in experimental methods are starting to provide a more detailed picture of the amyloid architecture, such as higher resolution in atomic force microscopy (AFM) and electron microscopy (EM) techniques and computer algorithms (Adamcik et al., 2010; Goldschmidt, Teng, Riek, and Eisenberg, 2010).

There are different types of protein aggregates such as amorphous aggregates, worm-like aggregates, rod-like aggregates, filaments, and finally amyloid fibrils. In this study the onset of aggregated species (nucleation) preceding amyloid fibrilation is of primary interest. From a general point of view and based in initial studies, it can be seen that among all the different structures which appear during protein aggregation, only the ones eventually leading to amyloid fibrils (see section 1.2.2) can be linked with neurodegenerative diseases. Proteins able to form these kinds of fibrils are also called amyloidogenic proteins and the mechanism by which the native state is converted into amyloid fibrils is named as the amyloid pathway. As in any chemical mechanism, several different conformational and aggregation states are observed within this amyloid pathway. This has usually led to an inconsistent terminology over the literature in this field. For this thesis, nomenclature used by Regina Murphy (2007) is followed. The definitions for the different structures that appear during protein aggregation are defined as follows:

(i) MONOMERS

Monomers may have non-native or native secondary structure. The native secondary structure of monomeric amyloidogenic proteins can be of different types: it may be α-helical, β-sheet or a mix. Conversion to amyloid usually requires the dissociation

of the monomer; in these cases, the monomer is altered from the native fold and it is therefore unstable.

(ii) OLIGOMERS

Oligomers typically lack of a defined secondary structure and are globular in shape. The sizes and conformational characteristics of oligomers are variable.

(iii) PROTOFIBRILS

Protofibrils are linear aggregates that appear early in the amyloid pathway. Protofibrils, as oligomers aforementioned, lack the well-defined and high level ordered secondary structure present in the mature amyloid fibril. Protofibrils grow by monomer addition at the tip, the incorporation of small spherical oligomers, and/or by end-to-end coalescence. These structures are considered intermediates in the amyloid pathway. Recent studies also show that their toxicity is higher than that of the mature amyloid fibrils (Walsh and Selkoe, 2004; Haass and Selkoe, 2007; Necula, Kayed, Milton, and Glabe, 2007).

(iv) PROTOFILAMENTS and FILAMENTS

Protofilaments and filaments are linear aggregates characterized by a definite β-sheet structure. The distinction between protofilaments and filaments is not exactly defined; typically, filaments are elongated protofilaments. Protofilaments are distinguished from protofibrils in that the structural characteristics of the former are closer to those of fully mature fibrils. Amyloid fibrils represent the next step in the complexity of protein aggregates. Due to their relevance for this work, they will be detailed in a separate section.

(v) NONFIBRILAR AGGREGATES

Large non-fibrilar aggregates are observed in some preparations alongside amyloid fibrils. These non-fibrilar aggregates are considered occasionally to be amorphous but may have β-sheet content although they lack the obvious regular propagating structure of fibrils.

1.2.2. Amyloid fibrils

The term amyloid was first introduced by Rudolph Virchow in 1874 (Sipe and Cohen, 2000). Amyloid fibrils, usually referenced in literature as simply 'fibrils', have typically a diameter between 7 and 12 nm and a length of 1 μm or more, with a defined secondary structure characterised by a cross-beta core. They are usually formed by a few filaments (between 2 and 6) that may be wrapped or twisted around each other with regular helicity (Adamcik et al., 2010). Generally, fibrils are rather straight and linear, but occasionally have some branching (Murphy, 2007). Mature fibrils are the insoluble end-product of amyloidogenesis and they are known as amyloid fibrils.

Amyloid fibrils share some common characteristics in their morphology, tinctorial properties, and their core structure independently of the disease. However, it has been shown that amyloid fibril formation is not limited to proteins associated with diseases, which has lead to the viewpoint that amyloid fibril formation is a common property of globular proteins under appropriate conditions (Guijarro, Sunde, Jones, Campbell, and Dobson, 1998; Chiti et al., 1999; Yang, Zhang, Zhou, Chan, and Liang, 2006). However, this amyloid formation process is not fully understood yet despite the plentiful research in the field.

Amyloid is defined in terms of empirical observations from X-ray fibre diffraction, electron microscopy (EM) and specific chemical staining (Sipe and Cohen, 2000; Nilson, 2004; Whittaker, Spence, Grossman, Radford, and Moore, 2007). Amyloid fibrils have a predominantly cross β-sheet core (antiparallel) with a specific distance of 4.7-4.8 Å between β-strands and 10-13 Å between β-sheets reported from X-ray diffraction studies (Blake and Serpell, 1996). These two distances are perpendicular to one another, i.e. the β-strands are perpendicular to the fibre axis (the so-called cross β-core of the amyloid fibrils). Therefore, the dominant structure of amyloid fibrils is β-sheet even if the native state of the fibril-forming protein is predominantly α-helix. Amyloid fibrils are detected by experimental methods sensitive to secondary

structure such as circular dichroism (CD), attenuated total reflection – infrared spectroscopy (ATR-IR) or Raman spectroscopy. These fibrils bind the amyloid specific dye Congo Red (CR), changing the circular dichroism (CD) spectrum of the dye. Fibrils also bind the dye Thioflavin T (ThT) causing a large increase in its fluorescence intensity at wavelength of 480 nm when it is excited at 425 nm. In terms of morphology, it has been reported from transmission electron microscopy (TEM) and atomic force microscopy (AFM) studies that amyloid fibrils are long, straight (although occasionally twisted), unbranching fibrils with diameters in the region of a few nanometres. When looking for the exact values that characterized amyloid fibrils' diameter it is likely to find that these values vary depending on authors. For instance, while Hamada and Dobson (2002) reported a diameter for amyloid fibrils in the range 5-20 nm, Khare and Dokholyan (2007) found that it was in the range 7-10 nm. Common representation of this kind of protein aggregates is shown in Figure 1.6.

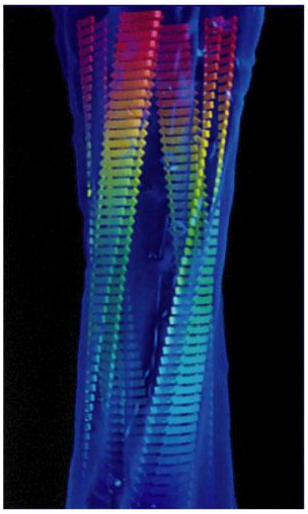


Figure 1.6 - Amyloid fibril. Adapted from the scientific website http://www.nature.com/horizon

At a macroscopic level, it is widely accepted nowadays that many proteins form amyloid fibrils via a nucleation-dependent growth mechanism, which will be described in the next section.

1.2.3. Nucleation Dependent Growth Mechanism

After plentiful research in this field, it is settled that amyloid fibril formation has the characteristics of a "nucleation growth" mechanism (Chiti and Dobson, 2006). This mechanism is characterised by a sigmoidal curve when plotting data from ThT-binding fluorescence intensity vs. time. In this curve, three different parts can be identified:

(i) A lag period

(ii) An exponential growth

(iii) Equilibrium phase

The lag phase, highly dependent on protein concentration, is the time required by the 'critical nucleus' to be formed. This critical nucleus is a thermodynamic nucleus, a very scarce species in the pathway but the only structure able to promote further growth leading to the formation of higher ordered aggregates. The formation of this nucleus is subjected to an activation energy barrier, characteristic of the nucleation growth mechanism (Figure 1.7).

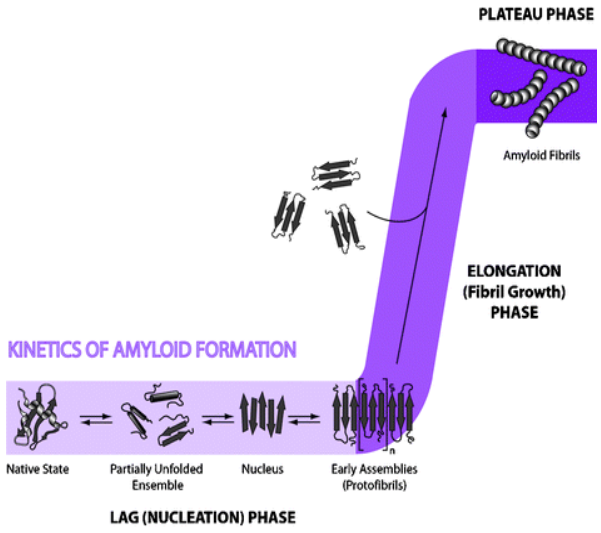


Figure 1.7 - Schematic representation of the amyloid fibril formation mechanism. Adapted from Chiti and Dobson (2006).

Once the nucleus is formed, the energy barrier is overcome and the fibril growth proceeds rapidly by further association of either monomers or oligomers with the nucleus. In this exponential growth stage, the rate of formation of polymers (amyloid fibrils) is set by the population of nuclei and the rate of elongation of the nuclei themselves, i.e. the rate of crossing the aforementioned energy barrier (Ferrone, 1999).

Finally, when there are no more monomers in solution, an equilibrium phase is reached (Pedersen and Otzen, 2008).

This mechanism is called nucleation-dependent because, in general, nucleation is the slowest phase, and so the rate limiting step for the reaction. As with many other processes dependent on nucleation, such as crystallization, the lag phase can be reduced by seeding, i.e. by adding preformed fibrilar species to a protein under amyloid fibril formation conditions or by certain experimental conditions dependent upon the system under study (Come, Fraser, and Lansbury, 1993; Bolder, Sagis, Venema, and van der Linden, 2007). This lag period can be completely abolished when seeding is enough to eliminate the need of forming a critical nucleus. The lack of a lag phase, however, does not necessarily means that a nucleated growth mechanism is not occurring. This may simply be because the time required for fibril growth is sufficiently slow relative to the nucleation process; in this scenario, nucleation is no longer the slowest step in the conversion of a soluble protein into its amyloid state, although this is not the general case (Chiti and Dobson, 2006).

The nucleation-growth mechanism is fairly accepted in the science community to describe the amyloidogenesis process; however, hypotheses proposed in this respect have not been fully corroborated yet at a more accurate (molecular) level. For instance, the existence of a conclusive causative link between aggregation and cytotoxicity in diseases remains unclear. In these terms, special interest is aimed in the research of the lag period because, although amyloid fibrils are not plentiful during this phase, several lines of evidence indicate that it is not the insoluble mature

amyloid fibrils but the soluble oligomers and/or protofibrils formed in the first steps the ones responsible for cytotoxicity (Walsh and Selkoe, 2004; Haass and Selkoe, 2007; Necula, Kayed, Milton, and Glabe, 2007).

At the molecular level, Khare and Dokholyan (2007) reported it in a recent paper that, "taken together, the results from several important investigations suggest three broad molecular frameworks that may underlie the conversion of normally soluble peptides and proteins into insoluble amyloid fibrils:

I. Edge-strand hydrogen bonding.
II. Domain swapping.
III. Self-association of amyloidogenic fragments."

However, none of these seems to be the exclusive path for the amyloid fibril formation since it appears to involve a combination of all three.

Several proteins have been used as models for the elucidation of the amyloid fibril formation process, such as α-Lactalbumin, β-Lactoglobulin, bovine serum albumin (BSA) (Bhattacharya, Jain, and Mukhopadhyay, 2011), prion protein (Come et al., 1993; Redecke et al., 2007), whey protein isolate (WPI) (Ikeda and Morris, 2002; Bolder et al., 2007; Bolder, Vasbinder, Sagis, and van der Linden, 2007), insulin (Bouchard, Zurdo, Nettleton, Bodson, and Robinson, 2002; Manno et al., 2007; Javid, Vogtt, Krywka, Tolan, and Winter, 2007), β-Synuclein (Yamin et al., 2005), β2-microglobulin (Gosal et al., 2005; Sasahara, Yagi, Sakai, Naike, and Goto, 2008), Aβ-amyloid (Bull, 1998), etc. Protein aggregation was carried out by subjecting the model protein in each case to different conditions of pH, temperature, salt concentration, pressure or other conditions, which make the protein unfold and aggregate afterwards.

This work is based on the behaviour of the whey protein β-Lactoglobulin (β-Lg) when it is subjected to denaturation conditions and its further development towards the formation of nucleated species. Previous studies in the field will be reported in the next section.

1.3. WHEY PROTEINS

As a by-product of cheese manufacture from cow's milk, three globular proteins can be isolated from whey: β-Lactoglobulin (~65%), α-Lactalbumin (~25%) and serum albumin (~8%) (Schokker, Singh, and Creamer, 2000). They are called whey proteins and are soluble in their native forms, independent of pH. In this thesis, only β-Lactoglobulin was used as a model protein for the research of protein aggregation process. An overview of the bovine serum albumin and α-Lactalbumin as well as more detailed information about β-Lactoglobulin is given below.

1.3.1. Bovine serum albumin

Bovine serum albumin (BSA) consists of 583 residues and has a molecular weight of 66430 Da, with an isoelectric point (pI) of 4.7 at 25ºC in water. It has numerous biochemical applications including ELISAs, inmunoblots and inmunohistochemistry. It is also used as a nutrient in cell and microbial culture. BSA is characterised by a low content of tryptophan and methionine and a high content of cystine, lysine, arginine and both aspartic and glutamic acid. The secondary structure of BSA is predominantly α-helix (67%) with the remaining polypeptide occurring in turns and extended regions with no β-sheets, as has been shown by X-ray crystallographic analysis (Carter et al., 1989). This secondary structure is divided in three domains (I, II and III) with specific binding to ligands such as chloroform (binding preferably to domain I), bilirubin (binding to domain II) and diazepam (binding to domain III) (Tayyab, Sharma, and Khan, 2000). A three dimensional image of this protein can be seen in Figure 1.8. The most outstanding property of BSA is its ability to bind other molecules, both in a reversible or irreversible way. For instance, it is the principal carrier of fatty acids.

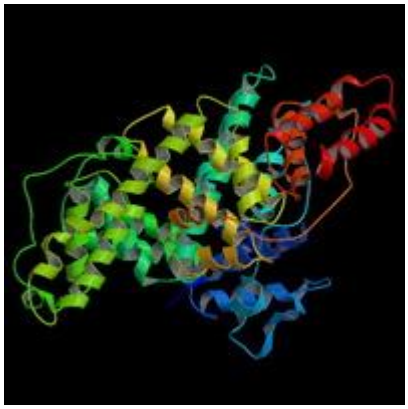


Figure 1.8 - Tertiary structure of bovine serum albumin (image retrieved from Bio-world company website: http://www.bio-world.com/productinfo/2_36_259/7632/BSA-Bovine-serum-albumin-mg-mL.html).

1.3.2. Alpha-Lactalbumin (α-La)

α-Lactalbumin (α-La) is a relative small milk protein (molecular weight of 14200 Da), with an isoelectric point of 4-5 (acidic protein). It is characterised by a Ca^{2+} binding site (Permyakov and Berliner, 2000). A general representation of this protein is shown in Figure 1.9. It has an important function in mammary secretor cells: it is one of the two components of lactose synthesis, which catalyzes the final step in lactose biosynthesis in the lactating mammary gland. α-La is frequently used as a simple model Ca^{2+} binding protein. Many researchers have selected this protein to be used in their studies because at acidic pH and in the apo-state (not Ca^{2+} loaded) and at elevated temperatures, α-La adopts the classic 'molten globule' conformation aforementioned (section 1.1.4).

Most of α-La proteins (e.g. human, guinea pig, bovine, goat, camel, equine and rabbit proteins) are formed by 123 amino acids residues. Acharya, Ren, Stuart, Philips, and Fenna (1991) reported that the tertiary structure of α-La is very similar to that of lyzozyme revealed by X-ray crystallography studies. Native α-La consists

in two domains: a large α-helical domain (~45%) and a small β-sheet domain (~9%) which are connected by a calcium binding loop (Permyakov and Berliner, 2000).

One of the most important features of α-La is its ability to bind metal cations. Additionally to its strong calcium site, α-La also has several Zn^{2+} binding sites (Permyakov et al., 1991) and can bind other physiologically significant cations such as Mg^{2+}, Mn^{2+}, Na^{+} and K^{+} which can compete with Ca^{2+} for the same binding site (Permyakov, Morozova, and Burstein, 1985). When Ca^{2+} binds α-La, it increases the stability of the protein when it is subjected to high temperatures or pressures (Veprintsev et al., 1997; Dzwolak, Kato, Shimizu, and Taniguchi, 1999). However, when α-La is Ca^{2+}-loaded and Zn^{2+} binds the protein, thermal stability decreases. Other cations binding α-La (e.g. Mg^{2+}, Na^{+} and K^{+}) increase the stability of the protein against denaturant agents such as urea or guanidine hydrochloride (GdnHCl) (Permyakov and Berliner, 2000). However, the degree of stability depends strongly on the concentration of the metal ions.

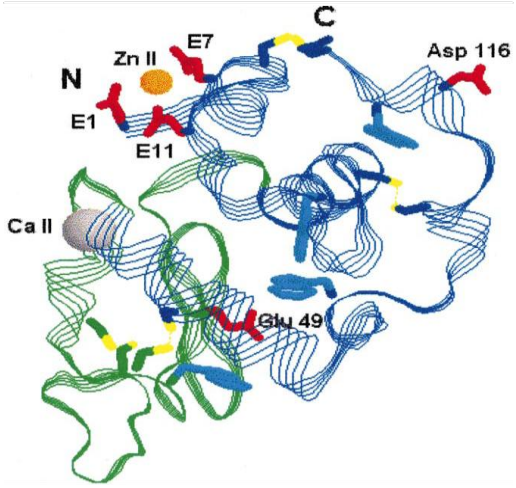


Figure 1.9 - X-ray α-La structure derived from native buffalo and recombinant bovine protein (taken from Brookhaven Protein Data Bank) (Permyakov and Berliner, 2000).

1.3.3. Beta-Lactoglobulin (β-Lg)

Bovine β-Lactoglobulin (β-Lg) is a key member of the lipocalin super family since it possesses the lipocalin motif (Euston, Ur-Rehman, and Costello, 2007). This protein

has been widely studied for several reasons; key factors have been the abundance and ease preparation of the protein (Dufour and Haertlé, 1990). β-Lg has great commercial importance to the milk industry, which is concerned with improving the processing of whole milk and milk derivates as well as searching for new uses for the protein. One of the most important processes in milk manipulation is heat treatment. At high temperature, the behaviour of β-Lg is of particular interest, as the aggregates it forms are supposed to be the responsible for fouling and less efficiency in heat exchangers (Majhi et al., 2006). β-Lg is a robust protein, easily isolated, which makes it a perfect model protein for different studies and for the testing of both biochemical and physical techniques.

Bovine β-Lg is defined as a globular whey protein with a molecular mass of 18400 Da and radius of gyration of about 1.3 nm for the monomer form (Pessen, Kumosinski and Timasheff, 1973) (for definition of radius of gyration see section 2.2.5). The monomer consists of 162 amino acids residues with two disulfide bridges (Cys^{60}-Cys^{160}, Cys^{106}-Cys^{119}) and one free cysteine (Cys^{121}) (Sawyer and Kontopidis, 2000; Yang, Power, Clark, Dunker, and Swanson, 2001). Each monomer contains two tryptophans in positions 19 and 61. The isoelectric point is ~5.2 and its secondary structure is composed of 15% α-helix, 40% β-sheet and 45% of turns and random coil together (Kabsh and Sander, 1983). Depending on pH and salt concentration, β-Lg at room temperature has been reported to exist as a monomer or a dimer (stabilised by hydrogen bonds). The monomer form predominates below pH 3 and above pH 9, but coexists in equilibrium with dimers in the pH ranges of 2.0-3.7 and 5.2-9.0 (Majhi et al., 2006), even though the dimer form dominates at neutral pH (Gimel, Durand, and Nicolai, 1993). The dimeric protein can be presented as two connected spheres of 1.7 nm radius, with a centre-to-centre distance of approximately 3.3 nm, given a gyration radius of approximately 2 nm for the β-Lg dimer (Gimel et al., 1993; Panick, Malessa, and Winter, 1999; Verheul, Pedersen, Roefs, and Fruif, 1999; Loupiac, Bonetti, Pin, and Calmettes, 2006). However, the structures at the different pH values possess the same basic topology, having nine

antiparallel β-strands and one short and one long α-helix at the carboxyl terminus. A significant feature of the protein is that the β-sheets are arranged in such a way as to give rise to a central cavity (a β-barrel or calyx) where small hydrophobic molecules may be bound (Euston et al., 2007). Figure 1.10 shows a representation of β-Lg monomer.

Figure 1.10 – Three-dimensional diagram of the structure of the milk protein β-Lg, showing the calyx-shaped container formed by a pair of β-sheets, providing a pocket for the protein to carry a lipid molecule (retrieved from the web site of The Astbury Centre for Structural Molecular Biology, University of Leeds: http://www.astbury.leeds.ac.uk/history/history2.htm).

The biological function of β-Lg is still unclear (Brownlow et al., 1997) but the protein has the ability to bind to extended hydrophobic compounds such as retinol, fatty acids or palmitate (Hamada and Dobson, 2002). Bovine β-Lg in its natural state consists of nine variants, named with the alphabet letters from A to J (Qin et al., 1998). The two main genetic variants, "A" and "B" only differ in the amino acids in position 64 and 118 of the peptide chain: asparagine and valine in variant "A", and glycine and alanine in variant B (Dong et al., 1996). Self-association is more pronounced for β-Lg-A than for β-Lg-B (Majhi et al., 2006).

Denaturation conditions for β-Lg depend on several parameters such as β-Lg concentration, pH, temperature, pressure, etc. Many studies have been carried out using β-Lg in the last 60 years based on the lost of stability of the protein which leads to the formation of β-Lg aggregates.

β-Lg has been widely used for fibrilation studies since it is a relative inexpensive and harmless protein. This protein can lead to the formation of fibrils by temperature or solvent-induced denaturation. However, the former has been more widely studied than the second has. There is a vast amount of literature concerning β-Lg denaturation and aggregation due to the many different variables that can be experimentally manipulated.

High temperatures and acidic pH have been the conditions most widely used to achieve β–Lg aggregation. Structure of aggregates depends on incubation time at high temperature and other ambient conditions, such as ionic strength or protein concentration. Aymard, Nicolai, and Durand (1999) studied β-Lg aggregates formed after heat-induced denaturation (80ºC) at pH 2 for different protein concentrations at different ionic strength finding that aggregates at pH 2 are essentially wormlike-chains with a persistence length that is strongly dependent upon the ionic strength. Additionally, they found that at this pH there is no evidence for a two-step aggregation process as was observed at pH 7 where denser aggregates were formed (Kella and Kinsella, 1988). None of these studies reported the formation of aggregates meeting the structural characteristics of mature amyloid fibrils. Long linear aggregates have been found at pH 2 (and occasionally at pH 2.5) following prolonged heating (80ºC over 6 hours) by Kavangh, Clark, and Ross-Murphy (2000), where the presence of salt led to the formation of much shorter aggregates. However, in this case, it was reported that no difference in aggregate structure could be found between acid and neutral pH conditions in contrast with the previous paper (Kella and Kinsella, 1988). Arnaudov, de Viries, Ippek, and van Mierlo (2003) found helical fibrils after prolonged heating at high temperature (80ºC) at pH 2 and at low

ionic strength. They reported that the fibril formation of β-Lg appears to be a "multistep process." In the end, two different products were found after several competing reactions: long linear aggregates and some low molecular weight "dead end" species. A theoretical modelling of the kinetics of fibrilar aggregation of β-Lg at pH 2 was also developed by this group, as well as the influence of ionic strength in β-Lg fibril aggregation kinetics and a report about a critical concentration needed for the formation of β-Lg fibrils: above 2.5 % w/w of β-Lg fibrils are formed, whereas under this concentration they are not (Arnaudov and de Viries, 2006; Arnaudov and de Viries, 2007). Gosal, Clark, and Ross-Murphy (2004a) found that β-Lg formed long and smooth aggregates under similar conditions. An interesting rheology study of heat-induced β-Lg gels was also carried out by this group (Gosal, Clark and Ross-Murphy, 2004c). Majhi et al. (2006) developed another kinetic study at low ionic strength. Maximum aggregation rates were observed at pH range of 4.3-4.8, i.e. below the isoelectric point (pI) of the protein, where it is positively charged. The rate of aggregation strongly increased when ionic strength decreases. They explained the aggregation of β-Lg as a two-step process in which a dimer, the initial reactant, rapidly forms intermediate aggregates that further associate to form longer clusters.

After reviewing the literature on temperature induced β-Lg denaturation, it is evident that published results contain an element of confusion due to terminology since the term amyloid fibril is sometimes substituted by the term 'fibrils' and in most cases β-Lg aggregates cannot be defined as mature and highly ordered amyloid fibrils. In fact, it is difficult to confirm whether the formation of β-Lg linear aggregates reported in published papers represent amyloid fibrils, as some researches did not apply relevant techniques such as ThT-binding fluorescence, Congo Red binding or structural analysis such as X-ray diffraction, for the analysis of their particular system. If amyloid fibrils are understood with the characteristics reported before in this thesis, the term is accurately used by Bromley, Krebs, and Donald (2005). In this work is reported that β-Lg aggregates with amyloid fibrils features were found at pH 2.5 by incubation at 70ºC in the absence of NaCl in solutions where protein

concentration was 4% w/w (40 mg/ml). Additionally, of significant interest are the results obtained by Krebs, Devlin, and Donald (2009) where powder X-ray diffraction pattern of the protein at pH 5.3 and concentration 3% w/w (30 mg/ml) after being heated at high temperature (80ºC) during a short period of time (10 minutes), revealed the presence of two peaks at ~4.7Å and ~10.6 Å, characteristic of the cross-β motif present in mature amyloid fibrils.

As well as the widely studied heat-induced denaturation, other methods have been investigated to achieve β-Lg amyloid fibrils formation. They include urea and guanidine hydrochloride (GdnHCl) which have been used as denaturants to unfold the protein as aforementioned. β-Lg amyloid fibrils were found by incubating the protein in the presence of urea at pH 7 and 37ºC for 10-30 days (Hamada and Dobson, 2002). The concentration of urea at which amyloid fibril formation is fastest was found to be 5M. Guanidine hydrochloride (GdnHCl) was also used with that objective although for the formation of amyloid fibrils in the egg white lyzozyme at pH 6.3. This denaturant was found to be more useful than urea for the same aim (Vernaglia, Huang, and Clark, 2004). Guanidine hydrochloride has also been used with β-Lg but solely for the investigation of β-Lg denaturation and not with the aim of β-Lg aggregation studies (D'Alfonse, Collini, and Baldini, 2002).

Organic solvents such as ethanol, methanol, trifluoroethanol (TFE) as well as others have been used in β-Lg aggregation studies (Dufour and Haertlé, 1990; Dufour, Genot, and Haertlé, 1994; Dufour, Robert, Bertrand, and Haertlé, 1994; Uversky, Narizhneva, Kirschstein, Winter, and Löber, 1997; Dufour, Robert, Renard, and Llamas 1998; Renard, Lefebvre, and Robert, 1999; Renard, Robert, Garnier, Dufour, and Lefebvre 2000; Robert, Lavenant, and Renard, 2002; Kanjilal et al., 2003; Gosal et al., 2004a; Mousavi, Chobert, Bordbar, and Haertlé, 2008; Krebs et al., 2009). It was found that, at low protein concentration (in the range of micro molar) and at a neutral pH, β-Lg undergoes a change on its secondary structure when it is dissolved in buffer-ethanol mixtures. The degree of this change depends on the concentration

of ethanol (Dufour and Haertlé, 1990). Several studies over the last years using many different techniques such as retinol-binding florescence, CD, FTIR and ultrasonic studies corroborate that the addition of ethanol (or other organic solvents such as iso-propanol, TFE or methanol) to β-Lg leads to conversion of β-sheets into α-helix structures (Dufour and Haertlé, 1990; Dufour et al., 1994; Uversky et al., 1997; Kanjilal et al., 2003). The extent of this change has been reported to vary but it is clear that α-helix structure only appears when the concentration of the alcohol is relatively high (50% v/v) and by changing the compressibility of the protein (Kanjilal et al., 2003). Minimal changes in the secondary structure have been found to occur up to 20% (v/v) of ethanol while at higher concentrations like 50% (v/v) there have been both major and drastic changes in the secondary structure of the protein. The retinol-binding fluorescence studies at a 20% (v/v) ethanol concentration revealed the presence of a less compact native-like intermediate structure, the 'molten globule' state (Dufour and Haertlé, 1990; Dufour et al., 1994) explained previously in this work. This ethanol-induced transition from β-sheet to α-helix was found to be reversible (Dufour et al. 1990). Very similar conclusions were obtained by Uversky et al. (1997) working with several alcohols at acidic pH. Dependent on pH, β-Lg can also undergo irreversible gelation in the presence of alcohol (Renard et al., 1999; Dufour et al., 1998).

In recent years the objective has been to identify the structure of β-Lg aggregates both from a more quantitative point of view (deconvolution of IR and CD spectra) (Robert et al., 2002; Mousavi et al., 2008) and from a more "visual" approach by using techniques such as atomic force microscopy (AFM) and electron microscopy (EM). It has been found that β-Lg aggregates structure depends highly in the physico-chemical conditions of the protein solutions. Aggregates formed from solvent-induced denaturation at acidic and neutral pH seems to be worm-like with a "string of beads" appearance (Gosal et al., 2004a). Furthermore, these aggregates appear to have different structure from those obtained by temperature-induced denaturation. For instance, it was found that aggregates formed by incubation of β-

Lg in various alcohol-water mixtures at pH 2 and in TFE-water solutions at pH 7 were wormlike with a "string-of-beads" appearance. Longer, smoother and apparently stiffer fibrils formed in heating aqueous β-Lg solutions at pH 2 and low ionic strength (Gosal et al., 2004a; Gosal, Clark, and Ross-Murphy, 2004b).

In conclusion, this project shall present a study of aggregation of β-Lg subjected to the influence of an organic solvent with a low dielectric constant (ε ~ 24.3 at 25°C) - ethanol. The physico-chemical conditions (e.g. temperature, concentration of salt, pH) vary from the previously mentioned studies in an attempt to study the effect of a different combination of conditions in the formation of nucleated species of β-Lg, when the protein is present at several concentrations in solution. Both experimental techniques and protocols applied in this study will be detailed in the next chapter.

2. MATERIALS AND METHODS

2.1. MATERIALS

The following chemicals and reagents were required for the experimental studies.

2.1.1. Proteins

β-Lactoglobulin from bovine milk, 3 times crystallised and lyophilised, containing β-Lactoglobulin A and B (90% purity) desiccated and stored at 4ºC was purchased from Sigma Aldrich (catalogue number L0130-5G). The protein was no further purified.

2.1.2. Reagents

The following reagents were required for the experimental studies.

(i) Hydrochloric acid

A solution of 0.1 M of hydrochloric acid (HCl) was prepared from commercial stock: 32% w/w (10.2 M), purchased from Fisher Chemicals, UK (catalogue number H/1100/PB15). 2.45 ml of HCl were transferred to a 250 ml flask topped up with distilled water. This solution was used to adjust the pH of the samples to 7 when required.

(ii) Sodium hydroxide

A solution of 0.1 M of sodium hydroxide (NaOH) was prepared from commercial stock purchased from Fisher Chemicals, UK (catalogue number BPE359-500). 0.4 grams of NaOH were weight and dissolved in 100 ml of distilled water. This solution was used to adjust pH of samples to 7 when required.

(iii) Sodium phosphate buffer

Sodium phosphate buffer 0.2 M and pH 7 was prepared from two different solutions of salts: sodium dihydrogen orthophosphate dihydrate ($NaH_2PO_4 \bullet 2H_2O$) and di-sodium hydrogen orthophosphate anhydrous (Na_2HPO_4), both purchased from Fisher Chemicals (catalogue numbers S381-100 and S375-500 respectively). 3.123 g of the former and 2.840 g of the latest were weighed and dissolved in 100 ml of distilled water to prepare two different solutions of concentration 0.2 M and pH 7. 19.5 ml of

$NaH_2PO_4 \cdot 2H_2O$ solution (0.2 M) and 30.5 ml of Na_2HPO_4 solution (0.2 M) were mixed in a 100 ml flask and topped up with distillate water giving 100 ml of final phosphate buffer 0.2 M and pH 7. Phosphate buffers 10 mM and 20 mM were prepared from buffer stock 0.2 M by dilution. pH of less concentrated buffers was adjusted to 7 by adding HCl 0.1 M or NaOH 0.1M when required. pH was measured using a pH 212 HANNA Microprocessor pH meter. All buffers were subjected to sterilization and kept in the fridge for no longer than two weeks to avoid biological contamination.

(iv) Thioflavin T

Thioflavin T (ThT) was used for fluorescence spectroscopy measurements as the dye binds cross β-sheet like structures. Thioflavin T was purchased from Sigma Aldrich (catalogue number: T3516-5G). A stock solution of ThT 100 μM was prepared by weighing the corresponding quantity of ThT powder and dissolved in 50 ml of phosphate buffer 20 mM. This solution was stored in an opaque plastic bottle in the freezer. When fluorescence measurements were carried out, a fresh ThT solution of 20 μM was prepared by dilution of the stock with phosphate buffer 20 mM and mixed with β-Lg samples in the fluorescence quartz-cell to achieve a protein solution to ThT solution ratio of 1:9 by volume. This volume ratio corresponds to different protein to ThT molar ratios depending on the concentration of the sample under study. The final concentration of ThT in the quartz cell was 18 μM in all cases.

(v) Ethanol

Ethanol absolute purchased from Sigma Aldrich was filtered by a PTFE syringe filter 0.45 μm pore-size before being added to the samples when required. No further treatment was done with the denaturant solvent.

2.2. METHODS

2.2.1. UV absorption spectroscopy

UV spectroscopy was used to calculate the concentration of protein samples. The corresponding quantity of protein powder was weighed for the preparation of samples with the desired protein concentration by weight, although the accurate

concentration of β-Lg by UV absorbance spectroscopy was required for CD spectra correction (see section 2.2.3). I note that in this work the concentration of protein will always refers to its concentration by weight, unless otherwise stated.

2.2.1.1. Theory

Proteins in solution absorb ultraviolet light with absorbance maxima at 280 and 200 nm. Amino acids with aromatic rings such as tryptophan, tyrosine, and phenylalanine (although the latest has a small absorption in this region (Luthi-Peng and Puhan, 1999)) are the primary reason for the absorbance peak at 280 nm. Peptide bonds are primarily responsible for the peak at 200 nm. Secondary, tertiary, and quaternary structures all affect absorbance. Therefore, factors such as pH or ionic strength affecting protein structure can also alter the absorbance spectrum. Protein concentration was calculated using Equation 1:

$$C = \frac{A_{280} - A_{310}}{\varepsilon \cdot (cell - path\,length)} \qquad \text{Equation 1}$$

Where:

- C is the protein concentration (mg/ml).
- A_{280} and A_{310} are the absorbance values at 280 and 310 nm respectively.
- ξ is the absorption coefficient that depends on the type of protein. For β-Lg it has a value of 0.96 $ml \cdot mg^{-1} \cdot cm^{-1}$ (Bhattacharjee and Das, 2000).
- Cell-pathlength is measured in centimetres (cm).

2.2.1.2. Sample preparation and spectral acquisition

UV-absorbance spectra were recorded using a Jasco spectrophotometer (V-550) (University of Glasgow) and a UV-VIS SHIMADZU spectrophotometer UV-mini-1240 (University of Strathclyde).

Quartz cells of 0.5 cm path length were used in Jasco V-550 instrument. Absorbance spectra were recorded between 340 nm and 220 nm. Individual absorbance values at 280 nm and 310 nm were read to calculate protein concentration using Equation 1.

Disposable cuvettes purchased from Fisher Scientific (catalogue number FB55923) of 1 cm path length were used in SHIMADZU apparatus. In this case, no absorbance spectra were recorded, but the individual values at 280 nm and 310 nm using the photometric function of the instrument. In both cases, cells were filled with solutions of 1 mg/ml of protein. When protein solutions had a higher concentration, they were diluted to 1 mg/ml before measuring to assure a good absorbance response. Blanks were also prepared and measured as a control.

2.2.2. Fluorescence Spectroscopy

Structural kinetics arrangements of β-Lg were monitored by fluorescence spectroscopy using Thioflavin T (ThT) – binding fluorescence. This dye was reported to bind amyloid fibrils in solution specifically, and has been widely used to determine the presence of amyloid fibrils in vitro. However, the specificity of ThT to bind amyloid fibrils must not be understood in a strict way, since information about ThT being able to bind other organic molecules such as nucleic acids is available in literature (Khurana et al., 2005). Detailed information about ThT and its binding to amyloid fibrils will be shown in the following section.

2.2.2.1. Theory

When a molecule is excited with electromagnetic energy, electrons can jump up to a higher level of energy by absorbing a photon (where they are unstable). When this electron returns to its own level of energy, it is accompanied by the emission of a photon, but this re-emitted radiation has less energy than the radiation needed to excite the electron. This phenomenon is known as the Stokes shift. This re-emitted radiation can be seen as a spectrum for different wavelengths, independently of the excitation wavelength (Lakowicz, 2006, pp 1-7).

Substances that have enough fluorescence signals to be useful are those with conjugated double bonds, especially aromatic species. In proteins, only aromatic amino acids have sufficient fluorescence intensity to be measured directly in solution. In this work, intrinsic fluorescence of β-Lg was not measured, but the

emission spectra of the dye ThT when it is excited at 425 nm in the presence of β-Lg. ThT is a cationic benzothiazole with a hydrophobic end with a dimethylamino group attached to a phenyl group (structure shown in Figure 2.1).

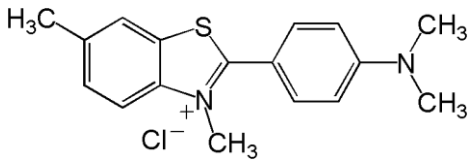


Figure 2.1 - Molecular structure of Thioflavin – T created with the program ACD/ChemSketch.

This combination of polar and hydrophobic regions creates a possibility for Thioflavin T molecules to form micelles in solution with hydrophobic interior and the positively charged N toward the solvent. This dye has been widely used to diagnose amyloid fibrils formation, both *ex vivo* and *in vitro* (Naiki., Higuchi, Hosokawa, and Takeda, 1989; LeVine, 1999). It undergoes a strong increase in fluorescence quantum yield (around 480 nm when exited at 425 nm) upon binding to amyloid fibrils. Even though the exact way ThT binds amyloid fibrils is not clear yet, it has been suggested that ThT binds along the β-strands in the amyloid fibrils. β-sheets are ordered into β-strands within a fibril, forming channels running along the fibril. For a ThT molecule to bind parallel to the long axis of the amyloid fibril in a stable way, it is likely that ThT inserts itself into these channels in the cross β-core of the amyloid fibril. This is the most likely manner for ThT to bind amyloid fibrils (Krebs, Bromley, and Donald, 2005). It is worth to note here that since all β-sheet contains binding channels, even the native state of the protein may bind ThT molecules without being in a fibrilar architecture if its secondary structure has a major content of β-sheets (as it is the case for β-Lg). A small increase in fluorescence has indeed been observed when ThT was added to samples of β-Lg in its native state. Since aggregation process may involve the formation of several types of aggregates, which can contain stable and/or transitional cross β-sheets like structures increasing the fluorescence signal of ThT at 480 nm (but without being the high ordered

structures present in amyloid fibrils), results from this technique alone cannot be conclusive to assure the present of amyloid fibrils in solution without the support of other techniques such as AFM or X-ray diffraction. However, ThT-binding fluorescence is a suitable technique to monitor the structural arrangements taking place during the aggregation process, by analysing how the ThT-emission peak in the presence of the target protein changes with incubation time. When these results are plotted in a scatter graph, a sigmoidal like curve is obtained when the process follows a nucleation-dependent growth mechanism (see section 1.2.3). This data can be fitted later to calculate both lag time and growth rate of the mechanism studied.

2.2.2.2. Fluorescence instrument

The spectrofluorometer is similar to a spectrophotometer. Beam coming from the light source pass through a monochromator that selects the excitation wavelength (425 nm for this project). This monochromatic beam passes through the sample and a second monochromator, situated in an angle of 90 degrees respect to the first one, selects the emission wavelength of the radiation coming from the sample. Afterwards, the fluorescence radiation is quantified by a photomultiplier. A Jasco-6500 instrument was used to obtain ThT-binding fluorescence data. A diagram of the optical system of this model of spectrofluorometer is shown in Figure 2.2. The light from the light source L (150W xenon lamp) is focused onto the entrance slit S1 of the excitation (Ex) monochromator by the ellipsoidal mirror M1 and spherical mirror M0. The light from the slit is dispersed by the diffraction grating G1 and monochromatic light is taken out by the exit slit. Part of the monochromator light is led to the monitoring photometric photomultiplier tube PM1 by the beam splitter BS. The rest of the monochromatic light that has been transmitted by the beam splitter is led to the sample chamber by the plane mirror M3 and ellipsoidal mirror M4, where it is focused onto the entrance slit S2 of the emission (Em) monochromator by the ellipsoidal mirror M5 and plane mirrors M6 and M7. Monochromator beam is taken out from the light dispersed by the diffraction grating G2 of the emission monochromator after going through the exit slit and is led to the photometric photomultiplier tube PM2 by the spherical mirror M8.

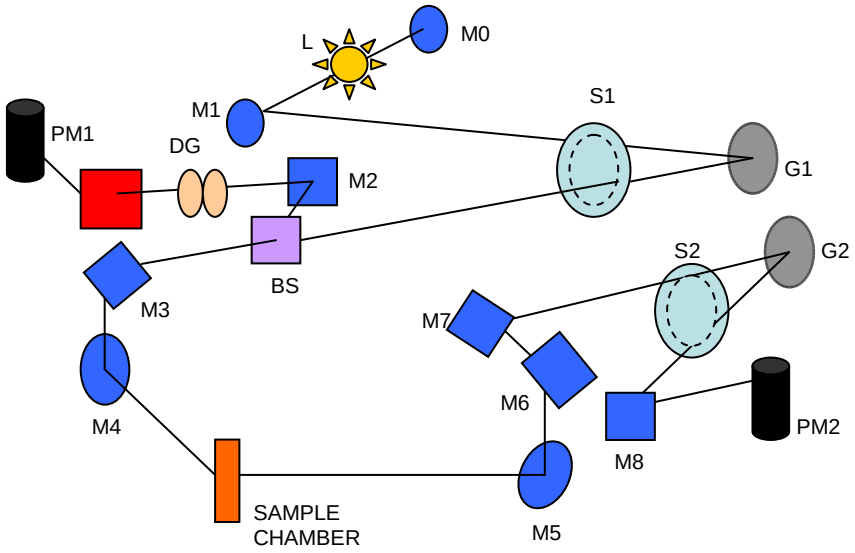


Figure 2.2 - Optical system of the Model FP-6500 spectrofluorometer. See text for explanation (adapted from Jasco-Europe spectrofluorometer model FP-6500 manual).

2.2.2.3. Spectral acquisition and analysis

Thioflavin-T fluorescence emission spectra were measured with excitation wavelength of 425 nm and recording between 430 nm and 550 nm at a scanning speed of 100 nm/min (data pitch 1 nm) and band width of 3 nm using a Jasco FP-6500 instrument as aforementioned, provided with a water bath Julabo F-12 allowing measurements at a constant temperature.

ThT stock solution (100 μM in 20 mM buffer) was prepared and kept in the freezer in an opaque bottle since ThT aqueous solutions are sensitive to light and to relative high temperatures. Fresh ThT solutions of 20 μM were prepared by dilution of the stock with buffer 20 mM before fluorescence measurement and kept in the fridge between measurements covered with aluminium foil. Samples for ThT-binding measurements were prepared by using the 20 μM solution of ThT and β-Lg samples to achieve a protein to ThT ratio of 1:9 by volume (molar ratios of 1:3.3, 1:0.83, 1:0.33, and 1:0.083 for protein concentration of 1, 4, 10, and 40 mg/ml, respectively).

Fluorescence spectra were measured in quartz cells with 1 cm path-length where ThT and samples solutions were mixed and left rest for 5 minutes before measuring to allow ThT-binding. All spectra were recorded at 22°C.

ThT fluorescence emission spectra were normalised by using Spectra Manager Package, provided with the instrument. Fluorescence intensities at 477 nm were read in every spectrum and plotted versus incubation time to allow the study of the structural arrangements kinetic of β-Lg in the presence of 50% v/v ethanol at different protein concentrations. These kinetic data were described in terms of nucleation-dependent growth model when possible and fitted to a sigmoidal equation (equation 2) with OriginPro.8 software package:

$$I = I_1 + \frac{I_0 - I_1}{1 + \exp(\frac{t - t_0}{g})} \qquad \text{Equation 2}$$

Where

- I is the observed fluorescence signal at 477 nm
- I_0 and I_1 are the initial and final plateau values respectively
- t is the time (hours)
- t_0 denotes the half-time for aggregation at the point of inflection of the fit (hours)
- 1/g is a parameter describing the magnitude of the increase of the signal ($hours^{-1}$).

From the fitted data, maximum structural rearrangements rate was calculated as the slope of the straight line through the point of inflection (value of the derivative of the sigmoidal curve in t_0). The lag times were defined as the intercept of this line with the abscissa at the plateau average value (Nielsen et al. 2001; Grudzielanek, Janse, and Winter 2005).

2.2.3. Circular Dichroism (CD)

Circular Dichroism (CD) is a spectroscopic technique used to analyse changes in the secondary and tertiary structure of the protein. Spectra recorded in the far UV (190-260 nm) gives information about the secondary structure of the protein, while analysis in the near UV (260-320 nm) shows information about changes in its tertiary structure. Quantitative analysis of far UV spectra allows knowing the percentages in which each type of secondary structure is present in the protein under study. CD is a fast (samples can be measured in the time scale of minutes), simple, and reliable technique although it gives less accurate information than other techniques used for the study of the tertiary structure of proteins such as nuclear magnetic resonance (NMR) or X-ray diffraction.

2.2.3.1. Theory

Plane-polarized radiation can be considered as being composed by two circular polarized components of equal magnitude: one rotating counter-clockwise (left handed, L) and the other clockwise (right handed, R). When the two components are combined without being altered (when L and R are not absorbed or are absorbed at the same level), radiation polarized in the original plane is regenerated. However, when an optically active sample (chiral) intercept this radiation, L and R are absorbed to different level and the resulting radiation would be said to possess elliptical polarisation (Figure 2.3) (Meyers and Robert, 1995).

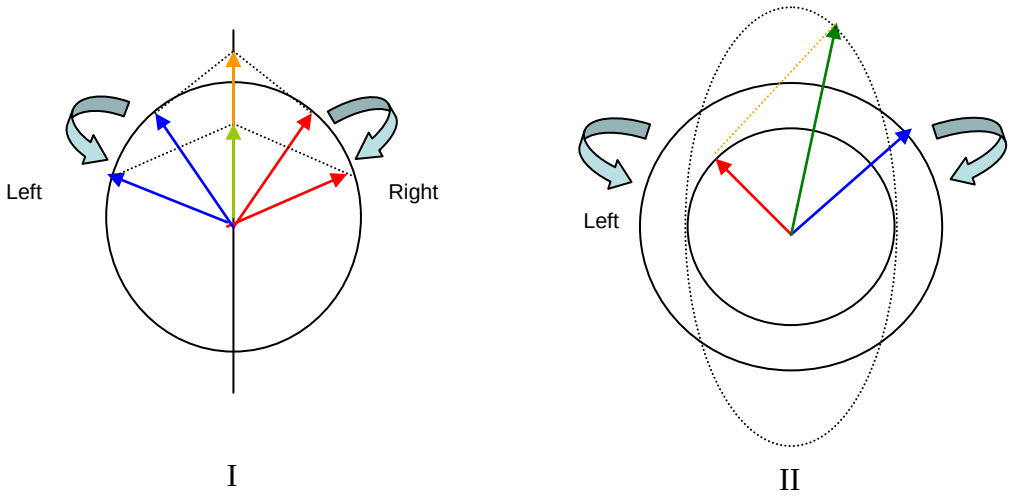


Figure 2.3 - Origen of the CD effect. (I) the two components have the same amplitude and when combined generate plane polarised radiation. (II) The components are of different magnitude and the resultant (dashed line) is elliptically polarised. Adapted from Kelly et al., 2005.

(i) *Tertiary structure analysis*

Analysis in the near UV (260-320 nm) gives information about the tertiary structure of the protein. Is in this region where aromatic amino acids absorb: tyrosine (c.a. 275-282 nm), phenylalanine (c.a. 255-270 nm), and tryptophan (c.a. 290-305 nm). Near UV spectra, besides providing a "fingerprint" of the tertiary structure, can also provide information about the presence of the "molten globule" state of the protein, characterised by a very weak near UV signal, reflecting the high mobility of the aromatic side chains (Ptitsyn, Bychkova, and Uversky, 1995; Kelly et al., 2005).

(ii) *Secondary structure analysis*

The different types of regular secondary structure found in proteins give rise to characteristic CD spectra in the far UV region (Kelly, Jess, and Price, 2005). The peptide bond is the principal absorbing group in this region. In Figure 2.4, the common shape of the far UV spectra for the different types of second structures is shown.

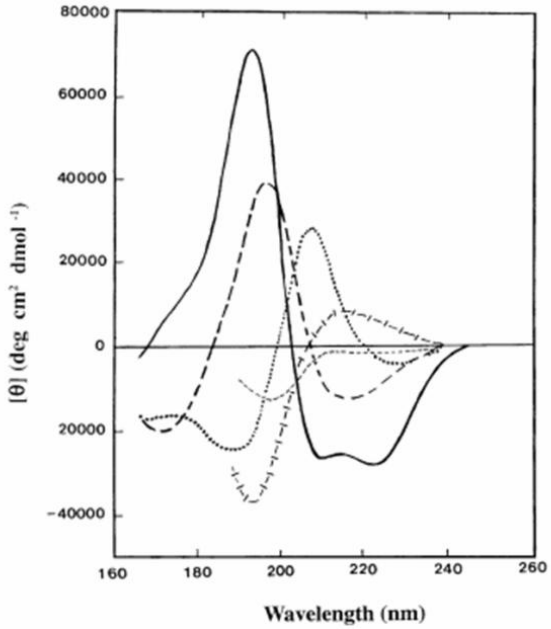


Figure 2.4 - Far UV spectra associated with various types of secondary structure. Solid line, α-helix: long dashed line, anti-parallel β-sheet; dotted line, type I β-turn; cross-dashed line, extended 31-helix or poly (Pro) II helix; short dashed line, irregular structure. Retrieved from Kelly et al., 2005.

It is possible to obtain estimates of the α-helical content of proteins from using the values of the CD signals at 208 nm and 222 nm (Barrow, Yasuda, Kenny and Zagorski, 1992). However, these estimates have to be treated with caution. For this thesis, quantitative analysis of far UV spectra was carried out using the online server DICHROWEB - hosted at Birkbeck Collage, University of London, U.K. This server allows data to be entered in different formats and it is compatible with the major CD instruments manufacturers. Different algorithms can be chosen to analyse CD data. For this project, CONTIN was selected as the algorithm to estimate the secondary structure composition of the protein under study (Provencher and Glockner, 1981; Van Stokkum, Spoelder, Bloemendal, Van Grondelle and Groen, 1990). For full details about DICHROWEB and quantitative analysis of far UV data, see APPENDIX A.

2.2.3.2. CD instrument

A general spectropolarimeter consists of a light source, a prism to produce the plane-polarized radiation, a monochromator to select the chosen wavelength, a modulator to resolve the components of the polarized radiation (L and R) and finally a photomultiplier to detect and read the radiation after passing through the sample. A Jasco J-810 instrument (University of Glasgow) was used to obtain CD data for this project. A block diagram of this machine is shown in Figure 2.5.

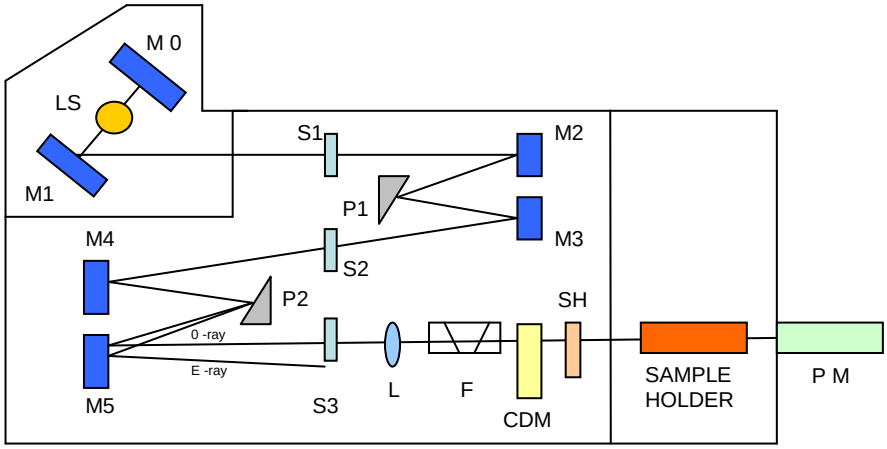


Figure 2.5 - Schematic representation of a spectropolarimeter (Jasco J-810). Plane polarised radiation is produced by passage of light from the source (LS) through two prisms (P1 and P2) and a series of mirrors (M0 to M5) and slits (S1 to S3). The ordinary ray (O) is focussed by a lens (L), and passed through a filter (F) to the modulator (CDM). The circularly polarised components are then passed through the shutter (SH) to the sample compartment, before detection by the photomultiplier (PM) (E represents the extra-ordinary ray). Adapted from Kelly et al., 2005.

The resulting dichroism will be displayed as either the difference in absorbance ($\Delta A = A_L - A_R$) or as the ellipticity in degrees (θ); $\theta = \tan^{-1}(b/a)$ where b and a are the minor and major axes of the resulting ellipse (part (II) of Figure 2.3) (Meyers and Robert, 1995). The CD spectrum is obtained when the dichroism is measured as a function of wavelength. Depending on the UV region under study - near UV or far

UV - different parameters were fixed in the spectropolarimeter J-810 to obtain the best response. These parameters are summarised in Table 2.1.

Table 2.1 - Summary of parameters used in CD measurements.

PARAMETERS	Near UV	Far UV
Sensitivity	Standard 100 mdeg	Standard 100 mdeg
Start	320 nm	260 nm
End	250 nm	185 nm
Data Pitch	0.2 nm	0.2 nm
Scanning Mode	Continuous	Continuous
Scanning Speed	10 nm/min	50 nm/min
Response	2 seconds	0.5 seconds
Band width	1 nm	1 nm
Accumulation[1]	3	5
Path length	0.5 cm	0.02 cm

[1]Number of scans run for each sample, from which Jasco J-810 automatically calculates the average as the final output from CD measurements.

2.2.3.3. Spectral acquisition and analysis

Different quartz cells were used for near and far UV regions. Cylindrical cells of 0.02 cm path length were used for far UV region measurements while square cells of 0.5 cm path length were used for studies in the near UV region. Same cells were used both for samples and blanks and during the same experiment to reduce variables that could lead to increase experimental errors. For CD analysis, quartz cells needs to be carefully washed to avoid unreliable data. Before each measurement, cells were

washed with tap water, followed by distilled water, concentrated nitric acid, distilled water once again and finally ethanol before drying with compressed air.

When protein concentration was higher than 1 mg/ml, these samples had to be diluted to achieve a final concentration in cells of 1 mg/ml with the aim of getting readable CD signals, both in the far and near UV region.

Raw data, given by the spectropolarimeter as CD (mdeg), was corrected to concentration using Spectra Manager software available with the instrument to give results in molar ellipticity ($deg \cdot cm^2 \cdot dmol^{-1}$). The mean residue ellipticity at wavelength λ ($[\theta]_{WRW,\lambda}$) is given by:

$$[\theta]_{MRW,\lambda} = \frac{MRW \cdot \theta_\lambda}{10 \cdot d \cdot c} \qquad \text{Equation 3}$$

Where:

- *MRW* is the mean residue weight (i.e. molecular mass of the protein (in Da) divided by the number of amino acids of the protein minus 1)
- θ_λ is the observed ellipticity (degrees)
- *d* is the path length of the cell (cm)
- *c* is the protein concentration (g/ml)

However, since the molar concentration of the protein is known from the UV absorbance measurements, the molar ellipticity at wavelength λ ($[\theta]_{WRW,\lambda}$) is given by:

$$[\theta]_{molar,\lambda} = \frac{100 \cdot \theta_\lambda}{m \cdot d} \qquad \text{Equation 4}$$

Where *m* is the molar concentration (mol/L) of the protein, calculated dividing the concentration obtained from UV absorbance using Equation 1 (in mg/ml) by the *MRW* of the protein. The rest of symbols have the same meaning as above.

For β-Lg, *MRW* was calculated to be 113.196 Da. When far and near UV spectra were normalised, concentration obtained from UV absorbance spectroscopy as it was explained in section 2.2.1 was divided by the MRW for β-Lg given the needed protein concentration in mol/L to be introduced in Spectra Manager software for CD data correction.

Before data correction, the corresponding blank was subtracted to raw data. Different smoothing algorithms provided with the software were selected to reduce the noise in CD spectra. Once data was corrected, every spectrum was normalised to zero in the X-axis. Far UV spectra were quantitatively analysed further using the online software DICHROWEB (see APPENDIX A).

2.2.4. Attenuated Total Reflection – Infrared Spectroscopy (ATR-IR)

Attenuated total reflection – infrared spectroscopy (ATR-IR) was used with the aim of complementing results from circular dichroism measurements for the analysis of the secondary structure of the protein. The absorbance spectrum obtained with this spectroscopic technique gives information about the secondary structure of the protein in the Amide I band (1600 – 1700 cm^{-1}).

2.2.4.1. Theory

Infrared spectroscopy (IR) provides information about the secondary structure of the protein (as the analysis in the far UV region with circular dichroism). IR spectroscopy works by shining infrared radiation on a sample and seeing which wavelength of radiation in the infrared region is absorbed by the sample. Characteristic bands found in infrared spectra of proteins include Amide I (~1600 – 1700 cm^{-1}) and Amide II (~1510 and 1580 cm^{-1}). These arise from the amide bonds that link the amino acids. The absorption associated with Amide I band leads to stretching vibrations of the C=O bond of the amide while absorption associated with the Amide II band leads primarily to bending vibrations of the N-H bond. These two bonds are involved in the hydrogen bonding that takes place between the different elements of the secondary structure of a protein (hydrogen bonding between turns

within α-helix structure or hydrogen bonds between strands in a β-sheet structure). The locations of Amide I and Amide II bands are sensitive to the secondary structure content of proteins. However, Amide II is not as good as Amide I predictor for quantitative analysis of the secondary structure. Elliot and Malcolm (1956) first demonstrated an empirical correlation between the frequency of the maximum of the C=O band and the secondary structure of a range of proteins, revealing a powerful technique for structural analysis of polypeptides and proteins. However, two major problems restricted this application of IR spectroscopy. The first was the strong absorption of water in the spectral region of interest. This problem was overcome by using cells with short path-lengths and mainly by using deuterated water (D_2O) as solvent instead of H_2O. The second problem was because a protein contains more than one type of secondary structure: each protein is composed of varying percentages of the three types of secondary structure, namely α-helix, β-sheets and turns, apart of the less ordered polypeptide segments. Each of these structures give rise to a C=O band and the relative widths and peak separations are such that these individual bands cannot be resolved by increasing the instrument resolution. The result is then a broad C=O band which needs to be deconvoluted to have reliable quantitative information about the content of the different types of secondary structure. With this aim, several band narrowing (also called resolution enhancement) methodologies have been developed in recent years. A large number of protein and polypeptides whose structures have been determined by X-ray diffraction have been examined by infrared spectroscopy and correlations made between the frequency of the C=O and the principle components of the protein secondary structure deduced from X-ray diffraction. Correlation of structure with C=O is relatively straightforward for proteins that have a high β-sheet or α-helix character. Thus, absorption in the regions 1630-1640 cm^{-1} and 1648-1657 cm^{-1} are generally attributed to β-sheets and α-helix respectively (Byler and Susi, 1986; Surewicz and Mantsch, 1988; Dong, Huang, and Caughey, 1990; Jackson and Mantsch, 1991; Zandomeneghi, Krebs, Mccammon, and Fandrich, 2004). However, there are evidences to conclude that IR spectroscopy is a suitable technique to analyse quantitatively the content of β-sheets structured in proteins but it is not so

good to analyse α-helix content (Byler and Susi, 1986). The wavenumbers at which different types of secondary structure are attributed also vary slightly in literature. The IR pattern shown by Sakuno, Matsumoto, Kawai, Taihei, and Matsumura (2008) is followed in this thesis for the analysis of IR spectra for β-Lg. Table 2.2 summarises the type of secondary structure associated with each wavenumber.

Table 2.2 - Secondary structure corresponding to each wavenumber in IR spectra.

	α-helix	**Intra-molecular β-sheet**	**Inter-molecular β-sheet**	**β-turn**	**Random coil**
Wavenumber (cm^{-1})	1652-1658	1620-1630	1681-1691	1662-1675	1640-1647

Attenuated total reflection (ATR) is a measurement technique used in conjunction with infrared spectroscopy allowing the examination of samples in the solid or liquid state. ATR uses a property of the optical phenomenon known as total internal reflection called the evanescent wave. An infrared beam passes through the ATR crystal in such a way that it reflects (at least once) off the internal surface in contact with the sample. This reflection forms the evanescent wave, which extends into the sample, typically by a few micrometers. A detector then collects the beam as it exits the crystal. ATR crystals work best when they are made from an optical material with a refractive index substantially higher than the refractive index of the sample under study. Typical materials for ATR crystals include germanium, thallium bromoiodide (KRS-5) and zinc selenide. The excellent mechanical properties of diamond make it an excellent material for ATR, especially for solid sample analysis, although its higher cost also implies its less wide use (Harrick, 1967, p. 342).

2.2.4.2. ATR-IR instrument

Infrared spectra were recording using a 10-bounce ATR Zn-Se crystal through MB 3000 apparatus from ABB. This machine consists in a typical Fourier-transform-

infrared-spectrophotometer, but with a horizontal attenuated total reflectance accessory supplied by PIKE technologies. A general diagram showing how a typical infrared spectrophotometer works can be seen in Figure 2.6. A schema of the ATR device is shown in Figure 2.7.

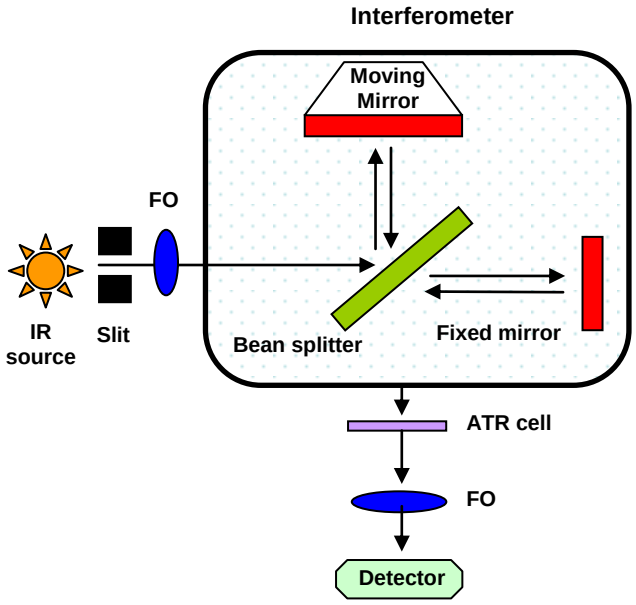


Figure 2.6 - Diagram of a typical IR spectrophotometer: an infrared light is directed by focussing optics (FO) at a beam splitter, which splits it into two beams. One beam is reflected off a flat, stationary mirror while the second beam is reflected off a moving, flat mirror. The equipment maintains the non-stationary mirror moving back and forth at constant velocity. Both mirrors reflect their light beams back to the beam splitter where they are recombined by an interferometer. The new beam, created by the merging of the two reflected beams, is called an interferogram, which is going through the sample. Another focussing system directs the beam to the detector, where the signal is read.

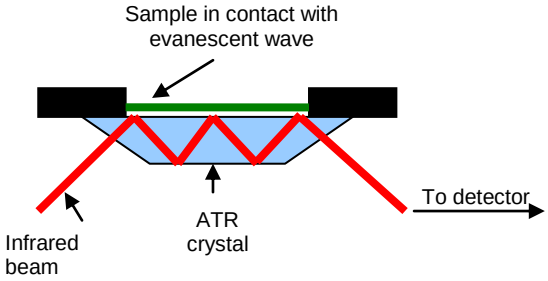


Figure 2.7 - ATR device base: the Zn-Se crystal reflects the IR beam off the surface in contact with the sample. This reflected beam (evanescent beam) extents into the sample and a detector collects the beam when it exits the crystal. Adapted from PerkinElmer - Technical Notes, 2005; available online at www.perkinelmer.com

2.2.4.3. Spectra acquisition and analysis

Infrared spectra were obtained using the instrument described above. For each spectrum, a 256-scan interferogram was collected in absorption mode with a resolution of 4 cm^{-1} at room temperature and analysed in the range of 1400 – 1800 cm^{-1}. Reference spectra (buffer – ethanol solutions 1:1 by volume) were recorded under identical conditions in single beam mode. Secondary structure analysis was performed using GRAMS 8 software provided with the instrument, which deconvolutes the broad Amine I band into specific peaks characterising secondary protein structures.

2.2.5. Dynamic Light Scattering (DLS)

Dynamic Light Scattering was the technique used to determine the average size of protein clusters in solution. Information provided from this technique allows analysing the kinetics of the formation of nucleated species of β-Lg at various protein concentrations.

2.2.5.1. Theory

Light scattering is a mechanism of absorption and re-emission of electromagnetic radiation. When light is coming into contact with a particle, if the wavelength of the light corresponds to an absorption band of the material, the particle absorbs the light. However, if that light has a different wavelength, it is radiated out in all directions. This secondary radiation is the scattered light. The incident light is a monochromatic laser source, i.e. a light source with a single wavelength or circular frequency. When the sample is formed by a perfectly homogeneous and isotropic material, the radiation scattered by the individual particles interferes destructively so that no scattered radiation is observed. Hence, scattered radiation is only observed when the measured sample is heterogeneous. Light scattering experiments can be of two different categories:

(i) Time average experiments, where only static properties of the sample can be studied such as the molecular weight determination or the radius of gyration of dispersed macromolecules or particles. Hence, these experiments are classified as static light scattering (SLS) experiments.

(ii) Time resolved experiments at one or several scattering angles, which yield information about time dependent properties (intensity fluctuations) of the samples, related to the fluctuations in the intensity of the scattered light. Therefore, these experiments are called dynamic light scattering (DLS) experiments and yield, most of all, the particle average diffusion coefficient, which is then related to an apparent particle size, the so-called hydrodynamic radius (Finsy, 1994).

The hydrodynamic radius of a molecule is defined as the radius of a hard sphere diffusing at the same rate as the given molecule including relevant hydration and shape effects (Finsy, 1994). Therefore, this term only refers to the size of particles in solution and it is different from the theoretical definition of the radius of gyration obtained from techniques such as X-ray diffraction or neutron scattering. The International Union of Pure and Applied Chemistry (IUPAC) defines the radius of gyration as a parameter characterising the size of a particle of any shape. For a rigid

particle, consisting in elements of mass, m_i, at a distance r_i from its centre of gravity, the radius of gyration is, by definition, the square rood of the mass-average of the r_i^2 for all the mass elements. For non-rigid particles, such as proteins, an average of all conformations has to be considered (IUPAC Compendium of Macromolecular Nomenclature, 1991, p.48).

In this project, DLS was used to obtain the hydrodynamic radius of the protein aggregates in solution, while collaborators obtained the radius of gyration from SAXS and SANS analysis (APPENDIX M).

Theoretical background of dynamic light scattering (DLS) comes from analysis of the electric field of the light scattered by a sample. However, it is the scattered light intensity that is directly measured instead of the electric field. Intensities and fields are related in a way that both contain the same information about the motion of the particles in solution and its size. Since particles are in constant movement (Brownian motion) changes in the electric field and the scattered intensity are time-dependent. Fluctuations of scattered intensity are mathematically described through the autocorrelation function (consisting in the comparison of a function with a delayed version of itself for every delay time (τ)). This autocorrelation function is applied to the electric field and the scattered intensity, and is presented in two forms: the normalised time correlation function of the scattered field, $g_1(\tau)$; and the normalised time correlation function of the scattered intensity, $g_2(\tau)$ (Pusey, 2002, chaps. 1 and 9). The normalised time correlation function of the scatted field ($g_1(\tau)$) is an exponentially decaying function of time delay τ. $g_1(\tau)$ and $g_2(\tau)$ are related to each other in the following way:

$$g_1(\tau) = \sqrt{(g_2(\tau) - 1)} = \exp(-\Gamma \cdot \tau) \qquad \text{Equation 5}$$

Where

- Γ is the decay rate of the autocorrelation function $g_1(\tau)$ (1/s)
- τ is the delay time (s)

The decay rate Γ is defined as:

$$\Gamma = D \cdot q^2 \qquad \text{Equation 6}$$

Where

- D is the diffusion coefficient of the particles in solution (related to Brownian motion) (m^2/s)
- q ($1/m^2$) is the scattering vector amplitude defined as:

$$q = \frac{4 \cdot \pi \cdot \eta}{\lambda} \cdot \sin\left(\frac{\theta}{2}\right) \qquad \text{Equation 7}$$

Where

- η is the refractive index of the solvent
- λ is the wavelength of the laser (m)
- θ is the angle at which the scattered light is measured (deg)

The average hydrodynamic radius R_h can be calculated from the Stokes-Einstein equation:

$$R_h = \frac{k_B \cdot T}{6 \cdot \pi \cdot \mu \cdot D} \qquad \text{Equation 8}$$

Where

- k_B is the Boltzmann constant ($(kg \cdot m^2)/(s^2 \cdot K)$)
- T is the absolute temperature (K)
- μ is the solvent viscosity at the given temperature ($kg/(m \cdot s)$).

By fitting the first decay of the natural logarithm (LN) of the autocorrelation function ($g_1(\tau)$) to a straight line by the means of the cumulant method, the decay rate (Γ) is obtained and D can be calculated since q is known for a specific experiment. By using the Stokes-Einstein equation, the apparent hydrodynamic radius of the particles in solution can be obtained. From DLS theory, it can be concluded that the decay rate

of the autocorrelation function (seen as the "shoulder" in Figure 2.8) is inversely proportional to the size of particles: the faster the decay, the smaller the particles in solution since they would be moving faster than big particles since they have higher diffusion coefficient (Equation 8).

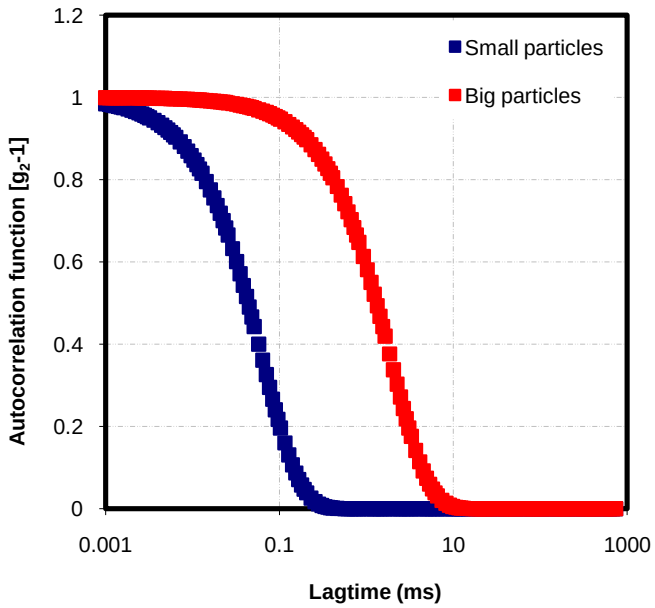


Figure 2.8 - Relation between autocorrelation function and particle size in DLS output data.

When particles of different sizes are present in enough quantity to develop a DLS signal, different decays may appear in the autocorrelation function as several different shoulders. To calculate the overall particle size distribution, the overall decay pattern in the autocorrelation function needs to be analysed (Luthi-Peng and Puhan, 1999).

2.2.5.2. DLS instrument

The instrument used for DLS measurements consists of a laser beam (the light source) which passes along a set of optical elements to guide and focus the beam before coming through the sample, which is settled in a vat containing the index matching liquid, acting as a reference. A second set of optical elements collects the

scattered light before the detector. This detector can be positioned in different angles for different types of measurements. It is connected to a correlator, which is linked to a PC. An electronic box control connects the instrument with a PC to allow the control of the system through the suitable software. A diagram of a general DLS instrument can be shown in Figure 2.9.

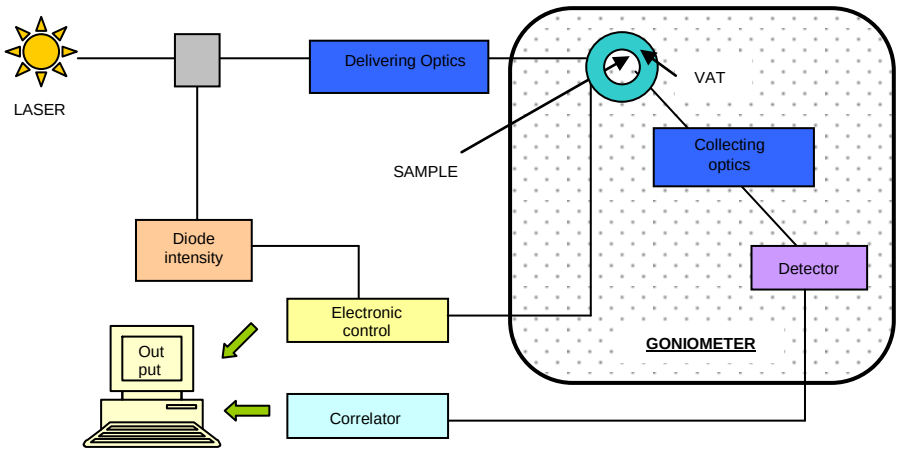


Figure 2.9 - Schematic diagram of DLS instrument.

2.2.5.3. Spectral acquisition and analysis

DLS experiments were carried out with an ALV/CG3 Compact Goniometer supplied by ALV and together with an ALV/LSE-5004 electronic system. The light source was a red laser beam with a wavelength of 632.8 nm. The index matching liquid inside the vat was toluene, whose refractive index is 1.496 at 20°C (Sigma-Aldrich product specifications). Cells for measurements are glass tubes of 10 mm wide (with lids) which were cleaned with filtered water (0.02 μm ANATOP syringe filter) before measurements. DLS requires special care in the treatment of cells and samples since there is a risk of contamination due to the presence of dust. Dust particles would scatter much more light than the particles of interest because they are usually much bigger than the particles of interest, leading to unreliable DLS data. Therefore, all stock samples and reagents subjected to DLS measurements were filtered

previous use within different devices depending on the nature of the experiment. Syringe filters of 0.45 μm (PTFE) and 0.1 μm (ANATOP) pore-size were used in filtration procedures.

Autocorrelation functions where recorded for each sample as the average of 3 runs lasting 10 seconds each. The output signal given by ALV / CG3 Compact Goniometer is $g_2(\tau)$-1 (see DLS theory in section 2.2.5.1). To calculate the apparent hydrodynamic radius (R_h) from these data, the cumulant method is used as it was mentioned above. The output data ($g_2(\tau)$-1) was transformed into $g_1(\tau)$ by Equation 5 and then the natural logarithm (LN) of $g_1(\tau)$ was calculated for each decay time.

$$g_1(\tau) = \sqrt{(g_2(\tau) - 1)} = \exp(-\Gamma \cdot \tau) \rightarrow LN(g_1(\tau)) = -\Gamma \cdot \tau$$

The cumulant method consists of fitting the first decay of the autocorrelation function (expressed in terms of LN ($g_1(\tau)$) to a polynomial equation up to the third order:

$$LN\ (g_1(\tau)) = a + b \cdot \tau + c \cdot \tau^2 + d \cdot \tau^3$$

By comparing both expressions above, it is clear that the first order coefficient (b) is proportional to the decay rate, Γ:

$$b \approx \Gamma$$

In the case of monodisperse samples, the autocorrelation function is fitted to a first order polynomial and LN ($g_1(\tau)$) vs. τ is a straight line whose slope is the decay rate with units 1/time. This decay rate (Γ) is related to the diffusion coefficient of the particles in solution as it is observed in Equation 6. By combining Equations 6 and 8, it is concluded that the decay rate (Γ) is inversely proportional to the hydrodynamic radius (R_h) of the particles. However, the probability of having monodisperse systems is low. Additionally, the higher the order of the fitting, the more accurate the

apparent hydrodynamic radius values will be. Therefore, every first decay of the autocorrelation functions measured were fitted to a second order polynomial function where the first order coefficient is proportional to the decay rate, Γ:

$$LN\ (g_1(\tau)) = a + b \cdot \tau + c \cdot \tau^2 \rightarrow b \approx \Gamma$$

$$\Gamma = D \cdot q^2 \rightarrow R_h = \frac{k_B \cdot T}{6 \cdot \pi \cdot \mu \cdot D} = \frac{k_B \cdot T \cdot q^2}{6 \cdot \pi \cdot \mu \cdot \Gamma}$$

The fit was done using Microsoft Excel package. However, as in every fitting procedure, there is an error associated with this fit. The mean error of the fit can be expressed as the standard deviation (σ) based on the least square theory as it is concluded from DLS theory (Pusey, 2002, chap. 9). This error takes into account the difference between the real (raw) data (LN (g_1(τ)) and the value of LN (g_1(τ)) calculated from the second order polynomial function used for the fit. To obtain the error in the apparent hydrodynamic radius calculated from the fitting parameter *b*, the error propagation formula has to be applied. For a detailed explanation about second order fitting and error calculations in R_h see APPENDIX B.

It is important to note here that, for samples containing ethanol, calculation of the apparent hydrodynamic radius must include the right viscosity and refractive index values to be introduced in Equation 8 (depending on ethanol concentration) and not just the values corresponding with the viscosity and refractive index for water, settled by default in the instrument. Several repetitions were done for experiments under the same conditions; the standard deviation of individual hydrodynamic radii obtained from different repetitions of experiments was calculated as indicated by Equation 9:

$$\sigma_{Rh} = \sqrt{\frac{\sum \left(x_i - \bar{x}\right)^2}{N-1}} \qquad \text{Equation 9}$$

Where

- x_i represents the individual values of R_{hi} for each repetition
- $\bar{x}$ is the average of the hydrodynamic radius obtained for each of these repetitions
- N is the number of individual values of R_{hi} (number of repetitions) (see APPENDIX B for further details).

The final apparent hydrodynamic radius values, R_h, are given finally by:

$$R_h\ (nm) \pm \sigma_{R_h}(nm)$$

The wide range of the spectroscopy and scattering techniques above mentioned will provide a robust protocol to characterise extensively the formation of nucleated species of β-Lg under the studied conditions. This characterization shall include analysis of the lag period in the aggregation process as well as structural and size transitions undergone by the protein within the experimental time. Experimental protocols for the preparation of samples are shown in the following sections.

2.3. EXPERIMENTAL CONDITIONS AND PROCEDURES

In this project, a general and systematic experimental procedure was developed to assure minimal experimental errors and to optimise experimental time and analysis. This systematic procedure involves the preparation of protein stocks solutions from which the samples under studies were prepared. When experiments differed from the general design, required variations were done to meet the objectives of the experiment in each case.

Experiments done during this project can be classified into three major groups. Results for each of this type of experiments will be shown in the following chapters in this thesis. These three groups of experiments include:

(i) analysis of β-Lg aqueous solutions were the protein is expected to be in its native state (using both distilled water and phosphate buffer as solvents),

(ii) analysis of β-Lg samples in the present of different percentages of ethanol for 10 and 40 mg/ml of protein (1 and 4 wt %),

(iii) analysis of β-Lg samples in the presence of 50% (v/v) of ethanol for different concentrations of β-Lg.

Unless otherwise stated, all samples were incubated at room temperature (22°C) and the pH of the protein solution was always adjusted to pH 7 using a pH 212 HANNA Microprocessor pH-meter previous addition of ethanol when required. Sample preparation will be explained in detail for every type of experiment in the following sections.

2.3.1. Experimental procedure for experiments involving β-Lg aqueous solutions: β-Lg in its native state.

Samples of β-Lg containing only the protein and the solvent were needed to have the appropriate background to fully analyse the effect when a denaturant solvent is added

to the protein. These β-Lg native samples were prepared in distilled water and in different concentrations of phosphate buffer (10 mM and 20 mM at pH 7) to see if the presence of various concentrations of salts would affect the behaviour of the protein in its native state.

Since native samples with different concentrations of protein were required (1, 4, 10 and 40 mg/ml – 0.1, 0.4, 1 and 4 wt % respectively) a protein stock solution of 80 mg/ml (8 wt %) was prepared by weighing the corresponding quantity of protein powder depending in the volume of stock needed. This protein powder was then dissolved during several hours in the corresponding volume of distilled water, phosphate buffer 10 mM – pH 7 or phosphate buffer 20 mM – pH 7 (see preparation of phosphate buffers in section 2.1.2) using a magnetic stirrer. This high concentrated protein stock solution was then diluted to obtain less concentrated stocks of 2, 8, and 20 mg/ml of protein. Dilutions were made with the corresponding amount of distilled water, phosphate buffer 10 mM – pH 7 or phosphate buffer 20 mM – pH 7 according with the solvent present in the high concentrated stock of 8 wt %. When stocks were prepared in distilled water, pH was adjusted to pH 7 by adding HCl 0.1 M or NaOH 0.1 M as required. The four protein stocks were filtered using a syringe pump (Cole Parmer Instrumental Company. 74900 Series) and syringes of 5 ml with a diameter of 1.29 cm. Two filtrations were done for each stock: firstly, protein solutions were filtered by a 0.45 μm syringe filter (PTFE) at a rate of 40 ml/hour; secondly, the filtered solutions were filtered a second time through a syringe filter of 0.1 μm (ANATOP) at a rate of 5 ml/hour. Both types of syringe filters were purchased from Whatman, plc UK.

Filtered protein stocks of 2, 8, 20, and 80 mg/ml were diluted further to achieve β-Lg native solutions of 1, 4, 10, and 40 mg/ml. Last dilutions were done again with distilled water, phosphate buffer 10 mM - pH 7 or phosphate buffer 20 mM-pH 7 depending on the experiment in mind. These solvents were filtered through a 0.45 μm (PTFE) syringe filter before being added to the protein stock.

It is important to note here that the preparation of aqueous solutions of β-Lg was usually part of a more complex experiment including also the preparation of β-Lg samples with 50% of ethanol (v/v) for different concentrations of protein prepared from the same stock. That is why high concentrated stocks of 2, 8, and 20 mg/ml were prepared from stock 80 mg/ml before the final β-Lg aqueous solutions and samples including 50% (v/v) of ethanol with protein concentrations of 1, 4, 10, and 40 mg/ml respectively (section 2.3.3). However, when experiments were only designed for the preparation of aqueous solutions of β-Lg, intermediate stocks were also prepared to follow a systematic sample preparation with the aim of reducing the number of variables that could affect reproducibility of results.

Aqueous solutions of β-Lg with concentrations 1, 4, 10 and 40 mg/ml were then measured according to the experiment carried out. These samples were analysed by ThT-binding fluorescence, CD, IR and DLS. In Figure 2.10 the experimental plan, solely including the preparation of aqueous solutions of the protein (also called native samples) can be seen.

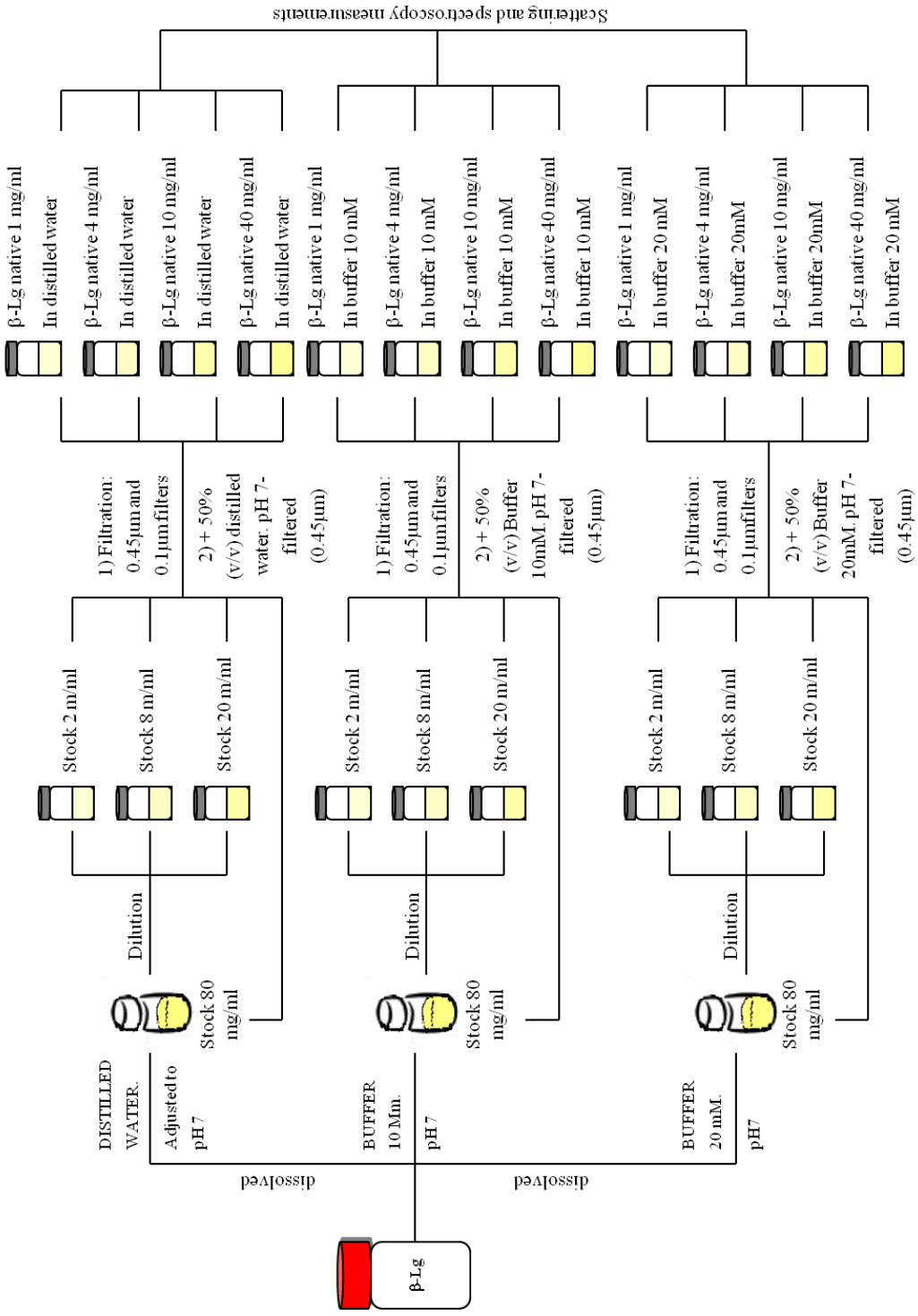


Figure 2.10 - Diagram of the experimental plan for β-Lg aqueous solutions (native samples) preparation.

2.3.2. Experimental procedure for experiments involving β-Lg aqueous solutions with different percentages (by volume) of ethanol

In other to study how different concentrations of ethanol may affect the aggregation behaviour of β-Lg, several samples with a constant concentration of protein with and without phosphate buffer (with a final concentration of 10 mM when used) and different percentages of ethanol were prepared.

With the aim of reducing experimental time and errors, a protein stock solution was made from which samples with different percentages of ethanol were prepared. This experiment was done for two concentrations of β-Lg: 10 and 40 mg/ml (1 and 4 wt % respectively). Stock solutions of 25 and 90 mg/ml were prepared by weighing the corresponding quantity of protein powder and dissolved in distilled water during several hours (using a magnetic stirrer) for experiments carried out for 10 and 40 mg/ml of protein respectively. Protein stocks were filtered following the same procedure described above: filtration through syringe filters of 0.45 μm PTFE at a rate of 40 ml/hour, followed by filtration through 0.1 μm ANATOP syringe filters at a rate of 5 ml/hour. The corresponding amount of these stocks was used for each sample to achieve the desire final concentration of protein in the fixed volume of sample to be prepared. When it was the case, phosphate buffer 0.2 M - pH 7 was also added to the corresponding volume of stocks, to achieve a final concentration of 10 mM. Distilled water was further added to complete the difference between the desire final volume of the sample and the volume of ethanol, protein stock solution, and buffer when used as required for each case. The pH of the sample was adjusted to 7 by adding HCl or NaOH before ethanol addition. In Tables 2.3 and 2.4, a summary of the volumes of all stocks and reagents used in these experiments is shown.

Table 2.3 - Summary of the volumes of stocks and reagents for experiments with different concentrations of ethanol, for samples with protein concentration of 10 mg/ml and a final volume of 4 ml, with and without buffer.

Stock 25 mg/ml (ml)	1.6 ml		**Stock 25 mg/ml (ml)**	1.6 ml	
Phosphate buffer 0.2M pH 7	0.2 ml		**Phosphate buffer 0.2M pH 7**	none	
% ethanol	**Ethanol (ml)**	**Water (ml)**	**% ethanol**	**Ethanol (ml)**	**Ethanol (ml)**
0	0	2.2	**0**	0	2.4
10	0.4	1.8	**10**	0.4	2
20	0.8	1.4	**20**	0.8	1.6
30	1.2	1	**30**	1.2	1.2
40	1.6	0.6	**40**	1.6	0.8
50	2	0.2	**50**	2	0.4

Table 2.4 - Summary of the volumes of stocks and reagents for experiments with different concentrations of ethanol, for samples with protein concentration of 40 mg/ml and a final volume of 4 ml, with and without buffer.

Stock 40 mg/ml (ml)	1.778 ml		**Stock 25 mg/ml (ml)**	1.778 ml	
Phosphate buffer 0.2M pH 7	0.2 ml		**Phosphate buffer 0.2M pH 7**	none	
% ethanol	**Ethanol (ml)**	**Water (ml)**	**% ethanol**	**Ethanol (ml)**	**Water (ml)**
0	0	2.022	**0**	0	2.222
10	0.4	1.622	**10**	0.4	1.822
20	0.8	1.222	**20**	0.8	1.422
30	1.2	0.822	**30**	1.2	1.022
40	1.6	0.422	**40**	1.6	0.622
50	2	0.022	**50**	2	0.222

2.3.3. Experimental procedure for experiments involving 50 % (v/v) buffer-ethanol solutions of β-Lg

To study the formation of nucleated species by β-Lg it is necessary to denature the protein, that is, to disturb its tertiary structure (native state) in solution. Denaturation of the protein would involve the exposure of its hydrophobic core to the solvent, making the protein prone to interact with other protein molecules in solution leading to the formation of nucleated species (protein's aggregates). As it was mentioned in chapter 1, denaturation of the proteins can be induced by different ways including high temperatures and/or organic solvents.

The main system studied in this project consists in solvent-induced denaturation of the whey protein β-Lg and subsequent formation of clusters at room temperature and under conditions of neutral pH and high concentrations of ethanol (50% (v/v)) as the denaturant solvent.

The experimental procedure for the preparation of samples involved once again the preparation of protein stocks from which the actual samples under study were prepared by dilution.

A concentrated protein stock of 80 mg/ml (8 wt %) was prepared by weighing the corresponding quantity of the protein powder. The protein was dissolved in the corresponding volume of phosphate buffer 20 mM – pH 7. A magnetic stirrer was used to dissolve the protein, stirring the solution during several hours (over night when possible) to assure complete solution of the protein. This stock was diluted with phosphate buffer 20 mM to prepared three stocks of protein with concentrations 2, 8 and 20 mg/ml (0.2, 0.8 and 2 wt % respectively). The four stocks were filtered following the same filtration procedure explained in section 2.3.1 to achieve optically clear β-Lg solutions: filtration through 0.45 μm PTFE filters followed by filtration through 0.1 μm ANATOP syringe filters using a pump to assure the filtration is taking place at a constant rate (40 ml/hour and 5 ml/hour for the first and second filtration respectively). From these filtered stocks, final samples of β-Lg including

50% (v/v) of ethanol were prepared (as well as the native samples including only β-Lg and buffer as it was explained in section 2.3.1). For samples including 50% (v/v) of ethanol, the same volumes of stocks 2, 8, 20 and 80 mg/ml and ethanol (previously filtered through 0.45 μm PTFE filters) were mixed in glass vials given samples of β-Lg with protein concentrations 1, 4, 10 and 40 mg/ml, 10 mM phosphate buffer and 50% (v/v) ethanol. Since the protein powder was dissolved in phosphate buffer 20 mM whose pH was adjusted to 7 when prepared (see preparation of buffer in section 2.1.2), the pH of stocks was not further adjusted. However, slight changes in the pH of the stocks in comparison to the pH of the single buffer might be assumed due to the slightly basic nature of the protein powder. The addition of ethanol, however, would shift this pH to a higher value: Gosal et al. (2004) reported that the change in pH by adding ethanol to a β-Lg solution at pH 7 was in the rage 7.5-7.8. Indeed, this shift was also observed in samples prepared for this project. The values of pH reported here will be referred always to the protein solution before the addition of the denaturant solvent.

To have a better insight into the main experimental procedure for the preparation of β-Lg samples with and without ethanol for different concentrations of proteins, see Figure 2.11. This diagram represents the main systematic procedure followed during this research project.

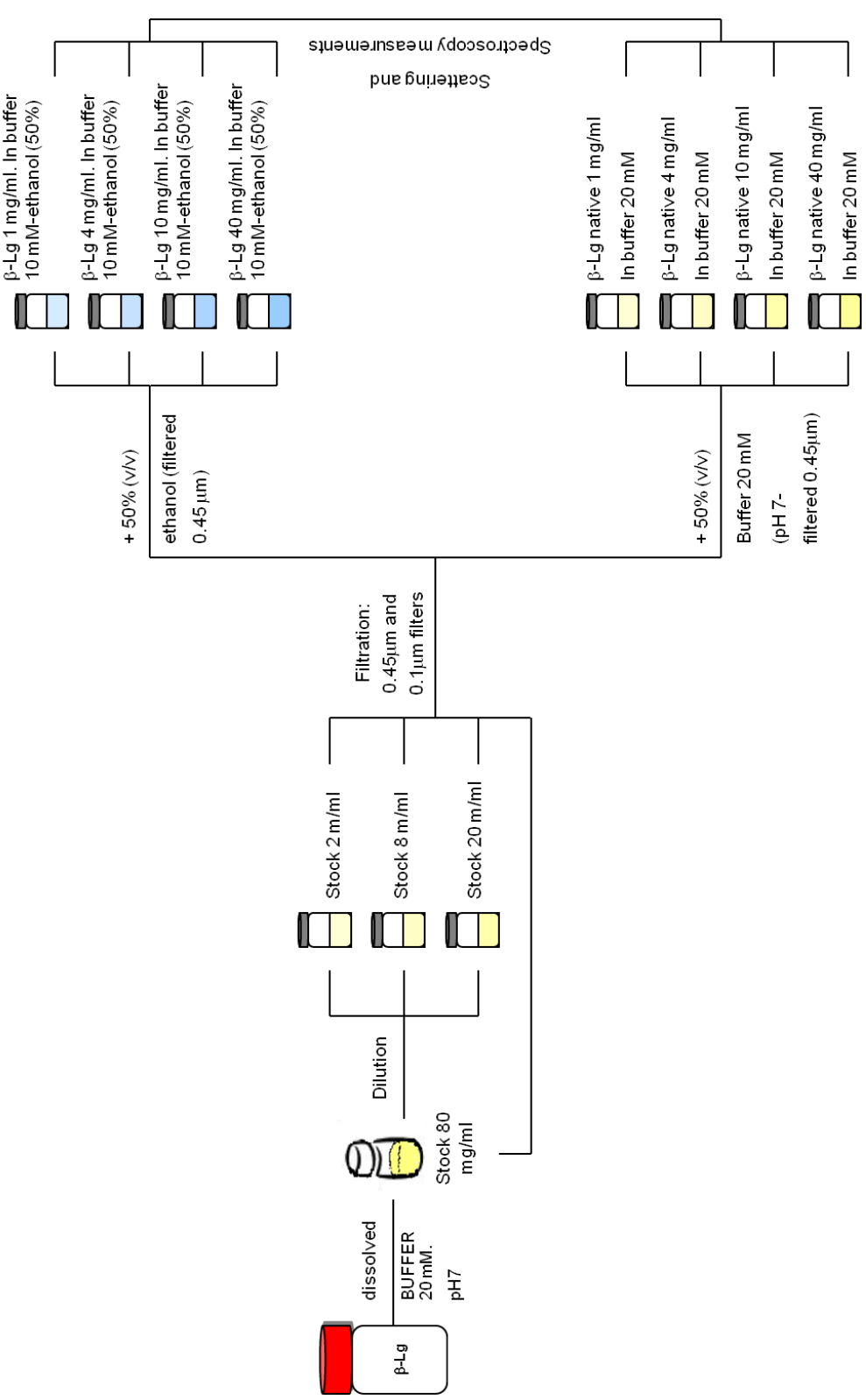


Figure 2.11 - Diagram of the main experimental plan followed during this research project – β-Lg native and actual samples prepared from same stocks.

β-Lg samples in the presence of 50% (v/v) ethanol at different concentrations of protein were then incubated at room temperature (22°C) during 10 days. Different analytical techniques were applied at different intervals of time after sample preparation. Measurements just after adding ethanol were labelled as "Day 0" in results (chapter 3). Samples were also measured after 1, 3, 7, and 10 days of incubation at room temperature.

This time-scale, together with the conditions of neutral pH and room temperature, constitute the main characteristics of the system analysed in this project. The concentration dependence of the formation of β-Lg clusters in the presence of high concentrations of ethanol was studied. However, based on results and the aim of focusing the investigation in the lag period of this aggregation process, samples would require to be analysed at different times, especially shortly after adding ethanol. In the next chapter, results obtained from the analysis of the samples with all the techniques mentioned above will be presented, and the necessity of analysing the samples in shorter periods will become evident.

3. RESULTS

This chapter will present the results obtained from the analysis of the systems under study upon applying the spectroscopic and scattering techniques detailed in section 2.2.

The chapter is organised in three subsections:

(i) Subsection 3.1 will present results obtained for aqueous solutions of β-Lg (dissolved in distilled water and phosphate buffer with concentrations 10 mM and 20 mM). In this environment, the protein is expected to remain in its native state.

(ii) Subsection 3.2 will present results from dynamic light scattering (DLS) measurements done for samples in which different percentages of ethanol (from 0 to 50 % (v/v)) were added to the dissolved protein.

(iii) Subsection 3.3 will finally present the results obtained from the analysis of the main system under study: samples of β-Lg dissolved in 10 mM phosphate buffer (pH 7) with 50 % (v/v) ethanol, incubated at 22ºC over a period of 10 days and measured at different times during the experimental time (after addition of ethanol and after 1, 3, 7 and 10 days). Due to its length and complexity, this section is further divided in subsections for a better comprehension of the reader.

3.1. AQUEOUS SOLUTIONS OF β-Lg: EFFECTS OF PROTEIN AND SALT CONCENTRATION ON AGGREGATION

Since the objective of the project is to analyse the effect of the denaturant solvent (ethanol) on the aggregation behaviour of β-Lg, we first analysed aqueous solutions of the protein (in phosphate buffer and distilled water) without the presence of ethanol. This will allow obtaining reference against which to analyse structural changes and clusters size undergone by the protein when the denaturant solvent is added to the sample. It is important to clarify here that, since the protein would remain in its native state within an aqueous solution environment, these samples are also called "native samples" throughout this and subsequent chapters.

Aqueous solutions of β-Lg in its native state with phosphate buffer concentration 10 mM were analysed with the same spectroscopic techniques as the samples of β-Lg in the presence of ethanol and the same concentration of buffer. The data obtained for the native protein will provide the necessary reference to study the structural transitions and cluster formation undergone by the protein when ethanol is added to the sample.

In Figure 3.1, ThT-binding fluorescence spectra for native samples of 1, 4, 10, and 40 mg/ml containing 10 mM buffer are shown. As it can be seen in this figure, spectra corresponding to the control sample (ThT plus buffer) gave a rather horizontal line as expected, indicating that the samples does not contain structures able to bind ThT. For the rest of the samples, ThT-binding emission spectra for different protein concentrations show a peak around 480 nm (maximum values appear at 477 nm in this study) of wavelength. The intensity of this peak depends on protein concentration; however, neither fibrils nor protofibrils as described in section 1.2.1 are expected to be present in aqueous solutions of the protein since it would be in its native state. This increase in ThT emission spectra can be attributed to the majority content of β-sheets in the native state of β-Lg, whose quantity increases when increasing protein concentration. In fact, other researchers (LeVine, 1995) have

indeed observed the presence of fluorescence intensity for native proteins containing majority of β-sheets in its secondary structure. Given that native β-sheets are unlikely to contain the regular cross β-sheet pattern characterising amyloid fibrils structure with the potential of binding ThT, the binding channels in native β-sheets probable lack this high ordered structure, being on the other hand, distorted and irregular.

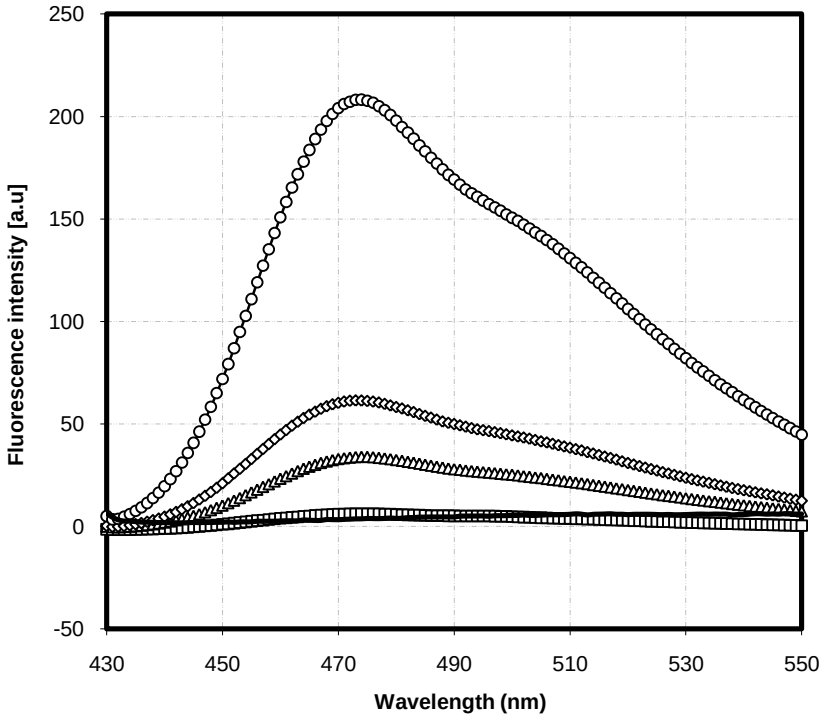


Figure 3.1 - ThT-emission fluorescence spectra for native samples of β-Lg dissolved in 10 mM buffer (pH 7, 22ºC) for different concentrations of protein: Buffer (as the control, thick black line); 1 mg/ml (squares); 4 mg/ml (triangles); 10 mg/ml (diamonds) and 40 mg/ml (circles). Repetition "A" for each concentration. For complete data of β-Lg aqueous solutions, see APPENDIX C.

In Table 3.1.1 values of ThT-binding fluorescence intensity at 477 nm for native samples of β-Lg are summarised. These intensities correspond to the average of all fluorescence intensity values obtained from different repetitions done for native samples with different concentrations of protein (see APPENDIX C). From these

data it could be concluded that fluorescence intensity for native samples seems to be approximately proportional to concentration. These values will be taken into account when the kinetics of structural re-arrangements induced by ethanol is monitored using ThT-binding fluorescence during 10 days (section 3.3.1).

Table 3.1 - ThT-binding fluorescence intensity [a.u] at 477 nm for native samples of β-Lg (buffer concentration 10 mM - pH 7, 22°C) for different concentrations of protein.

	1 mg/ml	**4 mg/ml**	**10 mg/ml**	**40 mg/ml**
Fluorescence Intensity [a.u] at 477nm	6.30 ± 0.46	34.5 ± 0.87	59.5 ± 0.92	216 ± 12.3

Changes in the secondary structure of the protein caused by the addition of ethanol will be also studied by circular dichroism (CD) and attenuated total reflection infrared spectroscopy (ATR-IR). Therefore, native spectra of the protein measured with both techniques will be also required as reference. In Figure 3.2 the CD spectra of β-Lg (corrected to concentration) in its native state in the near (left, A) and far (right, B) UV regions are shown. By quantitative analysis of far UV spectra using DICHROWEB software (APPENDIX A), the percentages of the different types of secondary structure present in the native state of β-Lg can be obtained. They were calculated to be (14 ± 1.1) % of α-helix; (37 ± 1.1) % of β-sheet; (22 ± 0.3) % of β-turns, and (27 ± 0.7) % of random coil. These values are in very good agreement with previous studies reporting the characterisation of β-Lg secondary structure in its native state (Kabs and Sande, 1983; Robert et al., 2002).

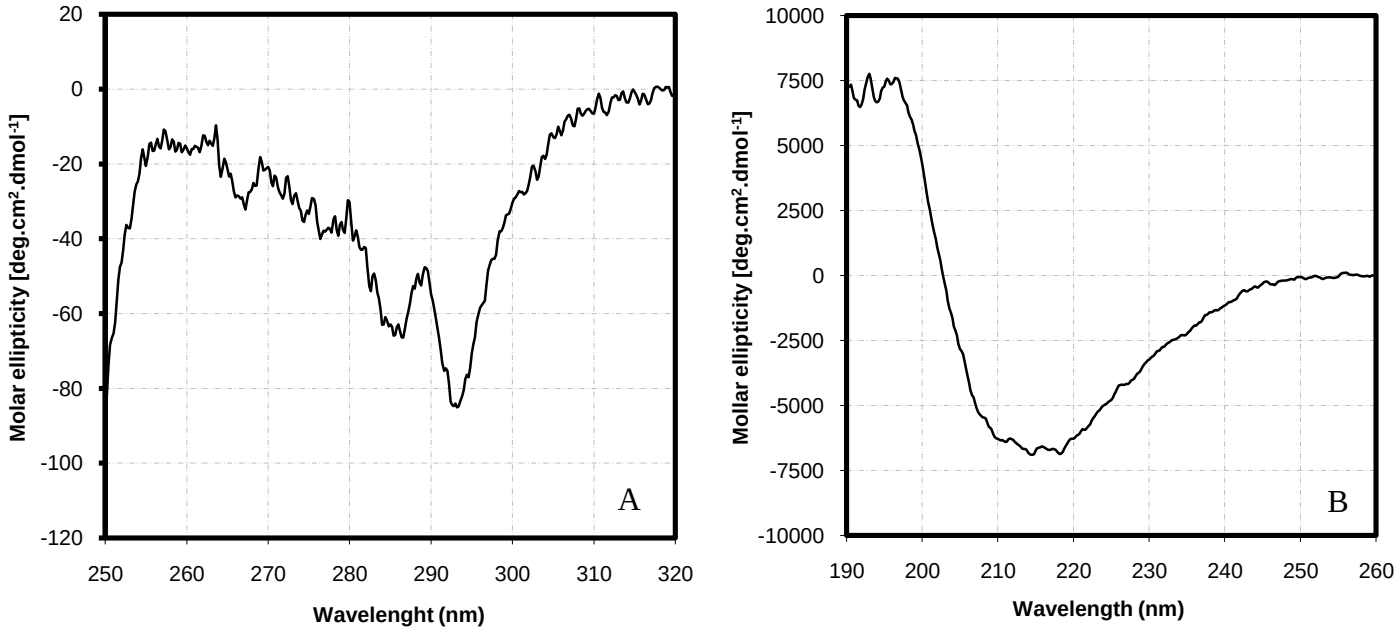


Figure 3.2 - Near (A) and far (B) UV CD spectra for aqueous solution of 1mg/ml of β-Lg, in 10 mM buffer (pH 7, 22°C). Spectra corrected to concentration.

Near UV spectrum (Figure 3.2 (A)) shows the characteristic peaks for β-Lg at 293 nm and 285 nm in this region. Peaks around 200-220 nm in the far UV CD spectrum are characteristic for β-sheet structures, while peaks at 208 and 222 nm indicate α-helix. In Figure 3.2 (B) one can clearly see the peak between 210 and 220 nm indicating the major content of β-sheet in the native state of β-Lg.

Native spectra using ATR-IR technique were recording for different concentrations of β-Lg (10 and 40 mg/ml) in 10 mM buffer by preparing the corresponding buffer in D_2O to avoid problems caused by the presence of water (section 2.2.4). Background spectra were obtained by measuring phosphate buffer 10 mM prepared in D_2O (pH 7) applying single beam mode and with the same parameters used when absorbance spectra were obtained. Figure 3.3 shows these ATR-IR spectra for the native state of β-Lg.

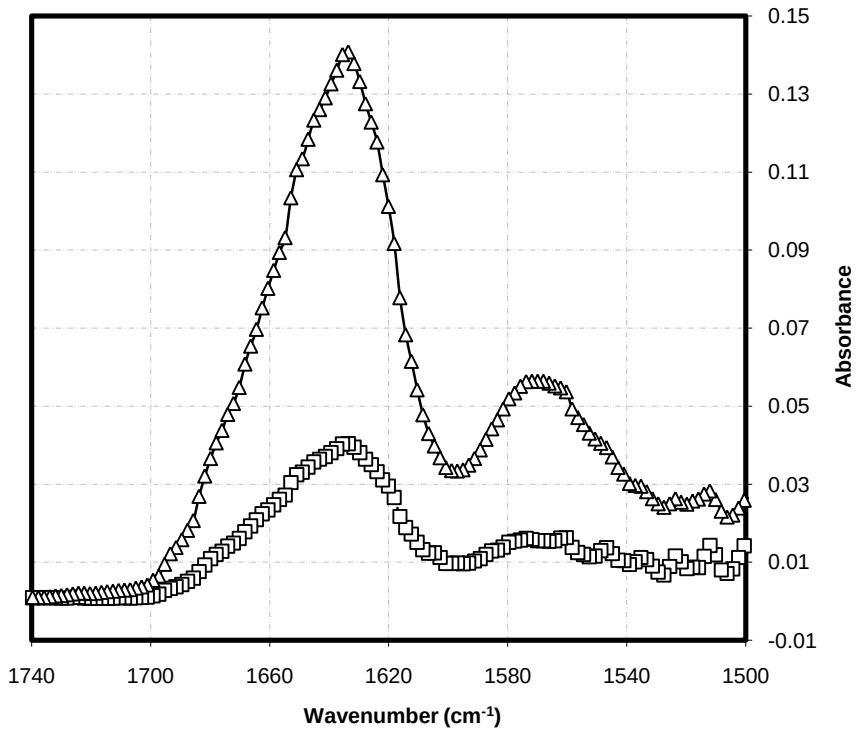


Figure 3.3 - ATR-IR spectra for aqueous solutions of β-Lg in 10 mM buffer prepared in D_2O (pH 7, 22ºC) for 10 mg/ml (squares) and 40 mg/ml (triangles) of protein. Resolution: 4. Coads: 256. Gain: 27.

In Figure 3.3 Amide I (1600 – 1700 cm^{-1}) and Amide II (~1510 and 1580 cm^{-1}) can be seen. The peak around 1635 cm^{-1} is attributed to intra-molecular β-sheet structure, indicating once again the majority content of this type of secondary structure in the native protein (Bhattacharjee et al., 2005). The absence of a clear peak around 1680-1691cm^{-1} indicates that intermolecular β-sheets are negligible in native β-Lg as expected (Shakuno et al., 2008). This experimental technique has been reported to be suitable for the determination of β-sheet content but not for the determination of α-helix content of the protein (Byler and Susi, 1986). However, IR absorbance at 1652-1658 cm^{-1} is attributed to α-helix content in the secondary structure of the protein. Measured absorbance intensities at the same wavenumbers for different concentrations of protein indicate that absorbance is proportional to protein concentration. The addition of a low dielectric constant solvent such as ethanol will

shift the characteristic IR peaks of β-Lg, as it will be shown later on in this thesis. Deconvolution of the IR spectrum obtained for 40 mg/ml of β-Lg in aqueous solution using GRAMS software gives the percentages for α-helix, intra-molecular β-sheets, and inter-molecular β-sheets, shown in Table 3.2.

Table 3.2 - Percentages of secondary structure obtained by deconvolution of ATR-IR spectrum for a sample of β-Lg (40 mg/ml) in 10 mM buffer prepared in D_2O (pH 7, 22ºC). Spectra collected in Figure 3.3.

	% of α-helix	**% of intra-molecular β-sheet**	**% of inter-molecular β-sheet**
40 mg/ml of β-Lg in phosphate buffer (10 mM)	17	46	<< 1

These results are in reasonable agreement with the percentages of α-helix and β-sheets (intra-molecular) obtained by deconvolution of far CD spectrum of β-Lg mentioned above: 14 % of α-helix and 37 % of β-sheet.

Up to this point, spectroscopic results for native samples of β-Lg have been presented. It can be concluded that these results agree with previous studies for the native state of β-Lg under similar conditions. Following, results of dynamic light scattering experiments done for aqueous solutions of β-Lg will be presented.

As it was reported previously in this thesis, native β-Lg appears mainly as a dimer when dissolved at neutral pH in aqueous solution (Gimel et al., 1993; Majhi et al., 2006). Aqueous solutions of β-Lg in distilled water were prepared and analysed by dynamic light scattering (DLS) for different concentrations of protein. Due to the fact that samples with ethanol were designed to have a final concentration of phosphate buffer 10 mM, native samples with the same concentration of buffer were also analysed with DLS to see the effect of the presence of salts as compared to results

when β-Lg is dissolved solely in distilled water. When the general diagram of the experimental plan was followed (Figure 2.11) native samples with a final concentration of buffer 20 mM were also analysed. Samples were prepared as it was explained in section 2.3.1. The apparent hydrodynamic radii were calculated as described in section 2.2.5. Since every experiment was done in triplicate, three different values of apparent hydrodynamic radius were obtained for each protein concentration dissolved in each solvent. These data (R_{hi}) are shown in APPENDIX D. The final value of the apparent hydrodynamic radius for each protein concentration was calculated as the mean of the repetitions (see APPENDIX B). Final values for the apparent hydrodynamic radii of the protein in its native state are tabulated in Table 3.3, together with the standard deviation corresponding to different repetitions.

Table 3.3 – Apparent hydrodynamic radii (R_h) for aqueous solutions of the protein dissolved in distilled water, phosphate buffer 10 mM and phosphate buffer 20 mM (pH 7, 22°C in all cases) for each concentration of β-Lg.

R_h (nm)	**1 mg/ml**	**4 mg/ml**	**10 mg/ml**	**40 mg/ml**
Water	3.1 ± 1.1	2.3 ± 0.5	1.9 ± 0.4	1.8 ± 0.2
Buffer 10 mM	4.0 ± 0.9	3.1 ± 0.2	3.1 ± 0.2	2.4 ± 0.1
Buffer 20mM	15 ± 3.5	3.2 ± 0.2	3.3 ± 0.2	3.0 ± 0.2

Apparent hydrodynamic radii in Table 3.3 for different concentrations of protein in distilled water and different concentrations of phosphate buffer are also shown in Figure 3.4.

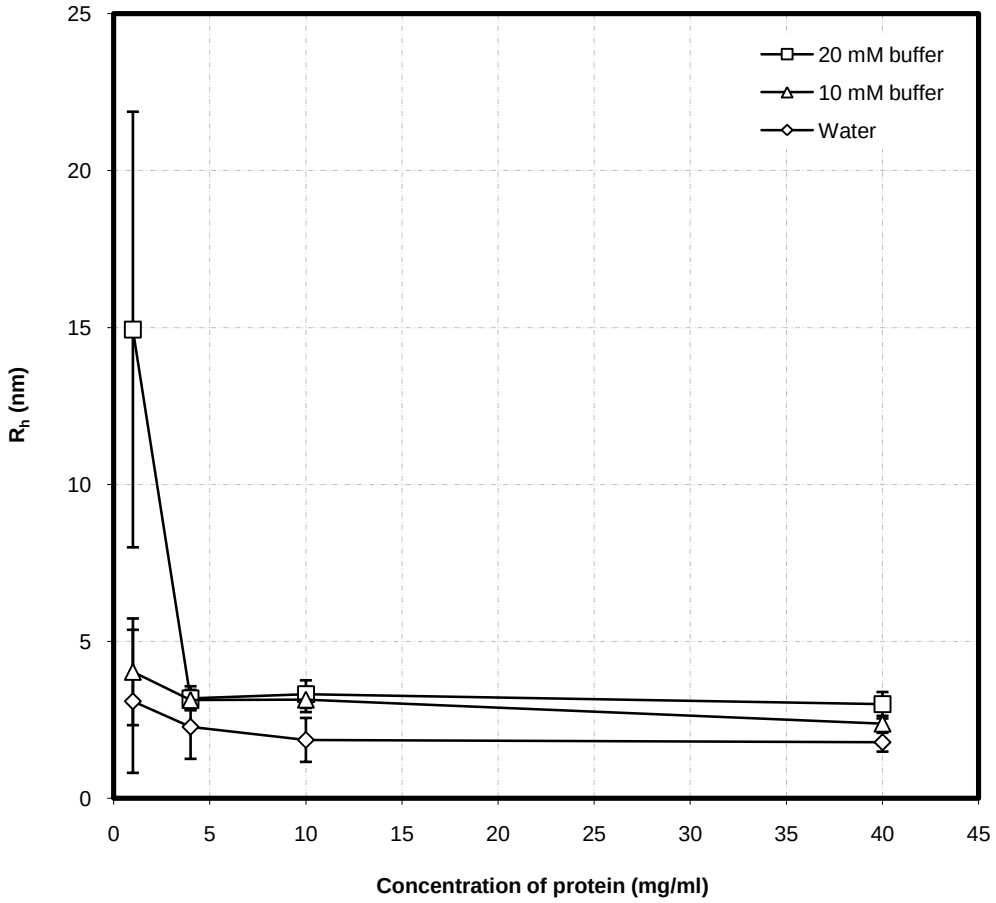


Figure 3.4 – Apparent hydrodynamic radii (R_h) vs. concentration of protein (mg/ml). β-Lg dissolved in 20 mM buffer (squares), in 10 mM buffer (triangles), and in distilled water (diamonds) (pH 7, 22°C in all cases).

Calculated hydrodynamic radii (R_h) for β-Lg solutions prepared in distilled water at pH 7, were 2.3 ± 0.5 nm, 1.9 ± 0.4 nm and 1.8 ± 0.2 nm for 4, 10, and 40 mg/ml respectively. These values are in good agreement with values of R_h reported previously for samples of β-Lg under similar conditions but calculated by applying different methodology (Kyung and Damodaran, 1991). On the other hand, these results seem to be slightly lower than the size for a native β-Lg dimer reported for dynamic light scattering measurements (approximately 3 nm) (Gimel et al., 1993; Aymard, Durand and Nicolai, 1996; Le Bon, Nicolai, Kuil and Hollander, 1999;

Takata, Norisuye, Tanaka and Shibayama, 2000; Beretta, Chirico and Baldini, 2000; Parker, Noel, Brownsey, Laos and Ring, 2001) and seem to be closer to the monomer size. There could be repulsive interaction effect (decreasing apparent R_h) at higher protein concentration similar as seen in section 3.3. However it is not straightforward to make a direct comparison among the results for the hydrodynamic radius of β-Lg in aqueous solutions, since previous studies have been carried out at a variety of ionic strengths, pH values, and protein concentrations. A summary of previous results for the hydrodynamic radius of β-Lg in its native state reported in literature for various conditions is shown in Table 3.4.

Table 3.4 – Apparent hydrodynamic radius, R_h (nm), of aqueous solutions of β-Lg obtained by previous studies involving different physico-chemical conditions.

Author	Method	Solvent	pH	Protein concentration (mg/ml)	R_h (nm)
Kyung and Damodaran, (1991)	Column chromatographic method	Distilled water (Colum equilibrated with 20 mM phosphate buffer)	8	10 mg/ml	1.63
Aymard et al. (1996)	Dynamic light scattering	Water with 0.1 M of ammonium acetate	7	Not specified	2.86 ± 0.04
Parker et al. (2005)	Dynamic light scattering	Water with 0.03 M NaCl	7.9	6 mg/ml	3.2
Gimel et al. (1993)	Dynamic light scattering	Water with 0.1 M of ammonium acetate	7	5 mg/ml	2.9 ± 0.1
Takara et al. (2000)	Dynamic light scattering	Distilled water (no salts reported to be present)	7	17.6 mg/ml	4
Takara et al. (2000)	Dynamic light scattering	Distilled water (no salts reported to be present)	2	17.6 mg/ml	2
Beretta et al. (2000)	Dynamic light scattering	Phosphate buffer 100mM	7.8	30mg/ml	2.59
Le Bon et al. (1999)	Dynamic light scattering	Distilled water with 0.1 M ammonium acetate	7	Extrapolated to zero	2.7

For samples with 1 mg/ml of protein dissolved in distilled water, the hydrodynamic radius from the first decay was calculated to be 3.1 ± 1.1 nm. However, we noted that fitting of the autocorrelation functions obtained for 1 mg/ml of β-Lg in distilled water is complicated by an apparent poly-dispersity of these systems. This different behaviour of the lowest protein concentration systems will be a constant present when dynamic light scattering is applied under various solvent conditions, as it will be shown later.

The hydrodynamic radius value for 10 mg/ml of β-Lg in 20 mM buffer was 3.3 ± 0.2 nm, which corresponds well with the values reported previously in literature for a dimeric β-Lg (e.g. Parker et al., 2005). By decreasing the buffer concentration to 10 mM, the size decreased slightly to 3.1 ± 0.2 nm. For higher protein concentration (40 mg/ml) the hydrodynamic radius was 3.0 ± 0.2 nm and 2.4 ± 0.1 nm in 20 mM buffer and 10 mM buffer, respectively. For 4 mg/ml, protein size does not show a noticeable change when buffer concentration is decreased from 20 mM to 10 mM: hydrodynamic radius was calculated to be 3.2 ± 0.2 nm and 3.1 ± 0.2 nm in 20 mM and 10 mM buffer, respectively. For the lowest protein concentrations, 1 mg/ml, the hydrodynamic radius was found to be very much dependent on buffer concentration, which can act as a charge screener for the charged protein molecules. In 20 mM buffer, β-Lg has a hydrodynamic radius of 15 ± 3.5 nm. In the presence of 10 mM buffer the size decreased to 4.0 ± 0.9 nm for the same protein concentration (1 mg/ml).

As a conclusion from the DLS results presented up to this point, it can be said that the hydrodynamic radii calculated for higher protein concentrations (4, 10 and 40 mg/ml) in the same solvent (distilled water, phosphate buffer 10 mM or phosphate buffer 20 mM) are close to each other. However, when β-Lg is present at the lowest concentration studied (1 mg/ml) the hydrodynamic radius undergoes a drastic increase as compared to the hydrodynamic radius obtained for higher concentrations of protein under the same conditions. This increase is most noticeable when β-Lg is dissolved in 20 mM buffer.

Dynamic light scattering results revealed that the trend followed by protein concentration and apparent hydrodynamic radius is in opposite directions: hydrodynamic radius increases when decreases protein concentration. With the aim of further corroborate these results, a dilution experiment was done where a sample of 40 mg/ml of protein concentration in 20 mM buffer was prepared. Subsequent dilutions to concentrations of 10, 4, and 1 mg/ml of protein (keeping a constant concentration of buffer - 20 mM) were done and every samples was measured by DLS. Hydrodynamic radii obtained in this experiment are shown in Figure 3.5.

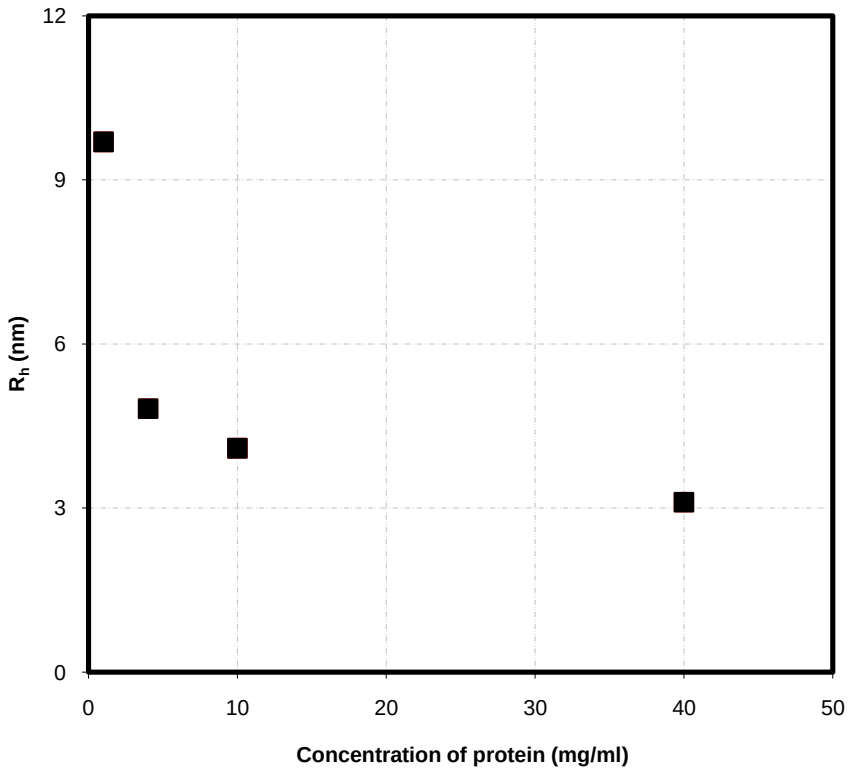


Figure 3.5 – Dilution experiment: hydrodynamic radius (R_h (nm)) vs. concentration of protein. Samples were diluted from 40 to 10, 4, and 1 mg/ml keeping a constant concentration of buffer of 20 mM (pH 7, 22ºC). See APPENDIX D for the autocorrelation functions obtained in this experiment.

Results in Figure 3.5 corroborate previous data obtained with individual β-Lg native samples in the presence of 20 mM buffer (square symbols in Figure 3.4). Therefore, this results give more evidence about the different behaviour that the lowest

concentrated samples (1 mg/ml of protein) shows in comparison with more concentrated samples, suggesting that the protein to buffer molar ratio appears to have a dramatic effect on the size of protein clusters**.** In fact, when a sample of 40 mg/ml in the same protein-molecules to buffer-molecules ratio as for samples with 1 mg/ml of protein in 20 mM buffer (for which phosphate buffer concentration has to be 0.4 M) was analysed, the calculated hydrodynamic radius was 4.0 ± 0.2 nm. This value is higher than any of the previous hydrodynamic radius calculated for a sample of 40 mg/ml of protein in distilled water, buffer 10 mM or buffer 20 mM indicating that the clusters size is affected by protein to buffer molar ratio. This hypothesis may be a possible explanation for DLS results obtained for the least concentrated samples (1 mg/ml), which seems to be already clustered/pre-aggregated in its native state. Another possible hypothesis to explain the unexpected big aggregates found for the least concentrated samples of β-Lg in aqueous solutions is that the commercial protein may contain some stabilising salt whose concentration becomes too low when the concentration of the protein is as low as 1 mg/ml. However, we were not able to corroborate this hypothesis within the scope of this project.

Apart of the unexpected big values of the hydrodynamic radius shown by DLS for 1 mg/ml of β-Lg, it was also observed that autocorrelation functions obtained for these samples did not show a shape consistent with a monodisperse particle size, but instead indicates that polydisperse population of aggregates is present in these solutions. For these reason, autocorrelation functions for 1 mg/ml of protein dissolved in distilled water, phosphate buffer 10 mM and phosphate buffer 20 mM were analyzed more deeply. In Figure 3.6 autocorrelation functions obtained for a native sample of 1 mg/ml of β-Lg in distilled water (A), in 10 mM buffer (B) and in 20mM buffer (C) are shown. Only one out of the three repetitions done is shown here as an example. The individual autocorrelation functions for each repetition done for 1 mg/ml systems can be seen in APPENDIX D. Figure 3.6 shows the normalised data obtained by DLS analysis together with the theoretical autocorrelation function calculated from the mean apparent hydrodynamic radius for each solvent shown in Table 3.3.

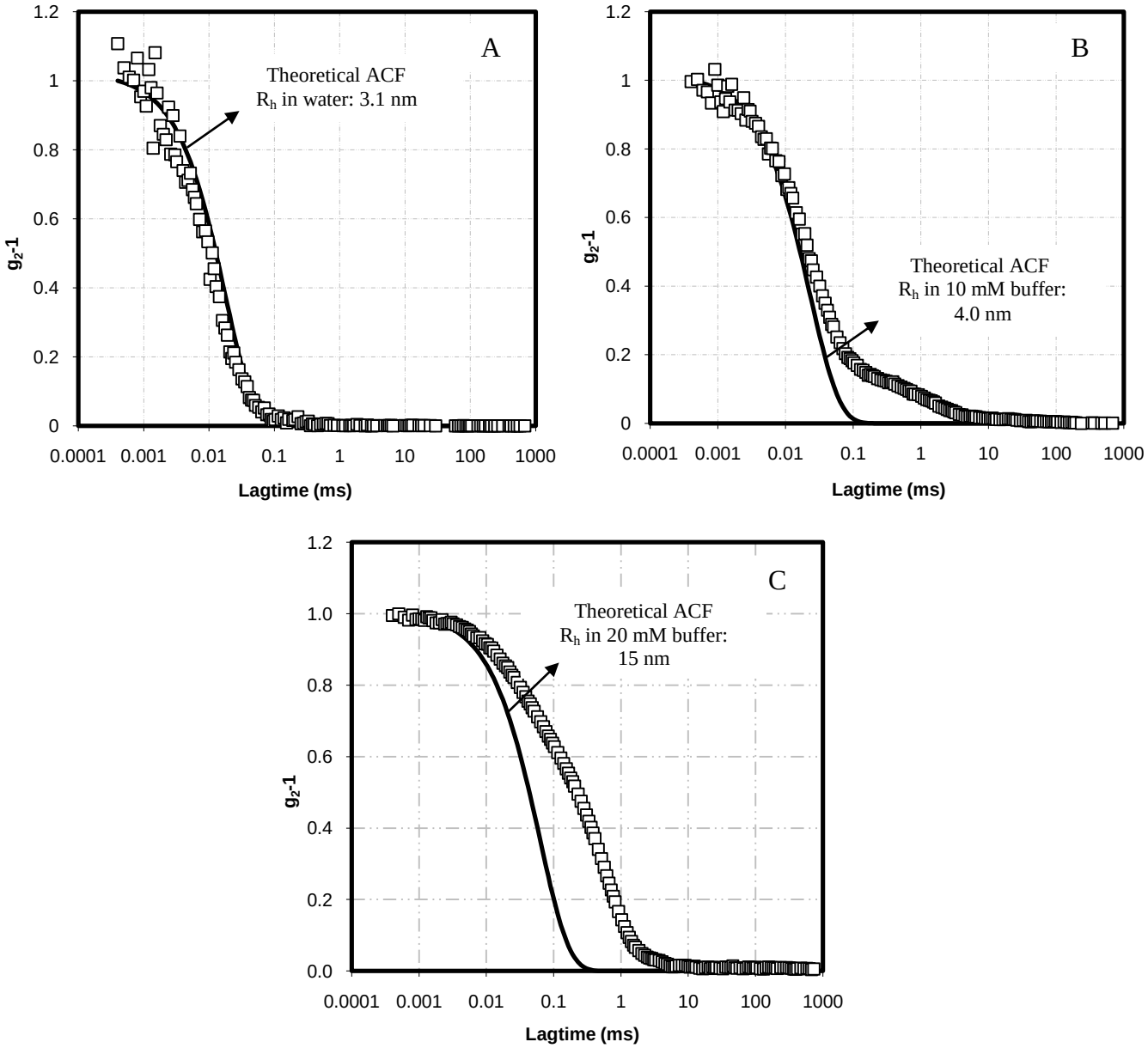


Figure 3.6 - Normalised autocorrelation functions obtained for a native sample of 1 mg/ml in distilled water (A), 10 mM buffer (B) and 20 mM buffer (C) (empty squares) - repetition A - and theoretical autocorrelation function (ACF) calculated from the mean apparent hydrodynamic radius obtained from every experiment in different solvents and shown in Table 3.3 (thick line). Whole set of data is shown in APPENDIX D.

Figure 3.6 reveals that the normalised autocorrelation functions obtained for 1 mg/ml system do not present a simple exponential decay as does the theoretical

autocorrelation function calculated for a monodisperse system with clusters of the same hydrodynamic radius (thick line in Figure 3.6).

By fitting the complete range of experimental points using the software provided with the DLS instrument (see section 2.2.5), the apparent particle size distribution was obtained (data not shown). It can be concluded that for the three solvents used, samples of 1 mg/ml present a polydisperse character, with small aggregates (with an average size of 3 nm obtained from the fitting of the first decay of the autocorrelation function) and bigger aggregates (whose average size depends on the solvent: 50 nm in distilled water and 200 nm for 10 mM and 20 mM buffer - although a small quantity of much bigger particles was also observed for the latest case (20 mM buffer) with size around 2 μm). Two major types of clusters are present in any case, but their proportion varies depending on the solvent. While the small clusters (corresponding to the first decay of the autocorrelation functions) are always dominant, the proportion of bigger clusters varies, being higher when buffer is present in the solution. Decays corresponding to various cluster sizes overlap with each other, resulting in a final autocorrelation function that cannot be accurately fitted by the cumulant method.

The polydisperse nature of this system is also evident from the tails presented in the autocorrelation functions for long decay times. Figure 3.7 shows the differences between the autocorrelation functions obtained from different repetitions done for the system consisting in aqueous solution of β-Lg with protein concentration 1 mg/ml (dissolved in distilled water (A), 10 mM buffer (B) and 20 mM buffer (C)), and the corresponding theoretical autocorrelation functions calculated from the mean apparent hydrodynamic radius obtained by the cumulant method shown in Table 3.3.

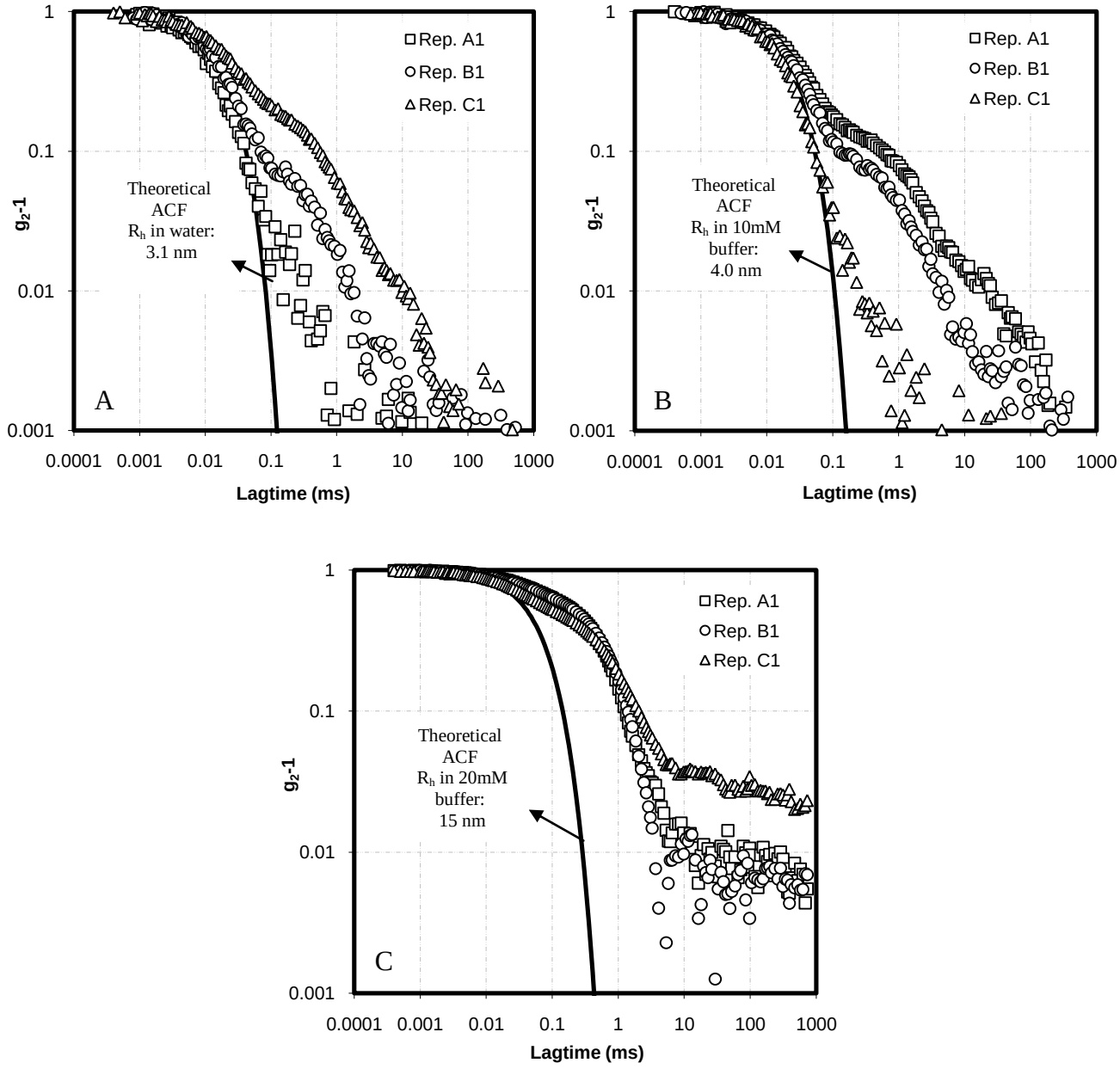


Figure 3.7 – Tails of the normalised autocorrelation functions obtained for 1 mg/ml system in distilled water (A), 10 mM buffer (B) and 20 mM buffer (C) for each repetition (A-C) together with the theoretical autocorrelation function (ACF) obtained from the mean hydrodynamic radius presented in Table 3.3 for each solvent (thick line). Whole set of data is shown in APPENDIX D.

The analysis of the last part of the autocorrelation function obtained by DLS has not been present in the literature up to this point; therefore, precise meaning of these tails remains still. For the propose of this project the analysis of the first decay was the one of interest, although a deeper insight in the whole autocorrelation function decay will be a subject of future studies.

The second conclusion that can be made from DLS results from aqueous solutions of β-Lg is the high influence that the present of different concentrations of ions in solution has on the size of protein clusters. It can be concluded that higher concentrations of ions leads to larger hydrodynamic radius for the same concentration of protein. Concentration of protein also plays a role in the clusters size, which decreases when increasing protein concentration. These results will be discussed in chapter 4 of this thesis.

The dimeric native form of β-Lg at neutral pH in aqueous solution will be disrupted with the present of a denaturant agent such as ethanol. The first hypothesis is that clusters size would depend in the quantity of ethanol that is added to the protein. In next section, results from DLS analysis in samples of β-Lg with different percentages of ethanol will be presented.

3.2. β-Lg SOLUTIONS WITH DIFFERENT PERCENTAGES OF ETHANOL

In this chapter, dynamic light scattering results for β-Lg solutions in the presence of different percentages of ethanol will be presented.

Solutions from 0 % (v/v) - aqueous solutions of β-Lg - and up to 50% (v/v) of ethanol were prepared both in the presence and absence of phosphate buffer 10 mM and dynamic light scattering measurements were applied right after preparation. Two series of experiment were done for different concentrations of protein: 10 and 40 mg/ml (for experimental procedure see section 2.3.2).

Several researchers have reported the formation of gels when β-Lg is in the presence of high percentages of ethanol and at high concentrations of protein at neutral pH and room temperature in the absence of buffer (Dufouer et al., 1994; Dufour et al., 1998; Renard et al., 2000). No gels were formed over a period of 10 days in any of the samples prepared for these experiments. Instead, transparent solutions were obtained in all cases, even for high concentrations of both protein and organic solvent. This may be due to the nature of the commercial protein itself, in terms of possible stabilization additives, which may have been different to that used in the mentioned studies.

The mean apparent hydrodynamic radii (R_h) obtained for samples with different percentages of ethanol in the absence and presence of phosphate buffer 10 mM for 10 and 40 mg/ml of protein are shown in Figure 3.8.

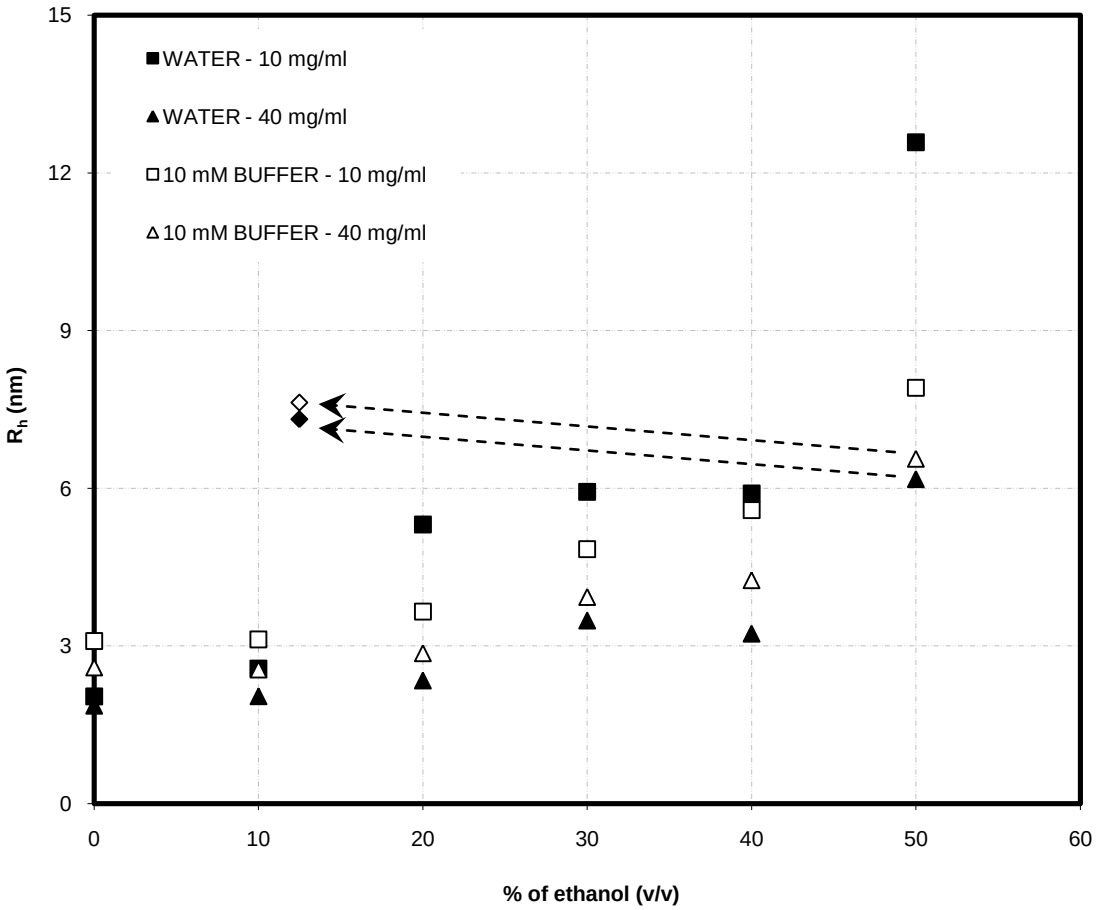


Figure 3.8 –Apparent hydrodynamic radius (nm) vs. concentration of ethanol (% by volume) for 10 mg/ml (squares) and 40 mg/ml (triangles) of protein concentration dissolved in distilled water (full symbols) and in 10 mM buffer (empty symbols) (pH 7, 22°C in all cases). The full and empty diamond correspond to the sample of 40 mg/ml with 50 % of ethanol (v/v) diluted to 10 mg/ml (concentration of ethanol 12.5 % v/v) in distilled water and 10 mM buffer respectively, as indicated by the arrows. See APPENDIX E for individual and average values of the hydrodynamic radius and autocorrelation functions.

Figure 3.8 shows an increase in the hydrodynamic radius as the concentration of ethanol is increased for both protein concentrations. Hydrodynamic radii for 0 % (v/v) of ethanol i.e., β-Lg natives samples, present similar values as the ones reported previously in section 3.1 for native samples of the protein dissolved both in distilled water and in 10 mM buffer and under the same conditions of pH and temperature (in

Figure 3.8 - 2.0 ± 0.2 nm and 1.9 ± 0.1 nm for 10 mg/ml and 40 mg/ml respectively in distilled water and 3.1 ± 0.2 nm and 2.6 ± 0.1 nm for 10 and 40 mg/ml of protein respectively in 10 mM buffer). Moreover, hydrodynamic radii obtained for 0% (v/v) of ethanol are smaller in distilled water than in 10 mM buffer for both concentrations of protein, according with results shown in Figure 3.4 in section 3.1.

By comparing values for both concentrations, the same trend as it was observed for native samples is followed here when ethanol is present in solution at every concentration: hydrodynamic radii are higher when protein concentration is less for both solvents. Hence, clusters sizes for samples of 10 mg/ml of protein are all higher than those for samples with 40 mg/ml of protein for the same percentage (by volume) of ethanol, both in distilled water and in 10 mM buffer.

When the sample of 40 mg/ml of protein and 50 % of ethanol (v/v) in distilled water (with hydrodynamic radius of 6.2 ± 1.1 nm) was diluted to a concentration of 10 mg/ml (and therefore the percentage of ethanol was reduced to 12.5% (v/v)) five minutes after preparation, the hydrodynamic radius of the protein did not decrease with dilution, but the measured value was 7.3 ± 1.3 nm (increase of 18%) indicating that the aggregation process is irreversible and clusters formed do not de-aggregate upon dilution (full diamond in Figure 3.8). The same trend was observed when the sample of 40 mg/ml of β-Lg in 50 % (v/v) ethanol and 10 mM buffer was diluted five minutes after preparation to 10 mg/ml (and 12.5 % (v/v) of ethanol): hydrodynamic radius of the diluted sample was calculated to be 16 % higher than the hydrodynamic radius of the initial sample (6.6 ± 1.0 nm and 7.6 ± 0.9 nm for the protein before and after dilution, respectively). This result suggests once again that the aggregation process is irreversible under the studied conditions. Deeper insight in the physical meaning of the increase of the apparent hydrodynamic radius with dilution can be found in chapter 4.

With the aim of completing the whole range of ethanol percentages from 0 to 90% (v/v) samples with 60, 70, 80, and 90 % (v/v) of ethanol were prepared under the

same conditions (10 mM buffer and pH 7 at room temperature) for samples with 10 mg/ml of β-Lg. This experiment had a qualitative objective: to identify the formation of gels as reported in literature. Since no gels were obtained at 50 % (v/v) of ethanol on the contrary to what other researchers did observe (Dufouer et al., 1994; Dufour et al., 1998; Renard et al., 2000), it was decided to increase ethanol concentration and find out if gels were formed under these new conditions. Only clear samples were obtained in 60% (v/v) of ethanol. For higher concentrations, samples became cloudy (in 70% (v/v) of ethanol) and even turned to be milky with precipitated clusters for higher concentrations of ethanol (80 and 90 % (v/v)). Standard dynamic light scattering cannot be applied in turbid samples, so it was no possible to measure hydrodynamic radii for samples with high concentrations of ethanol. It was found therefore that no gels were obtained in samples with higher concentrations of ethanol at least over 10 days of incubation at room temperature, the time scale fitted for kinetics analysis in this project.

The major changes in clusters size were observed when 50% (v/v) ethanol was added to the protein. Since samples could be analysed under these conditions with all the techniques reported in chapter 2, the main core of this study encompasses the study of structural changes and cluster formation undergone by β-Lg when high concentrations of ethanol (50% by volume) are added to the protein. Next section will present results obtained from these samples.

3.3. 50 % (v/v) BUFFER-ETHANOL β-Lg SOLUTIONS

In this chapter results obtained for the main system under study in this project are presented. Samples of β-Lg with different concentrations of protein, containing 50 % of ethanol (v/v) and 10 mM phosphate buffer at pH 7 and incubated at room temperature (22°C) for up to 10 days were analysed with multiple spectroscopic and scattering techniques described in section 2.2, with the aim of studying the aggregation behaviour of the protein focusing in the first stage (lag period) of the process, where formation of nucleated clusters is expected.

As a general introduction to the chapter, it can be said that β-Lg solvent-induced denaturation and subsequent aggregation was studied from three different points of view, each of them corresponding to a major set of results, by applying different techniques:

(i) Study of structural re-arrangements of the protein by applying ThT-binding fluorescence assay.
(ii) Study of changes in the secondary and tertiary structure of the protein taking place during the cluster-formation process by applying circular dichroism (CD) and attenuated total reflection infrared spectroscopy (ATR-IR).
(iii) Study of clusters size during the aggregation process by applying dynamic light scattering (DLS).

With the aim of presenting results in a properly organised way, this chapter will be divided in three major subsections corresponding with the three sets of results mentioned above.

3.3.1. β-Lg restructuring kinetics: ThT-binding fluorescence spectroscopy

The kinetics of β-Lg structure-transitions was monitored by applying fluorescence spectroscopy using Thioflavin T (ThT)-binding fluorescence assay. As it was mentioned in section 2.2, this dye has been widely used to determine the presence of amyloid fibrils in solution since it is thought to bind specifically the characteristic cross β-sheet structure present in this type of protein aggregates (Naiki et al., 1989; LeVine, 1999; Krebs et al., 2005). However, results in this study suggests that ThT can also bind aggregates with suitably arranged β-sheet structures without being necessary for the aggregates to have the mature and high ordered structures known as amyloid fibrils (Chiti and Dobson, 2006).

Four systems, corresponding to four different concentrations of β-Lg, were measured at different times once the denaturant solvent was added to the protein - just after adding ethanol (labelled as "Day 0" in results) and after 1, 3, 7 and 10 days of incubation at room temperature. The ThT emission fluorescence spectra for each protein concentration together with the spectrum corresponding to the native state of the protein (aqueous solution in 10 mM buffer) in the absence of ethanol for the same concentration of protein and buffer can be seen in Figure 3.9.

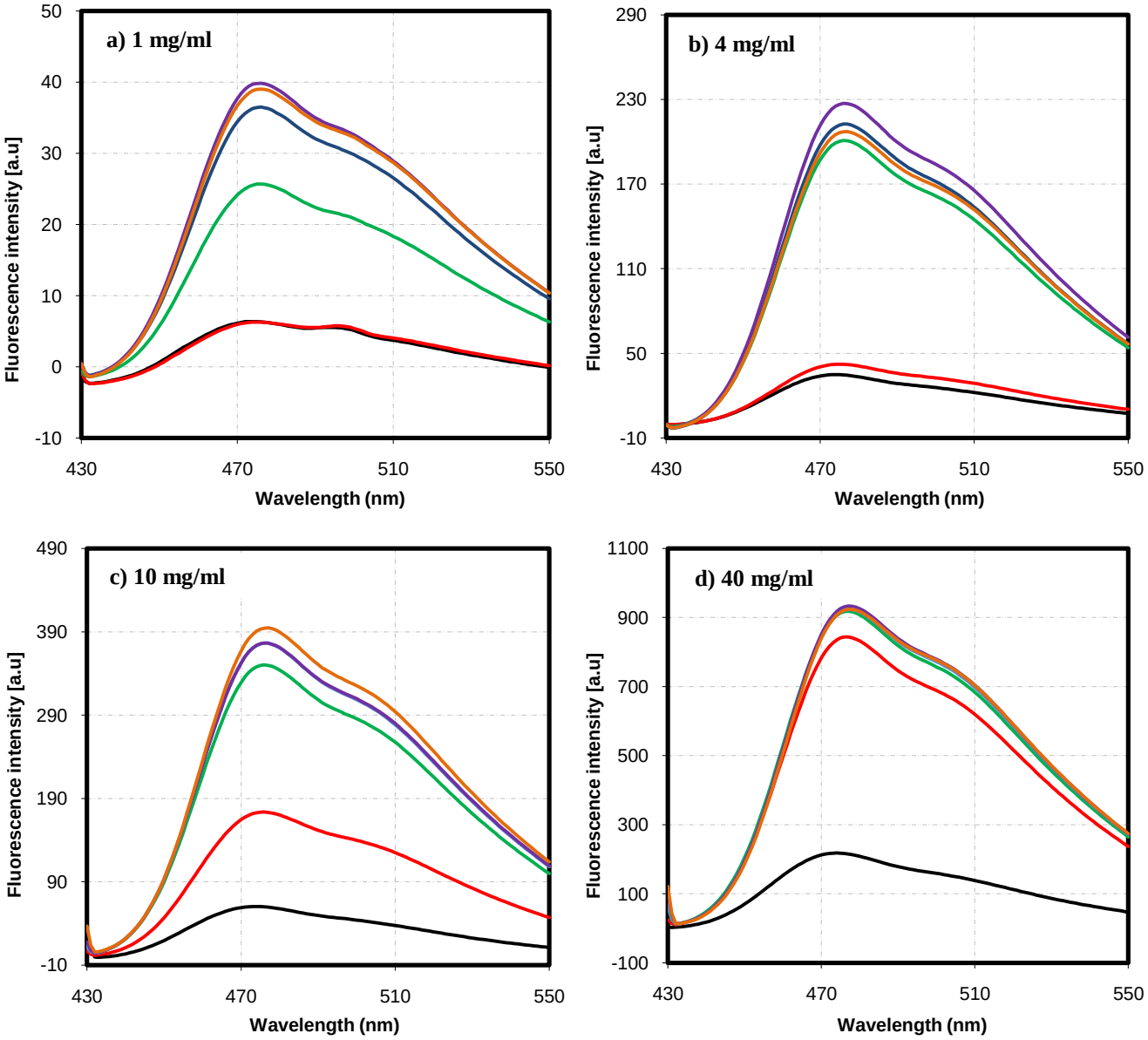


Figure 3.9 – ThT emission fluorescence spectra (λ_{exc}=425 nm) for different concentrations of β-Lg in the presence of 50% of ethanol (v/v) and 10 mM buffer (pH 7, 22°C). Samples were measured just after adding ethanol (Day 0 – red line) and after 1 (green line), 3 (blue line), 7 (purple line) and 10 (orange line) days of incubation at room temperature. The black line corresponds to the native protein. Data corresponding to repetition "A" for each protein concentration are shown. See APPENDIX C for complete fluorescence intensity data at 477 nm.

For the lowest protein concentration (1 mg/ml), no changes are observed in the fluorescence intensity of ThT emission spectrum after adding ethanol compared to the spectrum for the native protein at the same concentration. However, after one day of incubation at room temperature, an increase in the fluorescence intensity at 477 nm is observed: it is increased approximately 4 times with respect to the fluorescence intensity observed for the native sample of 1 mg/ml of β-Lg. A gradual increase is observed later during 10 days of incubation, being more noticeable between day 1 and day 3. The fluorescence intensity reaches its maximum value after 10 days of incubation having increased approximately 6 times compared to the native value, although the changes between day 3 and day 10 cannot be considered significant.

Samples with 4 mg/ml of protein follow a similar trend to the 1 mg/ml system. No significant changes were observed just after adding ethanol to the protein as it is shown by the fact that the native and day 0 ThT emission fluorescence spectra are nearly identical. After 1 day of incubation at room temperature, fluorescence at 477 nm increased about 5.5 times with respect to the native value and only smaller changes are observed subsequently up to 10 days after adding ethanol.

Systems of 10 and 40 mg/ml of protein concentration undergo a different behaviour to that observed for the lower concentrated systems, since rapid changes in the fluorescence spectra are clear just after ethanol addition.

For 10 mg/ml system, fluorescence intensity at 477 nm increases 3 times just after adding ethanol compared to that of the native protein. After 1 day of incubation at room temperature, fluorescence intensity reaches almost its maximum value, with a slight increase from day 3 to day 10, were it reaches a value 6.4 times higher than that of the native protein.

A drastic increase in fluorescence intensity was observed for the highest concentrated system (40 mg/ml of protein) shortly after addition of ethanol: fluorescence intensity increases 4 times regarding to the native protein intensity. However, it does not

change significantly within further 10 days of incubation at room temperature. The fluorescence intensity value after 10 days was 4.3 times higher than that of the native protein. If this intensity is compared with its value after adding ethanol (4 times the native value), it can be concluded that for the highest concentrated system aggregates with potential structures able to interact with the dye are formed immediately after adding the denaturant agent. Table 3.5 summarises the ThT-binding fluorescence intensities for every protein concentration at 477 nm, including the native protein, after adding ethanol and during the following 10 days of incubation at room temperature. Values in Table 3.5 correspond to the average of fluorescence intensities calculated from different repetitions for each protein concentration. Standard deviation corresponds to variations among these repetitions. Results show that the fluorescence intensity is approximately proportional to protein concentration even after 10 days of incubation at room temperature as it was observed for the protein in its native state (section 3.1). This can also be seen in Figure 3.10 where fluorescence intensity at 477nm is represented vs. protein concentration for the native state of the protein (aqueous solution in 10 mM buffer), shortly after adding ethanol (Day 0), and after 10 days of incubation at room temperature (fluorescence intensity values tabulated in Table 3.5).

Table 3.5 - Average ThT- binding fluorescence intensity [a.u] at 477 nm (emission wavelength) (λ_{exc} = 425 nm) for every concentration of protein, both in the native state and in the presence of 50 % (v/v) ethanol (10 mM buffer, pH 7, 22°C in all cases) after adding ethanol (day 0) and thereafter until 10 days of incubation at room temperature.

FLUORESCENCE INTENSITY [a.u] at 477 nm	**1 mg/ml**	**4 mg/ml**	**10 mg/ml**	**40 mg/ml**
Native	6.30 ± 0.5	34.5 ± 0.9	59.5 ± 0.9	216 ± 12
Day 0	6.25 ± 0.9	42.0 ± 2.5	173 ± 22	844 ± 29
Day 1	25.7 ± 3.5	201 ± 18	350 ± 18	918 ± 85
Day 3	36.4 ± 2.4	213 ± 15	376 ± 8.7	923 ± 43
Day 7	39.8 ± 3.4	227 ± 14	376 ± 24	933 ± 62
Day 10	39.0 ± 3.7	207 ± 9	395 ± 25	924 ± 39

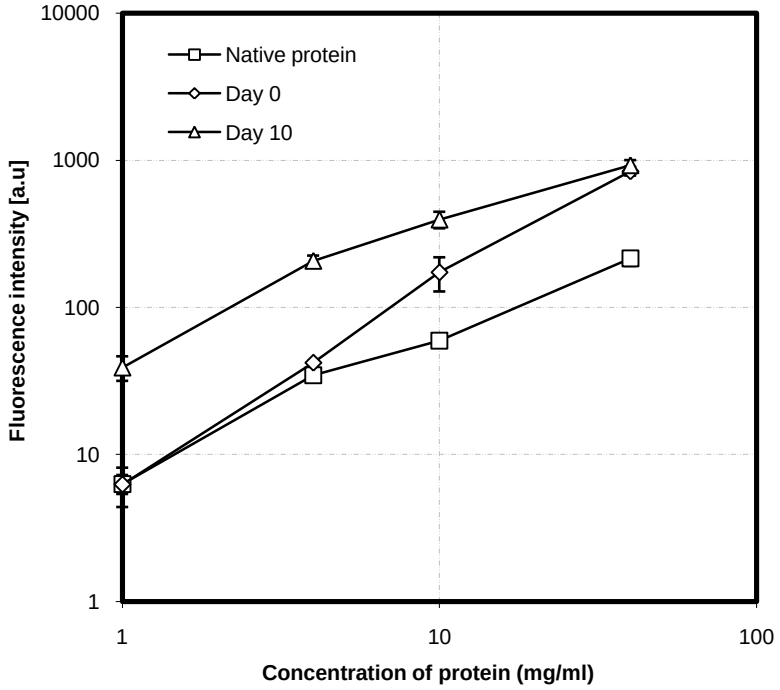


Figure 3.10 – ThT-binding Florescence intensity [a.u] vs. protein concentration (mg/ml) for native samples (aqueous solution in 10 mM buffer – squares), after adding ethanol (Day 0 – diamonds), and after 10 days of incubation at room temperature (triangles).

Fluorescence intensity spectra shown in Figure 3.9 were also normalised to concentration with the objective of finding further differences in the shape of fluorescence spectra between that for the native state of the protein, after adding ethanol (Day 0) and after 10 days of incubation at room temperature according to results in Figure 3.10. However, this did not reveal any significant information (see APPENDIX C).

Kinetics of structural re-arrangements undergone by the protein in the presence of ethanol can be monitored by plotting the fluorescence intensity values at 477 nm vs. incubation time. To have better information about the first steps taking place during cluster formation, additional measurements of ThT-binding fluorescence spectroscopy were also taken at smaller time steps. In Figure 3.11 the whole range of

data for ThT-binding fluorescence intensity vs. incubation time are shown for different concentrations of protein in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7), at room temperature.

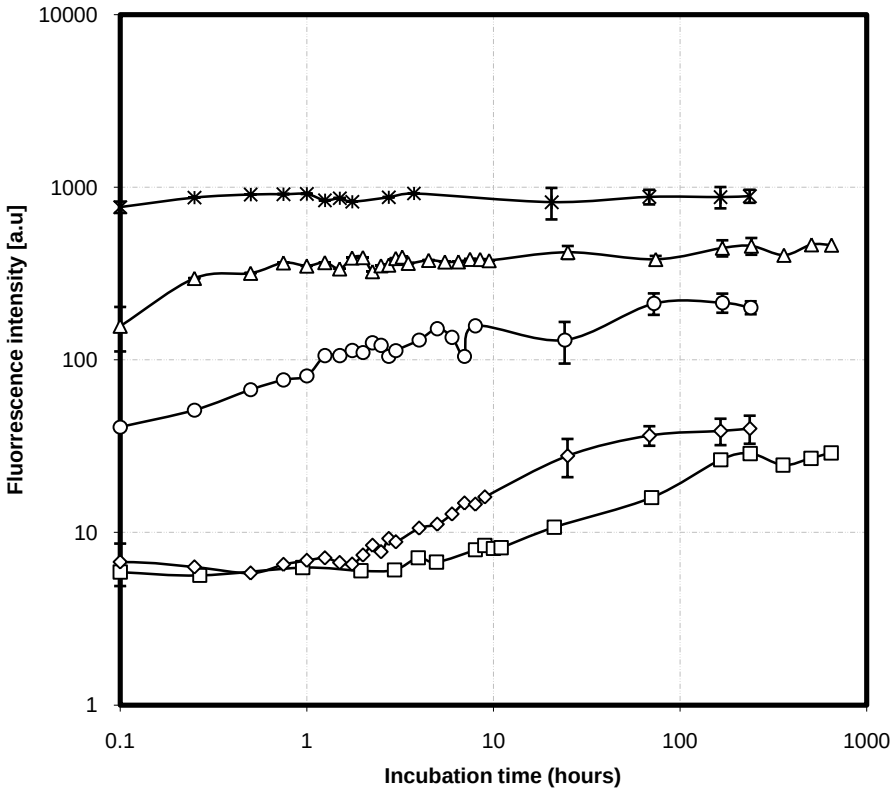


Figure 3.11 - β-Lg restructuring arrangements kinetics in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7, 22°C) for different concentration of protein: 0.5 mg/ml (squares); 1 mg/ml (diamonds); 4 mg/ml (circles); 10 mg/ml (triangles) and 40 mg/ml (stars). Points intercepting the vertical axis correspond to fluorescence intensities measured just after addition of ethanol (Day 0). Both axes are in log scale for clarity. For further details see APPENDIX F.

Figure 3.11 shows that for the least concentrated samples (0.5 mg/ml and 1 mg/ml of protein) the fluorescence intensity vs. the incubation time has a typical sigmoidal curve, composed of a lag period (counted from the moment ethanol is added to the protein), an exponential growth phase and finally a plateau or equilibrium phase. This kinetics is characteristics of a nucleation dependent growth mechanism, which

is also widely believed to be the path followed in amyloid fibril formation (Chiti and Dobson, 2006). The lowest concentration of protein (0.5 mg/ml) shows longer lag period and slower exponential growth than higher concentrated samples (1 mg/ml). The lag period for 0.5 mg/ml system was estimated to be 2.6 hours and the growth rate, 0.4 $hours^{-1}$. The plateau phase is achieved after 10 days of incubation at room temperature. For the system with 1 mg/ml of protein concentration, the lag period was estimated to be 1.6 hours and the growth rate, 1.5 $hours^{-1}$ (see section 2.2.2.3 for detailed calculation of lag period and growth rate from ThT-fluorescence data). The equilibrium phase is reached after 7 days of incubation at room temperature. These results are consistent with expectations for nucleation dependent growth mechanism, since lower concentrations of protein undergo aggregation with longer lag period (nucleation time) and slower subsequent exponential growth.

When the concentration of protein is further increased (4, 10 and 40 mg/ml) the typical shape of the sigmoidal curve with a lag phase cannot be observed anymore in our experiments, since fluorescence intensity rapidly increased within the first minutes after ethanol addition. However, we can see that 4 mg/ml system behaves in a transition way between lower and higher concentrated systems (10 and 40 mg/ml). A gradual increase in fluorescence intensity can be observed for each time measured after adding ethanol up to 72 hours of incubation at room temperature, when a plateau is reached and fluorescence intensity seems to not increase further for the system containing 4 mg/ml of β-Lg.

For protein concentration of 10 mg/ml, fluorescence intensity has also increased already after 15 minutes of ethanol addition and keeps increasing up to 45 minutes. After this time, it does not increase any more during the next 8.75 hours (525 minutes) of incubation at room temperature. The fluorescence intensity values obtained after 169 hours show a slightly increase as compared to the intensity between 45 minutes and 525 minutes of incubation. After 169 hours, fluorescence intensity does not increase any further during the rest of the incubation time of the experiment.

For 40 mg/ml system, the behaviour is similar to the samples with 10 mg/ml of protein although the increase in fluorescence intensity after 15 minutes of adding ethanol is only 1.1 times (while it is 2 times for the system with 10 mg/ml of protein). This fluorescence intensity for the highest concentrated samples does not increase any further during the next 10 days of incubation at room temperature.

By comparing results for 4, 10 and 40 mg/ml systems, it can be said that the common feature is the lack of a defined lag period as compared with lower concentrated systems (0.5 mg/ml and 1 mg/ml of β-Lg) were the lag phase is clearly shown, counted from the moment ethanol is added to the protein sample. It is worth to note here that a very short lag period (in the range of seconds or a few minutes) may exist for the high concentrated samples. However, due to the implicit restrictions of the experimental procedure of ThT-binding assay, fluorescence spectra could not be collected in shorter periods of time.

There are also differences in the behaviour of higher concentrated systems. As expected, the cluster formation process is slower when the protein concentration is lower: the plateau or equilibrium stage is achieved first for 40 mg/ml systems (after 15 minutes of incubation), followed by samples of 10 mg/ml of protein (after 45 minutes) and finally samples of 4 mg/ml of protein (after 72 hours). For the highest concentrated systems (10 and 40 mg/ml of β-Lg), the drastic increase in fluorescence immediately after adding ethanol regarding to the native fluorescence intensity (3 times for 10 mg/ml samples and 4 times for 40 mg/ml - see table 3.5) suggests an instantaneous formation of clusters containing suitable β-sheets structures able to bind the dye once the organic solvent is present in solution. However, for protein concentrations of 4 mg/ml it was found that after adding ethanol, fluorescence intensity does not change so rapidly respecting to the fluorescence intensity for the native protein as compared with higher protein concentrated system explained above. In this case, fluorescence intensity at 477 nm only increased 1.2 times regarding to the native sample after adding ethanol. This is the main difference between systems with 4 mg/ml of protein and higher concentrated ones (10 and 40 mg/ml). This

corroborates once again that less concentrated systems undergo a slower aggregation process and vice versa.

3.3.2. Structural changes in β-Lg: Circular dichroism and attenuated total reflection infrared spectroscopy.

Results regarding the structural changes undergone by the protein when it is in the presence of 50% (v/v) ethanol will be presented here. Two spectroscopic techniques were used with this aim: circular dichroism (CD), to monitor changes in the secondary and tertiary structure of the protein, and attenuated total reflection infrared spectroscopy (ATR-IR), to monitor changes solely in the secondary structure. In the following sections, results from both techniques will be presented separately.

3.3.2.1. Changes in the tertiary and secondary structure of the protein: circular dichroism (CD)

Circular dichroism (CD) enable the measurements of samples in minutes, being a fast and reliable technique providing information about the structure of the protein, both tertiary and secondary, depending on the region of the electromagnetic spectrum that is applied.

Near UV (260-320 nm) spectrum provides information about the tertiary structure of the protein due to the absorption of the aromatic amino acids. In this case, no quantitative analysis can be done, but near UV spectra provide a valuable fingerprint of the tertiary structure of the protein that can be used to identify, for example, the effect of adverse conditions in the protein environment on its folded (native) state as it is the case in this project.

Far UV (190-260 nm) spectrum contains information about the secondary structure of the protein due to the absorption of light by the peptide bond linking the amino acids. Different peaks at specific wavelengths are characteristic for α-helix or β-sheets. When a peak at around 220 nm is present in the far UV spectrum, it indicates a majority content of β-sheets in the protein, while if two peaks at 208 nm and 222

nm are present, then the secondary structure of the protein involves a majority content of α-helix (see Figure 2.4). Different algorithms have been developed to calculate the percentage of specific types of secondary structure (helix, sheets, turns, and random coil) as mentioned in section 2.2.3. Results presented in this section are based on the analysis of the far UV results using the CONTIN algorithm chosen as the analysis programme in the online server DICHROWEB. This analysis implies the correction of raw CD data to concentration (see APPENDIX A for more details about DICHROWEB).

Both far and near UV CD spectra were always corrected to concentration. Results shown in this section correspond to the corrected data; therefore far and near UV spectra are given in molar ellipticity ($deg \cdot cm^2 \cdot dmol^{-1}$) units. For this normalization, UV absorbance spectroscopy was used to obtain accurate values for the concentration of protein in solution inside the CD cells as detailed in section 2.2.1. APPENDIX G shows UV absorbance data obtained for each protein concentration.

3.3.2.1.1. Tertiary structure transitions: Near UV results.

Near UV spectra for β-Lg systems were recorded for a sample with 40 mg/ml of protein in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7) at room temperature immediately after ethanol addition (day 0) and after 1, 3, 7 and 10 days of preparation. With the aim of getting an acceptable near UV signal, concentration of protein in the quartz cell (path length 0.5 cm) was set to 1 mg/ml (see section 2.2.3.3); therefore, β-Lg sample of 40 mg/ml was diluted with buffer 10 mM/ethanol solution (50% v/v) before measurements. Near UV spectra was corrected to protein concentration once collected. Results are shown in Figure 3.12.

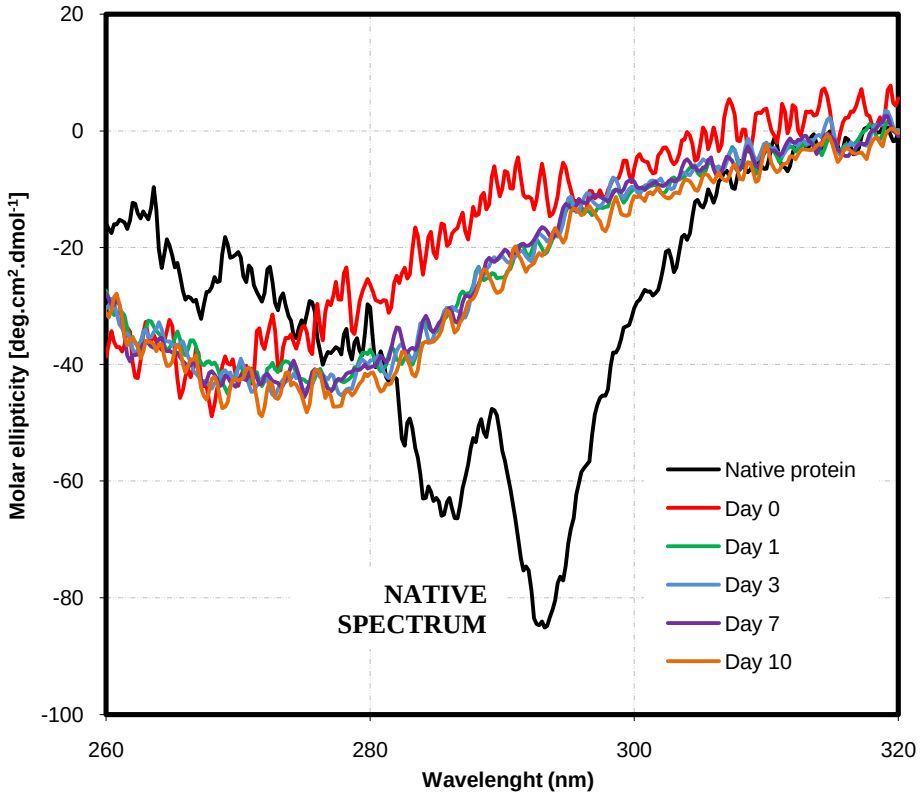


Figure 3.12 - Near UV spectra for a sample of 40 mg/ml (diluted to 1 mg/ml) of β-Lg in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7) incubated at 22°C. Spectra recorded after preparation (day 0) and after 1, 3, 7 and 10 days of preparation. Near UV spectrum of a native sample under the same conditions was also included for reference.

In Figure 3.12 it can be seen that the tertiary structure of the protein is completely lost already after ethanol addition (day 0): the native spectrum and any of the spectra recorded when ethanol is present in solution are very much different from each other. Near UV results obtained from other samples with different concentrations of protein (10 mg/ml – diluted to 1 mg/ml for CD measurements - and 1 mg/ml – without further dilution before measurements) were also investigated. Results were similar to those shown in Figure 3.12 indicating that no matter the concentration of the protein, its tertiary structure is completely destroyed once 50% (v/v) of ethanol is present in solution (for complete set of results see APPENDIX H).

With the aim of studying deeper the effect of the denaturant solvent in the tertiary structure of the protein, near UV spectra were recorded every 10 minutes for a sample of 1 mg/ml of β-Lg in the presence of 50% (v/v) ethanol and 10 mM buffer (pH 7, 22ºC). Results are shown in Figure 3.13.

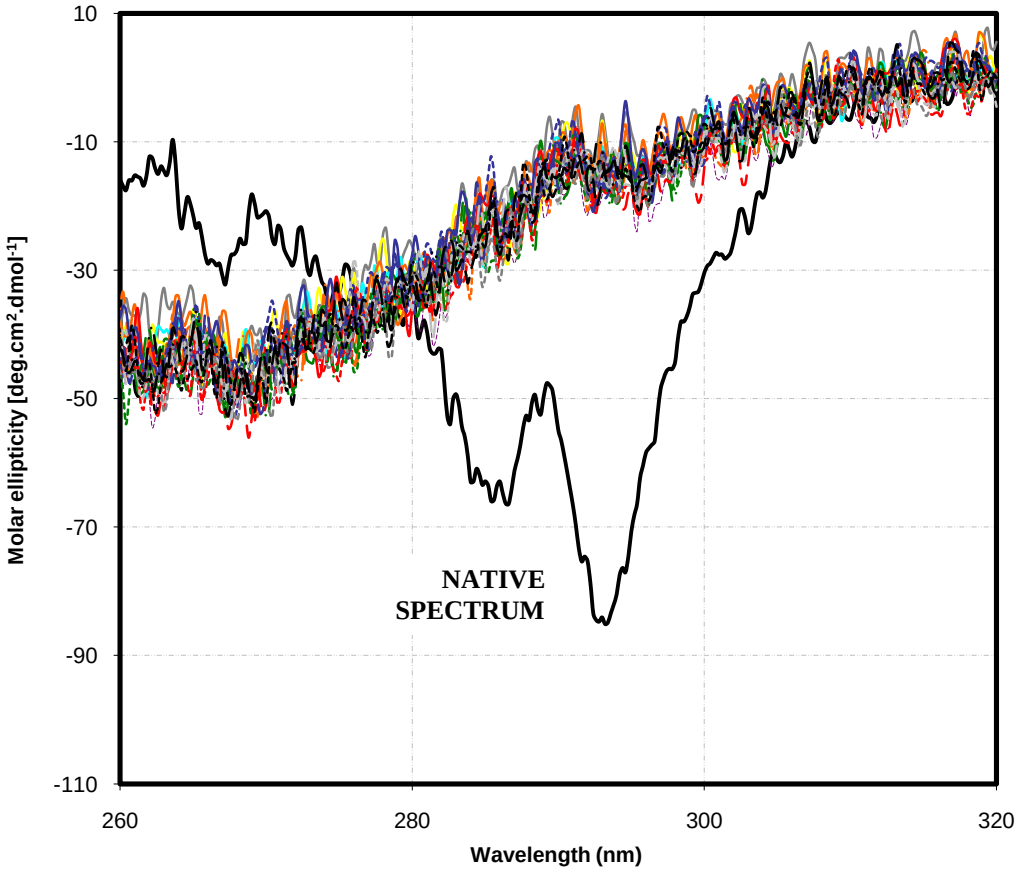


Figure 3.13 - Near UV spectra for a sample of 1 mg/ml of β-Lg in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7) incubated at 22°C. Spectra recorded every 10 minutes after preparation during 4 hours. Near UV spectra of a native sample under the same conditions was also included for reference.

Figure 3.13 clearly shows that the tertiary structure is lost immediately after the denaturant solvent is added to the protein sample without any observable transition between the native state and that characterized by the lack of a defined tertiary structure on the timescale of measurement (10 minutes). If the unfolded state of the

protein is considered to be that where the protein does not present both secondary and tertiary structures and therefore appears as a random coil in solution, it would not be accurate to say that the protein evolves from the native state to the unfolded state after ethanol addition, since it undergoes secondary structure transitions as shown by far UV results (section 3.3.2.1.2).

Near UV results suggest that no more detailed information can be provided from these spectra beyond the fact that the tertiary structure is lost immediately (or at least within the first ten minutes) after adding ethanol to β-Lg solutions. Therefore further focus was on far UV CD spectra that can also be analysed quantitatively.

3.3.2.1.2. Secondary structure transitions: Far UV results

Far UV spectra for β-Lg systems were recorded for several concentrations of protein (1, 10 and 40 mg/ml) in the presence of 50% (v/v) ethanol and 10 mM buffer (pH 7) at room temperature immediately after sample preparation (Day 0) and after 1, 3, 7 and 10 days of incubation. In order to acquire suitable far UV signals (as for near UV analysis), concentrations of 1 mg/ml of protein in the quartz cell (0.02 cm of path length for far UV) are required. Therefore, system of 10 and 40 mg/ml were diluted to 1 mg/ml of protein concentration before measurements. Every spectrum was corrected to concentration. In Figure 3.14 far UV results for every protein concentration (1, 10 and 40 mg/ml) are shown. Spectra for the native state of the protein (thick line) have been included as reference. In this case, because spectra are corrected to concentration, repetitions show very similar results for every concentration. In this chapter, only results for repetition A will be shown. The complete set of results for all repetitions can be found in APPENDIX H.

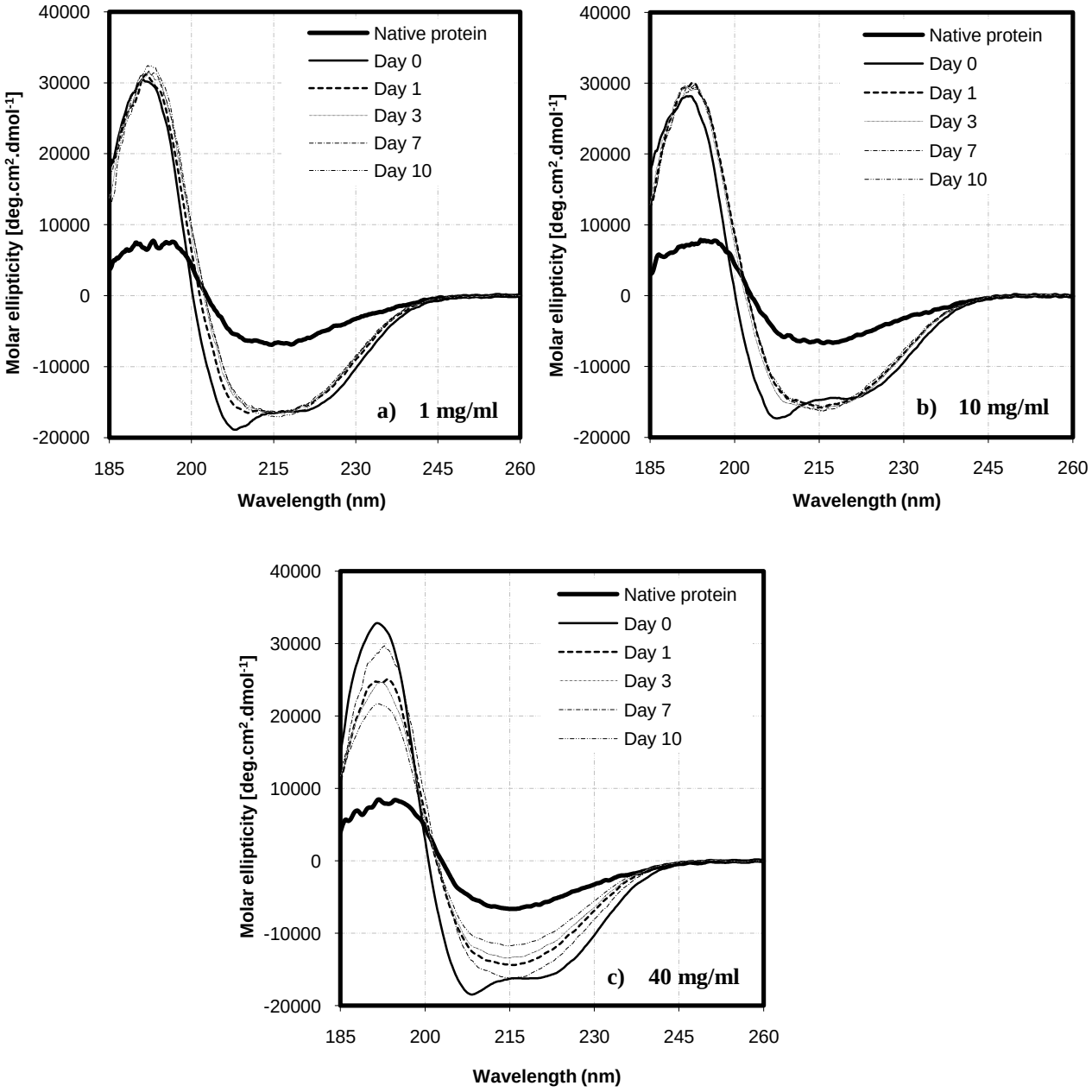


Figure 3.14 - Far UV CD spectra for samples of a) 1 mg/ml, b) 10 mg/ml and c) 40 mg/ml of β-Lg in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7) incubated at 22°C. Spectra recorded immediately after adding ethanol (Day 0) and after 1, 3, 7, and 10 days of incubation. Far UV spectra of the native protein under the same conditions were also included as reference (thick line). Data correspond to repetition A for each concentration. See APPENDIX H for complete set of data.

The far UV spectra for native β-Lg show a typical shape for a major content of β-sheets in the secondary structure of the protein as it was presented in section 3.1. Once the ethanol is added to the protein, far UV CD spectrum shifts to a shape characteristic for α-helix structures (peaks at 208 and 222 nm corroborate this) suggesting that there is a transition in the secondary structure of the protein from a majority content of β-sheets in its native state to a predominant α-helix structure once the ethanol is present in the protein solution. After 1 day of incubation at room temperature, far UV CD spectra have evolved to a shape more similar to that for the native protein, but with a more noticeable minimum at 216 nm. After that the spectra does not seem to change anymore up to 10 days of incubation. These results suggest the following secondary structural transitions in the protein in the presence of 50% (v/v) ethanol: from majority β-sheet content in its native state to predominant α-helix immediately after adding the denaturant solvent and back to an increase in β-sheet content within 1 day of incubation at room temperature (and subsequent decrease in α-helix). The formation of α-helix immediately after ethanol addition is in agreement with previous studies reporting the formation of this structure when organic solvents are added to the protein (Dufour et al., 1994; Dufour et al., 1998; Renard et al., 2000). In fact, thermodynamic studies in proteins indicate that the formation of α-helix in ethanol is not surprising, since in this way the protein is able to bury its peptide groups into the helix and leave the non-polar groups exposed to the solvent since they will have more affinity for the polar groups of ethanol (Pace, Treviño, Prabhakaran, and Scholtz, 2004). However, the evolution to a less content of α-helix after 24 hours suggests that this high quantity of α-helix after ethanol addition does not correspond to a stable state for the protein, which undergoes gradual secondary structural changes towards what seems a native-like secondary structure. In order to corroborate this hypothesis for the secondary structure transitions, a quantitative analysis of far UV spectra is required.

Percentages of the different types of secondary structures present in the protein once the ethanol is added to the protein solution and thereafter during the incubation time

at room temperature were calculated by using the online server DICHROWEB and CONTIN algorithm (see APPENDIX A for details). Table 3.6 shows the results obtained from the far UV spectra for each concentration as the average of the experiments, done in triplicate for each protein concentration.

Table 3.6 – Average percentage content of the different types of secondary structure of β-Lg induced by 50 % (v/v) of ethanol for every protein concentration just after addition of ethanol (Day 0) and after 1, 3, 7 and 10 days of incubation at room temperature (buffer 10 mM, pH 7). See APPENDIX H for data corresponding to individual repetitions.

Native Protein	**α-helix (%)**	**β-sheets (%)**	**Turn (%)**	**Random coil (%)**
	14 ± 1.1	37 ± 1.1	22 ± 0.3	27 ± 0.7
1 mg/ml of protein				
	α-helix (%)	**β-sheets (%)**	**Turn (%)**	**Random coil (%)**
Day 0	49.2 ± 2.8	7.15 ± 1.5	19.3 ± 1.5	24.4 ± 0.4
Day 1	42.1 ± 2.6	14.4 ± 0.9	17.4 ± 1.8	26.1 ± 0.4
Day 3	37.0 ± 1.5	17.6 ± 0.3	18.4 ± 1.3	27.0 ± 0.5
Day 7	35.7 ± 0.5	19.1 ± 0.4	18.1 ± 0.3	27.2 ± 0.2
Day 10	36.8 ± 1.9	18.6 ± 0.4	17.8 ± 1.2	26.8 ± 0.5
10 mg/ml of protein				
	α-helix (%)	**β-sheets (%)**	**Turn (%)**	**Random coil (%)**
Day 0	47.2 ± 0.5	8.83 ± 0.2	18.5 ± 0.2	25.2 ± 0.2
Day 1	33.7 ± 1.7	20.4 ± 1.1	18.5 ± 0.03	27.5 ± 0.6
Day 3	34.0 ± 1.7	18.9 ± 1.3	18.5 ± 0.3	27.6 ± 0.2
Day 7	34.2 ± 1.0	20.2 ± 0.2	18.6 ± 0.2	27.5 ± 0.3
Day 10	33.7 ± 1.0	20.1 ± 0.4	18.1 ± 0.2	27.6 ± 0.2
40 mg/ml of protein				
	α-helix (%)	**β-sheets (%)**	**Turn (%)**	**Random coil (%)**
Day 0	38.7 ± 2.3	10.8 ± 2.5	23.0 ± 3.1	28.6 ± 1.9
Day 1	34.6 ± 4.1	20.0 ± 1.8	17.6 ± 1.6	27.5 ± 0.7
Day 3	32.5 ± 2.0	20.7 ± 1.2	18.4 ± 0.7	28.1 ± 0.4
Day 7	29.0 ± 6.8	26.5 ± 6.3	19.2 ± 1.2	28.6 ± 1.0
Day 10	29.4 ± 2.1	22.9 ± 1.5	19.4 ± 0.6	28.4 ± 0.7

The kinetics of the structural changes for every type of secondary structure can be monitored by plotting the different percentages in Table 3.6 vs. incubation time. This is shown in Figure 3.15 together with the percentages for the secondary structure of the native protein: 13% of α-helix; 37% of β-sheet; 22% of turns and 28% of random coil (section 3.1).

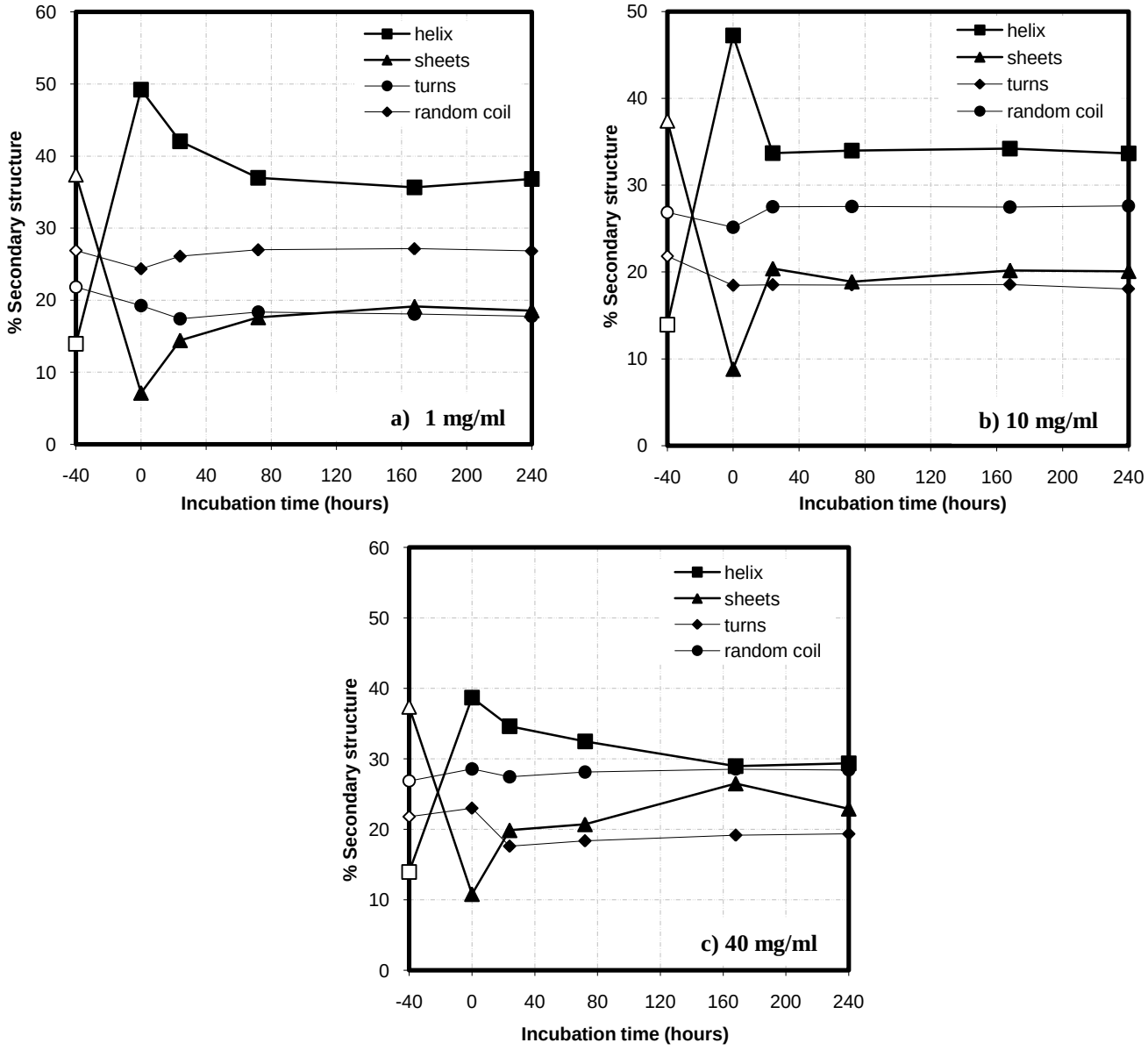


Figure 3.15 - Percentages of the different types of secondary structure vs. incubation time for each concentration of protein: a) 1 mg/ml; b) 10 mg/ml; c) 40 mg/ml in the presence of 50% (v/v) ethanol and buffer 10 mM (pH 7). Empty symbols, percentages for the native sample; full symbols, results just after adding ethanol and after 1, 3, 7, and 10 days of incubation at 22ºC.

Figure 3.15 suggests that kinetics of α-helix formation is practically instantaneous after ethanol is added to the protein. This is not a stable state for the protein, since the secondary structure evolves back subsequently towards more β-sheets. However, the native percentages of both α-helix and β-sheet secondary structures are not attained anymore. Instead, α-helix and β-sheets seems to reach equilibrium values higher and lower respectively, than in the native state irrespective of protein concentration. Therefore, the initial hypothesis based on qualitative analysis of the far UV CD spectra suggesting the evolution of the secondary structure of the protein to a native-like secondary structure was found to be false based on quantitative analysis of far UV CD spectra. These results suggest the formation of a new intermediate structure. After the first few days of incubation, α-helix and β-sheets content do not undergo any further changes, indicating that the formation of this intermediate is completed within the first days of incubation at room temperature. However, if previous fluorescence spectroscopy results showing the different behaviour for different concentrated systems are taking into account, it would be expected that the formation of this intermediate is faster for systems that are more concentrated. This hypothesis cannot be corroborated with far UV results up to this point, since the protein concentration in far UV cells is very similar (always around 1 mg/ml) due to the requirements to get a good far UV signal and hence any reversible rearrangements upon dilution cannot be excluded. We also note that the two other types of protein secondary structures (β-turns and random coil) remain mostly unchanged throughout the experimental time.

α-helix and β-sheet percentages for every concentration of protein obtained during the 10 days of incubation at room temperature and presented in Figure 3.15 were divided to obtain the α to β (α/β) ratios shown in Figure 3.16. Since incubation time has to be in logarithm scale for clarity, it has to be taken into account that incubation times corresponding to the aqueous solution of the protein and just after addition of ethanol are not actual incubation times but solely a licence to allow graphical representation.

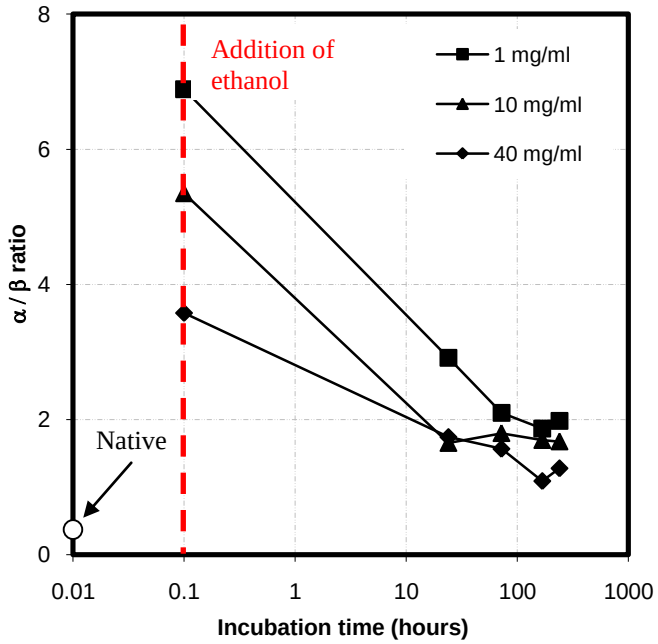


Figure 3.16 - α–helix to β–sheet percentages ratios for 1 mg/ml, 10 mg/ml, and 40 mg/ml of protein concentration in the presence of 50% (v/v) ethanol and 10 mM buffer (pH 7) incubated and measured during 10 days at 22ºC. Empty symbol corresponds to the native protein.

In order to investigate the early stage kinetics of protein structural transition a sample of 1 mg/ml was monitored over more than 5 hours after adding ethanol. Far UV spectra are shown in Figure 3.17. A total of 34 spectra (including the native protein and the measurement immediately after ethanol addition) were obtained, recorded every 10 minutes. Figure 3.17 shows the early far UV spectra obtained from 10 to 100 minutes of incubation at room temperature. The whole set of data is collected in APPENDIX H.

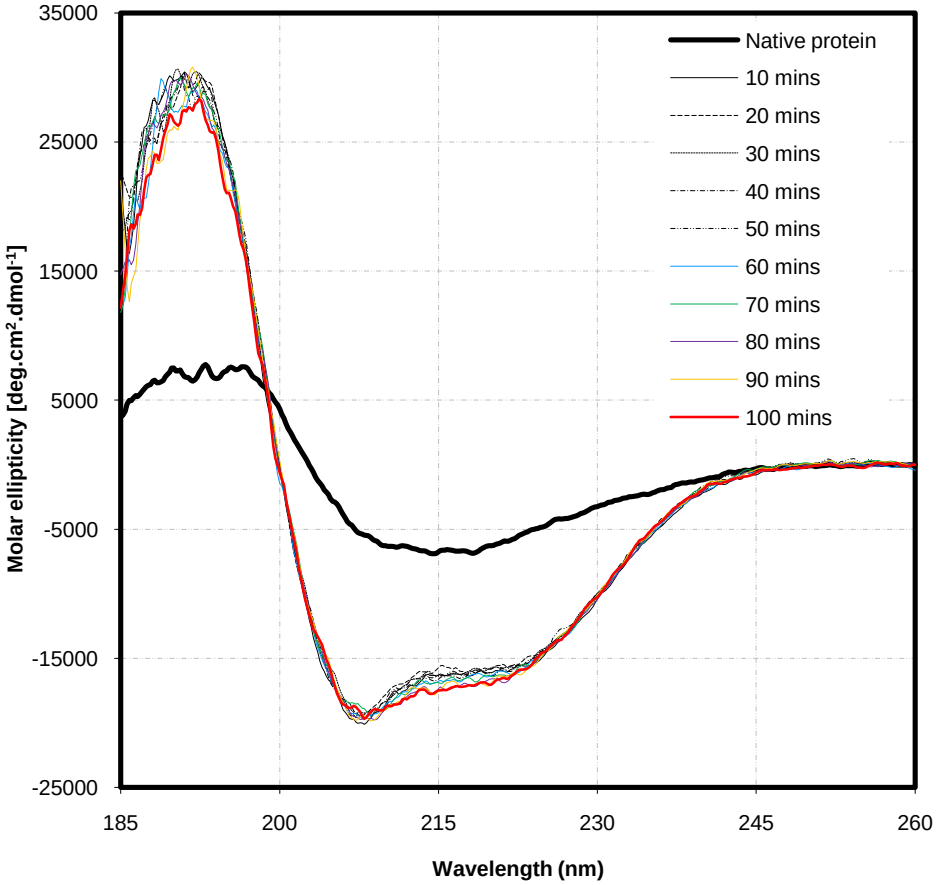


Figure 3.17 - Far UV spectra for a sample of 1 mg/ml of protein in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7, 22ºC) recorded just after adding ethanol and every 10 minutes thereafter up to 100 minutes after ethanol addition.

A quantitative analysis of each spectrum recorded during the first 5 hours of incubation after adding ethanol was done using DICHROWEB. The different percentages of α-helix and β-sheets obtained during the first two hours of incubation and previous results for the 10 days systems are shown in Figure 3.18, including the percentages of α-helix and β-sheets reported for the native protein. However, spectra obtained between 2 and 5 hours did not lead to reliable results upon fitting (NRMSD > 0.1). Incubation time is plotted in logarithmic scale for clarity because of which incubation time values for the native protein (0.01 hours) and just after adding ethanol (0.1 hours) are not actual times but a chosen representation to allow graphical presentation. APPENDIX H shows all these data tabulated.

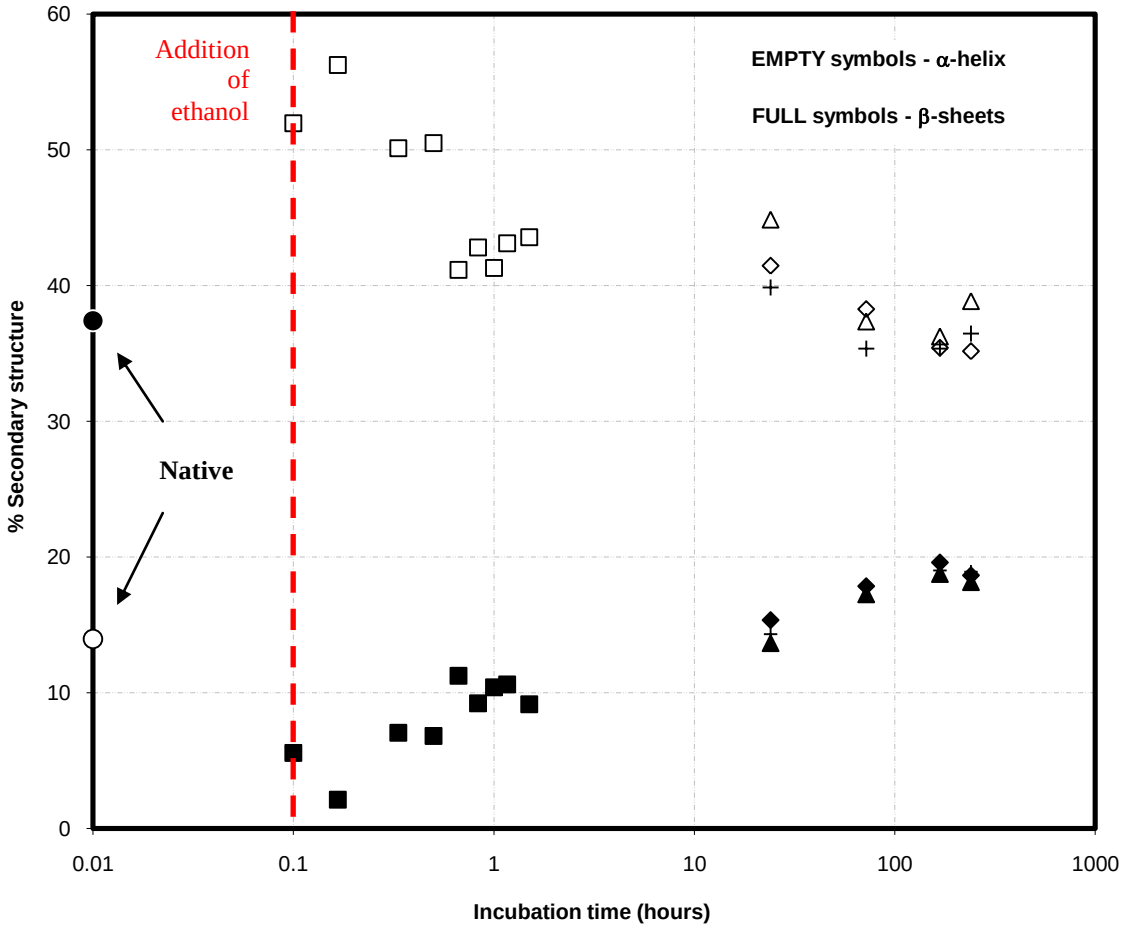


Figure 3.18 - α-helix (empty symbols) and β-sheets (full symbols) percentages for a sample of 1 mg/ml of protein in the presence of 50% (v/v) ethanol and 10 mM buffer (pH 7, 22ºC) measured after adding ethanol (squares), every 10 minutes thereafter during 2 hours (squares) and later after 1 day and up to 10 days of incubation at room temperature for different repetitions (triangles, diamonds and cross symbols correspond to measurements after 1, 3, 7 and 10 days of incubation at 22ºC). Circles in the Y axis represent percentages of both secondary structures reported for the native sample.

Figure 3.18 indicates that α-helix increases and β-sheet decreases drastically immediately after ethanol is added to the protein. Both structures show a gradual change afterwards (α-helix decreases while β-sheet increases) up to 10 days of incubation at room temperature (see also table H.5 in APPENDIX H). After about 3 days, α-helix and β-sheet structures seem to reach their constant values. From the quantitative analysis shown in Figure 3.18, it cannot be concluded that at some point the percentage of α-helix becomes less than the percentage of β-sheets since fitting

did not lead to reliable results between 2 and 5 hours. Therefore, it cannot be stated that there is a second transition between α-helix and β-sheets as hypothesised by qualitative analysis of the far UV spectra in Figure 3.17.

Far UV CD analysis gives very important information about the secondary structure of the protein allowing monitoring structural changes occurring during the incubation time. These changes can be analysed quantitatively to get the percentage content of the different types of secondary structures present in the protein at any time as shown above. However, these percentages indicate the total content of each structure within the protein but do not give any information about detailed nature of those structures for the case of β-sheets, i.e., far UV CD analysis does not allow knowing if the β-sheets structure present in solution correspond to inter-molecular (between different protein molecules) or intra-molecular (within one protein molecule) β-sheets. This information can be provided by infrared spectroscopy analysis. Results obtained with this spectroscopic technique are presented in the following section.

3.3.2.2. Changes in the secondary structure of the protein: attenuated total reflection infrared spectroscopy (ATR-IR)

Only systems with 10 and 40 mg/ml of protein could be analysed by attenuated total reflection infrared spectroscopy (ATR-IR) because systems with 1 mg/ml of protein were too dilute to get a reliable absorbance signal. As it was reported in section 2.2.4 this technique allows monitoring of changes in the secondary structure of the protein, as does CD, but it is more suitable for β-sheet secondary structure analysis since it is able to distinguish between intra- and inter-molecular β-sheets structures. Each secondary structure has a characteristic absorbance wavenumber as summarised in Table 2.2 (section 2.2.4.1).

As it was done with other spectroscopy techniques presented up to this point, every protein system was measured just after adding ethanol and thereafter for incubation times 1, 3, 7, and 10 days at room temperature. IR measurements at different times allow studying once again the kinetics of secondary structural changes undergone by

the protein. For IR measurements, samples needed to be done in D_2O to avoid the wide bands observed when water is present in solution. Figure 3.19 shows IR spectra obtained for samples with 10 and 40 mg/ml of β-Lg measured at different times over 10 days of incubation at room temperature including the spectra for the native protein as reference.

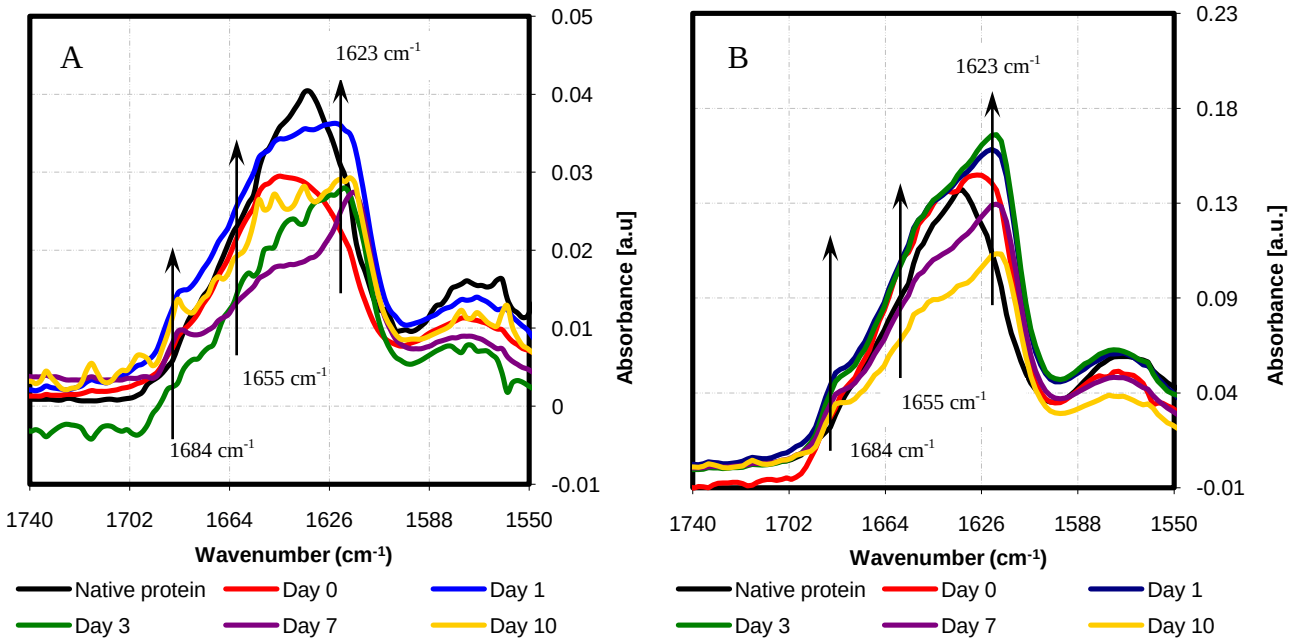


Figure 3.19 - ATR-IR spectra for samples of 10 mg/ml (A) and 40 mg/ml (B) of β-Lg in the presence of 50% ethanol (v/v) and 10 mM buffer (pH 7, 22°C), measured right after adding ethanol and after 1, 3, 7 and 10 days of incubation at room temperature. Native protein spectra (thick black line) included as reference. Background: buffer/ethanol 50% (v/v). Resolution: 4. Coads: 256. Gain: 27.

In Figure 3.19 it can be seen that the characteristic peak of intra-molecular β-sheet for native β-Lg in Amide I band at 1635 cm^{-1} is shifted to the right once the denaturant solvent is added to the protein for both protein concentrations to a value around 1626 cm^{-1}, suggesting the disruption of hydrogen bonds involved in β-sheet and the formation of new hydrogen bonds between β-strands and the solvent (Dufour et al., 1994). The increase in the peak at 1652 cm^{-1} indicates that the amount of α-

helix also increases from its content in the native protein when ethanol is added to the protein solution. This result is in agreement with far UV CD spectra, indicating formation of α-helix just after adding ethanol to the protein. The peak appearing at 1684 cm^{-1} is characteristic for the formation of inter-molecular β-sheet. This increase in inter-molecular β-sheet together with the formation of new hydrogen bonds between β-strands and the solvent (indicated by the peak at 1626 cm^{-1}) matches with fluorescence results for 10 mg/ml and 40 mg/ml of β-Lg immediately after addition of ethanol to the protein, where ThT-binding fluorescence intensity at 477 nm increases drastically from its value in the native protein, suggesting that ThT is able to bind the β-sheet (or β-sheet-like) structures present in solution. After 1 day of incubation at room temperature, the peak at 1626 cm^{-1} shifts to a value closer to 1622 cm^{-1}, characteristic of intra-molecular β-sheet structures.

To have a better insight of the secondary structure transition kinetics provided but ATR-IR measurements, IR spectra for systems with 10 mg/ml and 40 mg/ml of β-Lg were recorded just after addition of ethanol and every 10 minutes thereafter during 4 hours. Our colleague Dr. Nadeem Javid deconvoluted this data and results are shown in APPENDIX I.

Together with the structural transitions undergone by the protein under the influence of ethanol, the size of clusters formed under these conditions was also analysed. The last set of results presented in the following section will correspond with the study of clusters aforementioned by dynamic light scattering.

3.3.3. β-Lg aggregation: dynamic light scattering (DLS)

Study of protein clusters size evolution occurring during the aggregation process was studied by dynamic light scattering (DLS). Calculation of the apparent hydrodynamic radius of protein clusters present in solution was carried out as explained in section 2.2.5.

Protein systems with different concentrations of β-Lg (1, 10 and 40 mg/ml) in the presence of 50% ethanol (v/v) and 10 mM phosphate buffer (pH 7) were analysed right after adding ethanol and thereafter for incubation times 1, 3, 7 and 10 days at 22°C as it was done for other experimental techniques explained before. Experiments for each concentration were done in triplicate. The apparent hydrodynamic radii (R_h) presented here correspond to the average values obtained from the three repetitions done for each protein concentration together with its standard deviation unless otherwise stated. Relative deviations are not higher than 10%, except for 1 mg/ml, where the errors are in a range of 10-20 % for the native sample and after ethanol addition. This is in accordance with larger variability of results obtained for the least concentrated samples as it was already mentioned in section 3.1. The fitting of the autocorrelation function using the cumulant method leads to small fitting errors due to a very good quality of the fit. Table 3.7 summarises the mean apparent hydrodynamic radius (R_h) calculated for these three system (1, 10 and 40 mg/ml of β-Lg) measured over 10 days of incubation at room temperature. Values obtained for the native protein in the same concentration of buffer (10 mM) presented previously in section 3.1 are also shown here as reference. In APPENDIX J we show the individual values of the hydrodynamic radii (R_{hi}) calculated for every single experiment together with the fitting error. To have a better insight into the kinetic of cluster growth, apparent hydrodynamic radii tabulated in Table 3.7 were plotted vs. the incubation time at room temperature in Figure 3.20.

Table 3.7 – Apparent hydrodynamic radius (R_h, nm) calculated for each concentration of protein in the presence of 50% ethanol (v/v) and 10 mM buffer (pH 7) just after adding ethanol and after 1, 3, 7 and 10 days of incubation at 22°C. Native values presented in section 3.1 are also included for reference.

$R_h \pm \sigma_h$ (nm)	1 mg/ml	10 mg/ml	40 mg/ml
Native protein	4.0 ± 0.9	3.1 ± 0.2	2.4 ± 0.1
Day 0 – addition of ethanol	21 ± 2.7	7.7 ± 0.4	5.0 ± 0.4
Day 1	28 ± 1.7	8.1 ± 0.2	5.6 ± 0.1
Day 3	25 ± 1.0	8.2 ± 0.4	6.5 ± 0.1
Day 7	22 ± 1.7	8.1 ± 0.3	6.9 ± 0.2
Day 10	21 ± 1.7	8.2 ± 0.1	6.8 ± 0.2

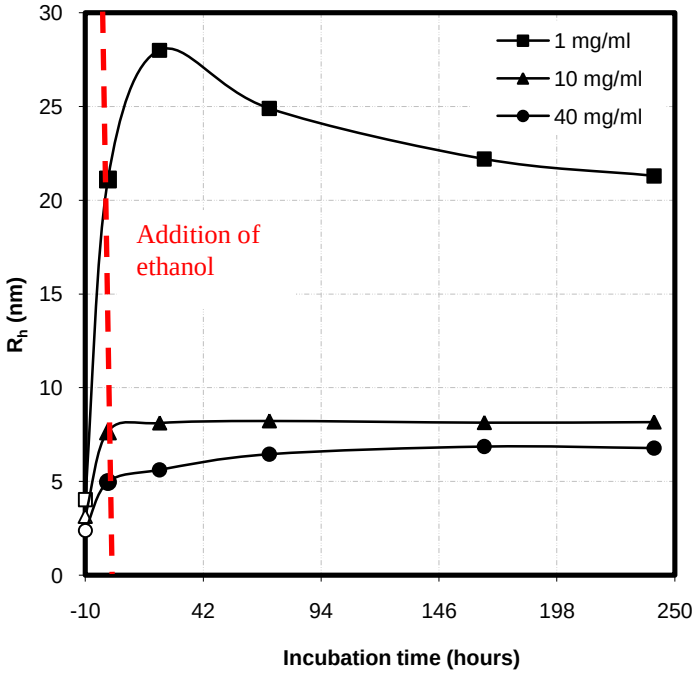


Figure 3.20 - Apparent hydrodynamic radius (R_h, nm) vs. incubation time for each protein concentration measured just after adding ethanol and after 1, 3, 7, and 10 days of incubation at 22°C. Empty symbols in the vertical axis represent values of R_h for the native protein shown in section 3.1. Errors bars can be seen in Table 3.7.

For the system with protein concentration 1 mg/ml relatively large aggregates grow immediately after adding ethanol to the protein, reaching a mean hydrodynamic radius 5.2 times larger than that for the native protein with the same concentration of protein and phosphate buffer. Clusters size keeps increasing after 1 day of incubation and starts decreasing slowly after the first day of incubation. This decreasing trend continues up to 10 days of incubation at room temperature as it can be observed in Figure 3.20 to a final value of 21 ± 1.7 nm that represents a size 5.3 times that of the native protein for the same concentration.

For higher concentrated systems (10 and 40 mg/ml) the growth of aggregates after the ethanol is added to the protein do not show such a drastic increase respective to the native size as it was observed for the 1 mg/ml system. Although for 10 mg/ml and 40 mg/ml systems the growth of aggregates seems to be fast, the increase in size is much smaller. For 10 mg/ml system the hydrodynamic radius increases 2.4 times after addition of ethanol respective to the native protein with the same concentration. After 1 day of incubation at room temperature, the hydrodynamic radius slightly increases to a value of 8.1 ± 0.2 nm, which remains almost unchanged during the 10 days of incubation at room temperature. The behaviour observed for the highest concentrated system (40 mg/ml of protein) is similar to 10 mg/ml system: when ethanol is added to the protein, hydrodynamic radius increases 2.1 times respective to the hydrodynamic radius of the protein in its native state with the same protein concentration. A gradual increase is observed afterwards up to 3 days of incubation, when the hydrodynamic radius reaches a value that remains unchanged up to 10 days of incubation at room temperature.

Figure 3.20 shows that clusters are bigger for lower protein concentrations at all times. This trend was also observed when β-Lg was dissolved solely in 10 mM buffer and ethanol was not present in solution (section 3.1). Indeed, if the hydrodynamic radii obtained for each concentration are plotted vs. protein concentration, this trend can be seen for the aqueous solution samples in 10 mM

buffer (the protein in its native state), just after adding 50% (v/v) of ethanol as well as after 10 days of incubation at room temperature (Figure 3.21).

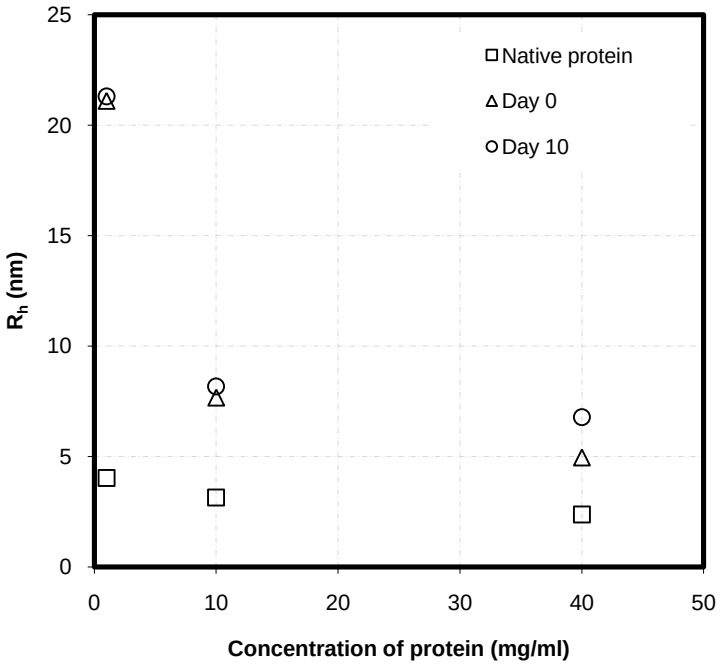


Figure 3.21 - Apparent hydrodynamic radius (R_h, nm) vs. protein concentration (1 mg/ml, 10 mg/ml and 40 mg/ml) for the protein in its native state (in 10 mM buffer) (squares), after adding 50% (v/v) of ethanol (day 0, triangles) and after 10 days of incubation at 22°C (circles).

Similarly to DLS results for the aqueous solutions of the protein (section 3.1), when ethanol is present in solution at high concentration (50 % v/v) large clusters were found when the concentration of protein was 1 mg/ml, suggesting that protein molecules were pre-aggregated before ethanol was added. On the other hand, the relatively small aggregates formed in 10 and 40 mg/ml systems when the denaturant solvent was added, seem to be stable and do not grow much further during the 10 days of incubation at room temperature, especially for 10 mg/ml system (Figure 3.20).

The least concentrated system, 1 mg/ml, shows a different behaviour from the higher concentrated systems not just in terms of the formation of larger clusters, but also in

the evolution of the mean hydrodynamic radius during the incubation time: hydrodynamic radius varies considerably with incubation time while for 10 and 40 mg/ml systems, clusters seem to attain a stable size shortly after adding ethanol which does not change significantly during the 10 days of incubation at room temperature (Figure 3.20). The less consistent DLS results for 1 mg/ml system can be better understood through analysis of the autocorrelation functions obtained for every system and comparison between them. Figure 3.22 shows the different normalised autocorrelation functions obtained for different protein concentrations (1, 10 and 40 mg/ml) in 10 mM buffer (protein in its native state), just after addition of 50% (v/v) ethanol, and after 10 days of incubation at room temperature (concentration of buffer in samples with ethanol remains unchanged respective to the aqueous solution. i.e., 10 mM. These three representations show the different behaviour of the least concentrated system (1 mg/ml) with respect to higher concentrated ones: 10 and 40 mg/ml of β-Lg. Autocorrelation functions in Figure 3.22 represent the results obtained in repetition A as an example. APPENDIX K shows every single autocorrelation function obtained for all individual experiments (repetitions A, B, and C) for each protein concentration and different incubation times, together with the theoretical autocorrelation function obtained from the mean hydrodynamic radii (R_h) shown in Table 3.7.

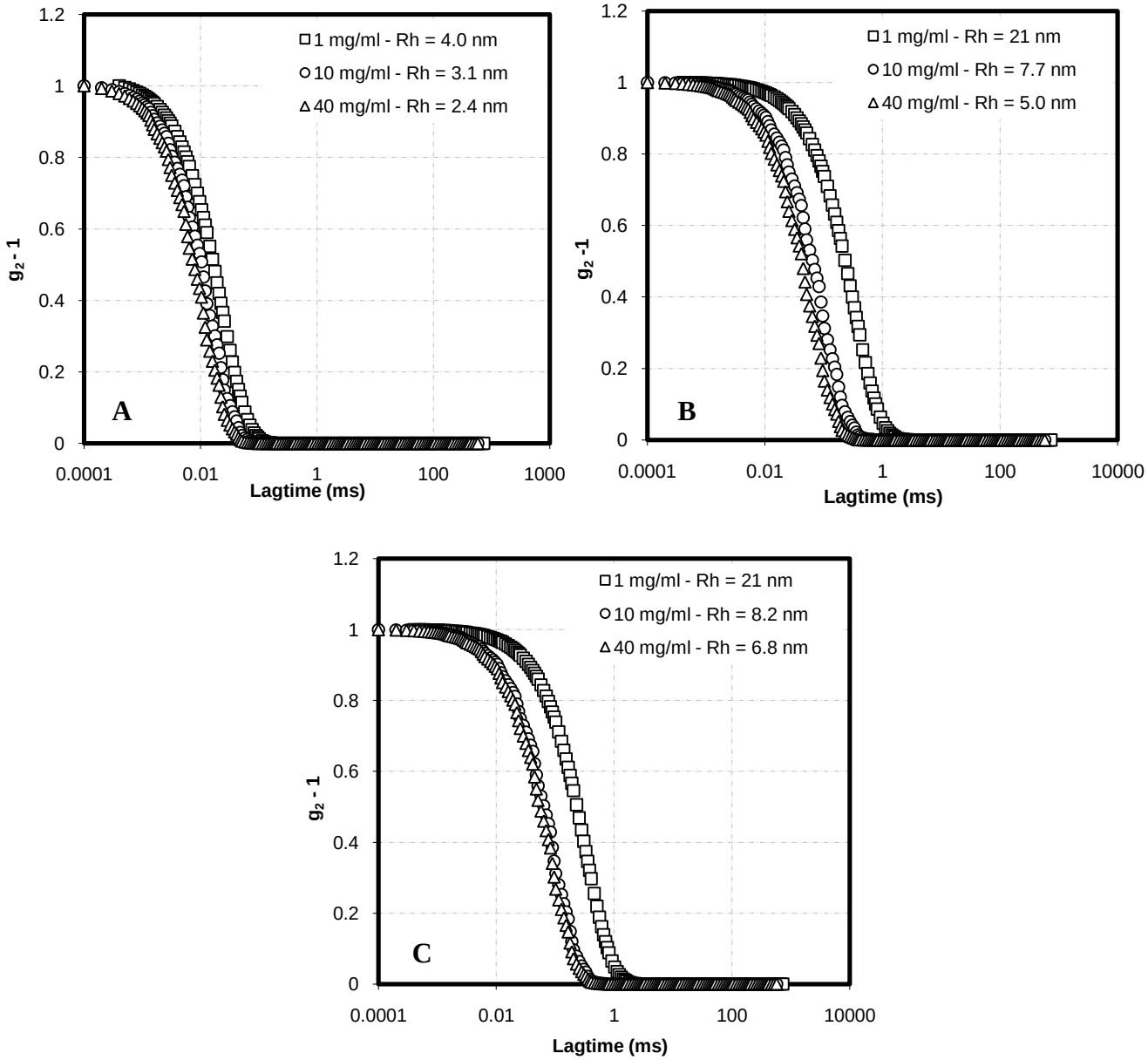


Figure 3.22 - Normalised autocorrelation functions (together with the apparent hydrodynamic radii (R_h)) obtained from different concentration of protein when dissolved in 10 mM buffer – native state (A); after adding 50% (v/v) of ethanol (B), and after 10 days of incubation at room temperature (C). Concentration of buffer remains the same for samples in the presence of ethanol (10 mM). Data correspond to repetition A of experiments. For whole collection of data see APPENDIX K.

In Figure 3.22 - A it can be seen that autocorrelation functions for every concentration of the protein dissolved in 10 mM buffer are similar to each other leading to hydrodynamic radius values which do not differ significantly (for more detail see section 3.1 for aqueous solution of the protein and APPENDIX D).

When 50% (v/v) of ethanol is added to the protein (Figure 3.22 - B), autocorrelation functions are similar for higher concentrated systems (10 and 40 mg/ml) but autocorrelation function obtained for 1 mg/ml system shows a first decay delayed respect to the first decay for higher concentrated systems. This leads to a higher value of the hydrodynamic radius once the first decay is fitted as it was explained in section 2.2.5.3. Indeed, there is a big difference between the mean hydrodynamic radius calculated for 1 mg/ml just after adding ethanol and the mean hydrodynamic radius obtained for 10 and 40 mg/ml systems at the same stage (Table 3.7 and Figure 3.20). After 10 days of incubation at room temperature, the significant difference observed between low (1 mg/ml) and high (10 and 40 mg/ml) concentrated samples after addition of ethanol remain constant: the autocorrelation function obtained for 1 mg/ml system has a first decay occurring later in time, leading to higher value of the mean hydrodynamic radius. From DLS results presented above for high concentrated systems (10 mg/ml and 40 mg/ml) it can be concluded that aggregates grow very fast after adding ethanol and only relatively small changes in cluster size are observed during the 10 days of incubation suggesting that the equilibrium size is already reached shortly after the denaturant solvent is added to the protein. Therefore the kinetics of the cluster growth process is too short to be studied conveniently in real experimental time for those systems.

Since less concentrated samples (1 mg/ml) shows more variability in the apparent hydrodynamic radius calculated at different time of incubation at room temperature, it was decided to run DLS experiments in which autocorrelation functions were recorded in shorter periods of time, with the aim of studying better the nucleation growth of clusters for systems with low concentrations of protein. Autocorrelation functions were recorded every 10 minutes during 6 hours for samples containing 1

and 4 mg/ml of β-Lg. The apparent hydrodynamic radii calculated by fitting every autocorrelation function were plotted vs. the incubation time for both concentrations. Results for the sample of 4 mg/ml of β-Lg are shown in Figure 3.23. Results corresponding to the sample of 1 mg/ml are shown in APPENDIX L together with the autocorrelation functions and hydrodynamic radii calculated by fitting the autocorrelation functions for both concentrations.

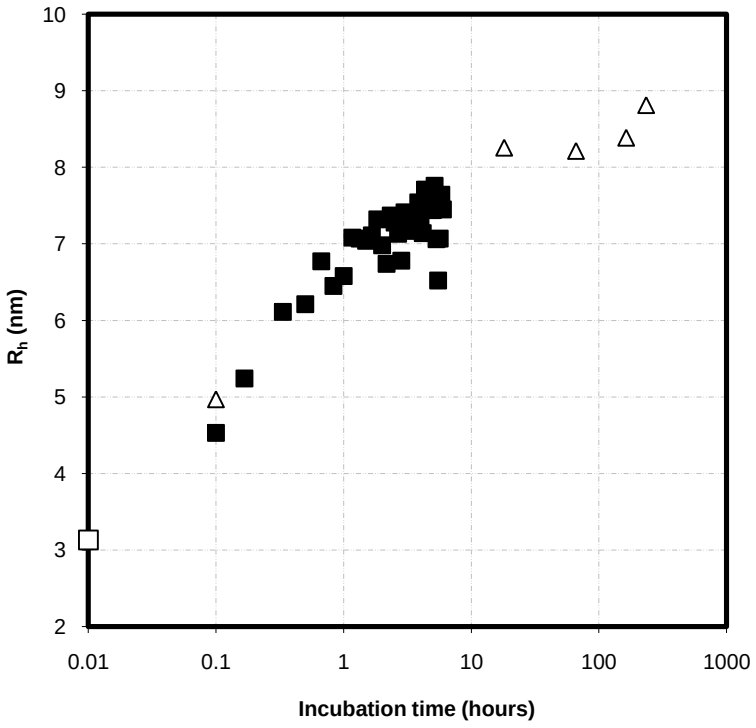


Figure 3.23 - Growth kinetics of aggregates for 4 mg/ml of protein in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7). The empty square intercepting the vertical axis corresponds to the hydrodynamic radius obtained for the aqueous solution of the protein (native state). Full squares correspond with the hydrodynamic radius calculated from short time DLS measurements. Empty triangles correspond with long-term DLS measurements of a second sample (during 10 days incubated at 22ºC).

A clear trend is observed for 4 mg/ml system: hydrodynamic radius increases 1.4 times respecting to the native state of the protein with the same concentration of

buffer (10 mM) when ethanol is added, and keeps increasing during the next 65 minutes, after which hydrodynamic radii values fluctuate considerably. Results from the sample measured just after addition of ethanol and after 1, 3, 7 and 10 days of incubation at room temperature was also included in Figure 3.23. It is clear that the hydrodynamic radius follows an increasing tendency during the whole range of incubation time. However, it is difficult to see a defined lag phase for the clusters growth since they keep growing after ethanol addition. After 1 day of incubation at room temperature, a plateau phase is attained were the calculated hydrodynamic radius remains constant up to 10 days. This apparent hydrodynamic radius was calculated to be 8.2 ± 0.1 nm. These results were not expected since hydrodynamic radius values for the system of 4 mg/ml did not fall between the hydrodynamic radii obtained for 1 mg/ml and 10 mg/ml according with the tendency observed with these systems (Figure 3.20): the higher the protein concentration, the smaller the hydrodynamic radius. Instead, DLS of 4 mg/ml system led to hydrodynamic radii close to those obtained for 10 mg/ml system. Even though only one repetition was done for this system, the smooth shape of the autocorrelation functions (APPENDIX L) did not suggest any major problem occurring during the experiment, so results must be considered as preliminary results interesting to be studied further.

Some of the samples prepared for this project were also analysed by small angle X-ray scattering (SAXS) at beamline SWING at Soleil synchrotron source in Gif-Sur-Yvette, France, by collages Dr. Nadeem Javid and Dr. Karsten Vogtt as a collaboration for this project. This results were published recently (Vogtt, Javid, Álvarez, Sefcik and Bellissent-Funel, 2011) and they will be shown in APPENDIX M in this thesis.

4. GENERAL DISCUSSION

There is still only limited understanding of concentration dependent mechanism and kinetics of structural and size transitions undergone by proteins in early stages of the nucleation process. It is widely believed however in the scientific community that intermediate species (oligomers / protofibrils) formed during the lag period of the amyloidogenesis process constitute the potential toxic species rather than the mature and higher ordered structures known as amyloid fibrils. It is also thought that oligomers may not be obligate intermediates on the fibril formation pathway but rather accompanying species that may co-exist with the protofibrils leading to the formation of mature amyloid fibrils. Moreover, it is not clear whether these oligomers may eventually transform into mature fibrils or if they constitute an alternative pathway that is different from the typical nucleation-dependent fibril assembly extensively believed to be the one followed in amyloidogenesis (Necula et al., 2007).

In this study, the structural and size transitions undergone by β-Lg, used as a model protein, triggered by the addition of a low dielectric constant solvent (ethanol) were followed for up to 10 days of incubation at room temperature. Protein aggregation was started by the disruption of the hydration shell of the protein, caused by the addition of the organic solvent. Kinetics of the dynamic structural and size transitions undergone by the protein during the nucleation process was monitored by applying several spectroscopic and scattering techniques. This constitutes the main core of this project. However, other related effects were also observed, involving not just the influence that the organic solvent has in the aggregation process of β-Lg but also the influence that different concentrations of salts (ions) have on the protein apparent hydrodynamic radius when ethanol is not present in solution and β-Lg is presumably in its native state. This will be the first point to discuss in this section, linked to results presented in section 3.1. The second part will deal with the behaviour of the protein in the presence of different concentrations of ethanol, corresponding to

results shown in section 3.2. Finally, the last part of this discussion will involve the results presented in section 3.3 and that represents the main part of this research project.

4.1. DISCUSSION OF RESULTS CORRESPONDING TO AQUEOUS SOLUTION OF β-Lg

β-Lg is known to be a globular protein in its native state and its quaternary structure in solution depends on pH. This study was carried out at pH 7. For this value of pH β-Lg is present in a dimeric form whose hydrodynamic radius has been reported in literature to be around 3 nm (Gimel et al., 1993; Aymard et al., 1996; Parker et al., 2005). Interesting results were obtained in this project with dynamic light scattering showing variation of the hydrodynamic radius when β-Lg at different concentrations was dissolved in pure water and at different buffer concentrations. It is worth mentioned that since experiments were carried out at relatively low buffer concentrations where protein remains well soluble and preserves its native structure at both secondary and tertiary level, salting out effects can be ruled out. Salting out typically occurs at very high salt concentrations (which are protein and salt dependent), leading to protein precipitation through increasing the strength of hydrophobic protein-protein interactions and subsequent aggregation (Curtis, Ulrich, Montaser, Prausnitiz, and Blanch, 2002).

For a constant concentration of protein, the apparent hydrodynamic radius was found to increase with increasing electrolyte concentration in solution. This can be explained due to repulsive electrostatic interactions between negatively charged protein molecules in the electrolyte solution at neutral pH (Monahan, German, and Kinsella, 1995; Shpigelman, Israeli, and Livney, 2010). These repulsive interactions cause protein molecules to move around faster and thus increase their diffusion rate, which leads to a smaller value of the apparent hydrodynamic radius according to the Stokes-Einstein equation (Finsy, 1994). When the concentration of electrolyte ions in solution increases, the electrostatic interactions among protein molecules become

more screened, i.e., they act over a shorter range (Hiemenz and Rajagopalan, 1997, pp. 499-533). This is due to the reduction of the double layer surrounded the charged protein molecules in solution. Therefore, protein molecules repel each other less strongly, leading to a lower diffusion rate and therefore to a higher value of the hydrodynamic radius. This is called the Debye screening effect, in which mobile ions in solution can reduce (screen) both attractive and repulsive interactions in proteins (Ramos and Baldwin, 2002). Eventually, when repulsive interactions are sufficiently screened such that their effect on protein molecule diffusivities becomes negligible, a "true" hydrodynamic radius for their corresponding tertiary structure (native state) can be obtained (assuming that no additional salt-induced clustering has taken place).

This would explain why the apparent hydrodynamic radius of β-Lg in its native state was found to grow with increasing concentration of electrolyte ions in solution, from distilled water to 10 mM and 20 mM of phosphate buffer concentration. It was also observed by DLS that even the small volumes of HCl and NaOH used on occasion to adjust the pH of samples to 7 had an observable effect on the hydrodynamic radius calculated from the fitting of the autocorrelation functions. This corroborates even more the importance that the charge screening has on the apparent protein size as measured by DLS.

A similar effect to that of increasing electrolyte concentration was observed when concentration of protein solutions was varied from the most concentrated (40 mg/ml) to more dilute, 10 mg/ml and 4 mg/ml, where the apparent hydrodynamic radius was increasing with dilution. Here repulsive interactions between protein molecules were weaker at lower protein concentrations because average distances between protein molecules were becoming longer with increasing dilution, and so protein molecules repelled each other less strongly and their diffusion rates became less affected by their mutual interactions. In summary, lower protein concentrations and higher electrolyte concentrations, within the bounds of values indicated above, lead to decreasing repulsive interactions and increasing hydrodynamic radius approaching

its limiting value of about 3 nm, corresponding to ideal Brownian diffusivity of a β-Lg dimer (e.g. Parker et al. 2005).

However, this was found to be no longer true at the lowest protein concentration of 1 mg/ml, where very surprising results were obtained. In comparison with higher concentrations of protein (4, 10 and 40 mg/ml), a drastic increase in the hydrodynamic radius is observed especially for the highest concentration of electrolyte, 20 mM, for the protein concentration of 1 mg/ml. This is in contrast with what would be expected from theoretical considerations above, since repulsive interactions should be weakest at the lowest protein concentration and we should get a limiting hydrodynamic radius of about 3 nm corresponding to freely diffusing non-interacting dimer (e.g. Parker et al., 2005). To explain this unexpected behaviour at the lowest protein concentration, a deeper insight in the autocorrelation functions obtained from dynamic light scattering measurements is required. As it was shown in section 3.1, autocorrelation functions obtained for 1 mg/ml of protein concentration in water and different concentrations of buffer do not present the single exponential decay that is observed for higher protein concentrations (Figure 3.6). Instead, autocorrelation functions show a distinct tail at longer decay times indicating a highly poly-dispersed system (Figure 3.7). Therefore, the fitting of the initial decay of the autocorrelation function to calculate the hydrodynamic radius for this system cannot be done accurately and results can only be considered as approximate values. The particle distribution reveals the existence of two main populations composed of smaller particles (of average size 3 nm, corresponding to the dimeric β-Lg) and much larger particles whose size depends on the electrolyte concentration in solution: about 50 nm for distilled water, and about 200 nm in the presence of buffer. Moreover, for the highest concentration of buffer (20 mM), a third population with particles of average size 2 μm was also observed but the proportion of this large population was very small in comparison with the other two populations. After an overall analysis of all samples with protein concentration of 1 mg/ml, it can be concluded that solutions at this concentration of β-Lg show a significant variability in the measured average hydrodynamic radius, as it can be seen by the standard deviation calculated from the

hydrodynamic radii obtained from different repetitions of experiments using native solutions of β-Lg in water and different concentrations of buffer (Table 3.3). This suggests that samples of 1 mg/ml are highly poly-dispersed systems, which seem to be pre-aggregated in its native state.

A possible speculative hypothesis to explain these results is based on the (unknown) composition of the commercial β-Lg powder used in this study. The stock powder was used without further purification and it may well contain some molecular or ionic additives used in the manufacturing process with the aim of stabilising and/or purifying the protein. It is thought that when the protein powder is dissolved at very small concentration (1 mg/ml), any additive compounds become diluted to such low concentration that they may not be able to prevent spontaneous protein clustering due to native aggregation and hence the DLS results obtained for 1 mg/ml systems. However, we were not able to get any information from the supplier regarding manufacturing related additives, since this is likely to be a trade secret, and therefore this hypothesis could not be corroborated. Another hypothesis to explain the different behaviour of the least concentrated samples could be the relative lower protein to electrolyte concentration ratio for samples with 1 mg/ml of β-Lg. It was thought that the higher concentration of electrolyte (especially 20 mM) relative to the scarce protein molecules in solution would seal the latest to a higher level; this higher solvation of protein molecules would affect the apparent hydrodynamic radius observed, since by definition, the hydrodynamic radius encompasses hydration and shape effects of the molecules in solution. This hypothesis is preliminary supported by the results found in the experiment where the same protein molecule to electrolyte ions ratio for samples with 1 mg/ml was done for a sample containing 40 mg/ml (for which the buffer concentration was needed to be 0.4 M). In this case, the apparent hydrodynamic radius observed was higher than that for less concentrations of buffer for the same protein concentration (40 mg/ml) as shown above. However, more analysis is required to corroborate this hypothesis and find a conclusive explanation about the unexpected behaviour of samples containing 1 mg/ml of β-Lg.

To summarise discussion of results for the native protein solution studies, some final statements can be reported. The main conclusion provided by dynamic light scattering applied to native protein solutions is the noticeable effect that repulsive electrostatic interactions between negatively charged protein molecules have on the apparent hydrodynamic radius obtained for different protein and electrolyte concentrations. Increasing protein concentration and decreasing electrolyte concentration leads to the apparent hydrodynamic radius decreasing with respect to its ideal value of about 3 nm due to freely diffusing native dimers present in these solutions at pH 7. However, at the lowest protein concentration of 1 mg/ml, extensive clustering was observed which is inconsistent with the behaviour of more concentrated systems. This may be related to unknown additives present in the stock protein powder or to a possible important solvation effect due to the low protein to electrolyte concentration ratio in these systems.

The native state of the protein can be disrupted in the presence of destabilising conditions in its solution environment. In this project, a low dielectric constant solvent (ethanol) was used to disrupt the hydration shell of the β-Lg and induce its denaturation. Results obtained in these experiments are discussed in the following section.

4.2. DISCUSSION OF RESULTS CORRESPONDING TO SOLUTIONS OF β-Lg WITH DIFFERENT PERCENTAGES OF ETHANOL

The disruption of the hydration shell of the protein by ethanol leads to a higher exposure of its hydrophobic regions to the solvent, allowing subsequent aggregation. Results from DLS measurements corroborate this since the hydrodynamic radius measured in the presence of different ethanol concentrations (from 10% (v/v)) was higher than that of the native protein dimer observed in the aqueous solutions with the same concentration of the protein and buffer. Two concentrations of the protein were studied in detail: 10 and 40 mg/ml (note that significant clustering was found in aqueous solutions of β-Lg at the concentration of 1 mg/ml as described above) with

and without buffer (10 mM), for different concentrations of ethanol (up to 50% (v/v)) as it was presented in section 3.2.

It was observed that in all cases (at both protein concentrations and with or without buffer) the measured hydrodynamic radius was increasing with ethanol concentration for up to 50% of ethanol (v/v) (Figure 3.8). For the lowest ethanol concentration (10% (v/v)) in the presence of 10 mM buffer the measured hydrodynamic radii were very similar to the ones obtained from the aqueous solutions (0% ethanol), so it can be concluded that there was no significant aggregation present at that concentration of ethanol and the protein is likely to remain in its native dimer state. From 20% (v/v) of ethanol onwards it can be seen that the measured hydrodynamic radius for both 10 mg/ml and 40 mg/ml of protein in the presence of 10 mM buffer is steadily increasing with increasing ethanol concentration. The hydrodynamic radii for 40 mg/ml samples were always about 22% lower than those measured for 10 mg/ml solutions at the same ethanol concentration, similarly to results obtained for the native protein aqueous solutions in the same conditions (Table 3.3). This can be attributed once again to increasing repulsive interactions between negatively charged protein molecules, which become more prominent at higher protein concentrations (due to decreased average distances between neighbouring molecules). Stronger repulsive interactions results then in faster diffusion and smaller apparent hydrodynamic radius, as discussed previously. In the absence of buffer, measured hydrodynamic radii for 40 mg/ml protein solutions were always lower that those measured in the presence of the buffer at all ethanol concentrations, again in agreement with behaviour observed in aqueous solutions. However, this was not the case for 10 mg/ml protein solutions, where the measured hydrodynamic radii in pure water were higher than those from buffered solutions for ethanol concentrations of 20% (v/v) and higher, where protein clusters formation is present. It is interesting that while the absence of buffer in 40 mg/ml solutions apparently resulted in stronger repulsive interactions among formed clusters (since the corresponding hydrodynamic radii were lower than those in the presence of buffer), for 10 mg/ml solutions the absence of buffer seems to have led to more extensive aggregation instead, forming

even lager clusters compared to those found in the presence of buffer at the given ethanol concentration (≥ 20 % (v/v)).

Apart of the quantitative DLS results obtained for samples with different percentages of ethanol, a comparison with previous studies reported in literature under similar conditions would be of interest. We note that there is not a relatively wide body of literature reporting studies of solvent-induced denaturation of β-Lg in water/buffer-alcohol mixtures. However, experimental conditions as well as experimental methods used varied considerably among different studies, so direct comparisons are sometimes difficult to make. Table 4.1 shows the conditions of previous studies on solvent-induced denaturation and gelation of β-Lg reported in literature.

Table 4.1 - Summary of conditions used in different studies done in solvent-induced denaturation and gelation of β-Lg.

Author	Alcohol	Conc. % (v/v)	pH	Temp.	Buffer
Dufour et al. 1998	Ethanol	50%	7 & 8	25°C	-
Renard et al. 2000	Ethanol	50%	7, 8 & 9	25°C	-
Dufour et al. 1994	Ethanol	50%	7 & 8	25°C	-
Gosal et al. 2004	Ethanol, Methanol, Propanol & TFE	50%	2 & 7	25°C	20 mM phosphate buffer
Mousavi et al. 2008	Ethanol	10 – 15%	1.8 & 6.5	20°C	20 mM glycine buffer & 50 mM phosphate buffer
Gosal et al. 2004	Ethanol & TFE	50%	2 & 7	25°C	20 mM phosphate buffer
Dong et al. 1998	TFE	50%	7.4	20°C	10 mM potassium phosphate buffer
Yoshida et al. 2010	Methanol, Ethanol, 1-Propanol, 2-Propanol, *tert*-Butanol, TFE & HFIP	10 – 80%	2	20°C	-

Most of previous works involving β-Lg solvent – induced denaturation studied the aggregation behaviour of β-Lg under different conditions of pH or by using other solvent instead of ethanol (trifluoroethanol (TFE) was found to be the most widely used in this type of studies at pH 7) (Dong, Matsuura, Manning and Carpenter, 1998; Gosal et al., 2004a; Gosal et al., 2004b; Mousavi et al., 2008). Moreover, even when studies are done with different percentages of ethanol as the denaturant solvent, they are focused in the study of the secondary and tertiary structural transitions undergone by the protein rather than growth of protein clusters. Since spectroscopic and scattering techniques were only applied for samples containing 50 % of ethanol (v/v) in this work, results cannot be compared at the structural level of the protein for ethanol concentrations varying from 10 to 40 % (v/v). However, very interesting is the fact that most of these studies reported the formation of β-Lg gels when the protein concentration and ethanol concentration were sufficiently high. In these terms, Dufour et al. (1995) found that β-Lg formed milky gels in 50% (v/v) water-ethanol mixtures at pH 7 for protein concentration as high as 45.5 mg/ml at ambient temperature. The same author later studied in detail the influence of pH and protein concentration in the gelation process of β-Lg concluding that β-Lg gelation was faster at neutral pH (pH 7) and that when increasing protein concentration, the turbidity of samples increased dramatically and the homogeneity in the gel decreased (Dufour et al., 1998). They reported the formation of gels in 50% (v/v) water-ethanol mixtures at pH 7 for protein concentrations in the range 14.5 – 40.2 mg/ml, with gelation time decreasing with increasing protein concentration. For protein concentrations similar to those used in this project (14.5 mg/ml and 40.2 mg/ml closest to 10 and 40 mg/ml, respectively) the gelation times were reported to be approximately 4 days and 1.5 minutes respectively. In the same way, Renard et al. (2000) observed the formation of turbid β-Lg gels at pH 7 in 50 % (v/v) water-ethanol solutions at room temperature for protein concentration 40 mg/ml after 8 days of incubation, concluding that the turbidity of these gels increased when increasing protein concentration, pH or ethanol concentration. In the present work no gels were formed in any case for 50% (v/v) water/buffer-ethanol mixtures after 10 days of incubation at room temperature even for the highest protein concentrated

systems, although some of these samples formed clear gels at timescale of months (see below). This lack of gel formation in relatively short periods of incubation time could be attributed to the nature of the β-Lg used in each case: while the protein used by Dufour and Renard contained 3% of α-Lactalbumin and 2% of serum albumin with a purity of 94%, the β-Lg purchased from Sigma Aldrich for this work had a purity of 90% but none of the other whey proteins according to commercial specifications. Moreover, no buffer was used in the preparation of samples for their experiments. A more recent work carried out by Yoshida, Yamaguchi, Osaka and Shibayama (2010) reported the formation of gels at concentrations of ethanol of 30, 40 and 50% (v/v) using the same protein as in this project. Their study was carried out at acidic pH (pH 2) and therefore results cannot be directly compared with the present work. It is worth to note here that gelation of β-Lg was indeed observed in this project after long periods of incubation times – several months. However, gels were only found to be present for the highest concentrated system (40 mg/ml of β-Lg) in the presence of 50% (v/v) buffer-ethanol mixtures. In these systems, transparent gels were observed after approximately 13 months of incubation at room temperature. Since samples were not checked regularly because it was not the objective of this work, it is possible that these gels were actually formed at some earlier time (shorter than 13 months). This is supported by the fact that one of the samples of 40 mg/ml in 50% (v/v) buffer-ethanol mixture was found to be gelled after only 27 days of incubation at room temperature, while several other samples under the same conditions did not result in gels over the same time period. It is not possible therefore to get to a general conclusion about the gelation process in the samples prepared for this project except for the fact that very slow gelation at timescale of months was occurring for buffered (10 mM) samples of 40 mg/ml in the presence of ethanol (50% (v/v)). Small differences in pH, final protein or electrolyte concentration and physical arrangements (such as presence or absence of magnetic stirrers in solution and different geometries of vials) can be the cause of the variability of observed timescales in β-Lg gelation process.

With the aim of checking if the turbidity observed by other authors would be present in the samples prepared for this project by increasing ethanol concentration, samples of 10 mg/ml in 10 mM phosphate buffer with percentages of ethanol from 60 to 90% (v/v) were prepared for a qualitative test. It was found that when ethanol concentration rose to 70% (v/v), a cloudy sample was obtained. When the percentage of ethanol was as high as 80% (v/v) a very milky sample was observed and when the ethanol represented the 90% (v/v) of the total sample, a milky sample with big clusters and their subsequent precipitation was observed. However, no solid gels were found in any case at least after 10 days of incubation at room temperature.

In summary, it can be concluded that the kinetics of β-Lg gelation at room temperature and neutral pH in 50% (v/v) water/buffer – ethanol mixtures, in the order of months as observed for our systems, is several orders of magnitude slower than that found by other authors under similar conditions. This can be due to differences in the composition of the protein stock powder and buffers used in various studies.

Since β-Lg aggregates formed in 50% (v/v) water/buffer – ethanol mixtures at neutral pH appeared to be stabilised against further aggregation and subsequent gelation for extended periods of time, we were able to characterise them in great detail, using a number of different analytical techniques. This allowed us to gain a novel insight into these intermediate species on the protein aggregation pathway, as is discussed below in detail.

4.3. DISCUSSION OF RESULTS CORRESPONDING TO 50% (v/v) BUFFER-ETHANOL SOLUTIONS

The kinetics of structural and size transitions undergone by β-Lg in the presence of 50 % (v/v) buffer (10 mM) – ethanol mixtures was monitored by applying the spectroscopic and scattering techniques as described in previous sections. To discuss the protein concentration dependent structural changes and kinetics of

nucleation/oligomerization from a qualitative point of view it was decided to develop a general representation involving all results from ThT-binding fluorescence, circular dichroism and dynamic light scattering analysis for the main protein concentrations under study – 1, 10 and 40 mg/ml vs. the incubation time. CD analysis was not done for samples with 4 mg/ml of β-Lg but both fluorescence and dynamic light scattering results have been also included here as part of the general discussion. Data correspond to those shown before in the results section: for ThT binding fluorescence, the fluorescence intensity (observed at 477 nm); for DLS, the mean hydrodynamic radius (measured at the scattering angle of 90°) and for CD, the α-helix to β-sheet content ratio calculated from the deconvolution of far UV CD spectra. Fluorescence and CD data were normalised to the native state of the protein, acting as the reference. Mean hydrodynamic radii were normalised to the actual size for the dimeric form of β-Lg at neutral pH, equal to 3 nm as discussed above; therefore the normalised hydrodynamic radii represent directly comparable values. All these results are shown in Figure 4.1. For clarity, both axes were presented in logarithmic scale due to the wide range of data values. For the case of the incubation time presented in horizontal axis, the values for native β–Lg and just after addition of ethanol conditions have been chosen symbolically as a licence to allow representation of data as it was explained above when similar charts were presented.

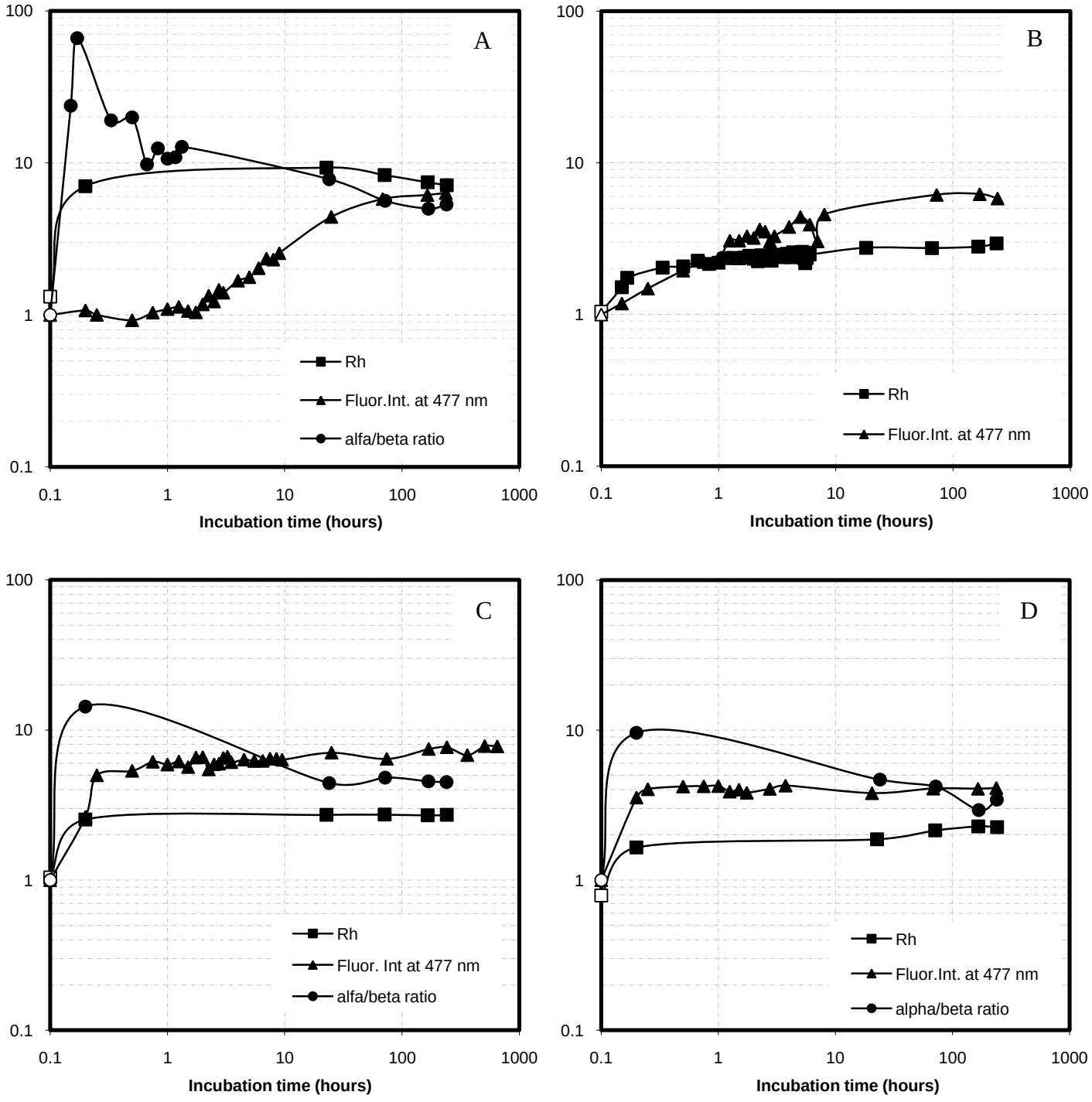


Figure 4.1 - Summary of fluorescence, circular dichroism and dynamic light scattering results for the main concentration of protein under study: 1 mg/ml (A); 4 mg/ml (B); 10 mg/ml (C) and 40 mg/ml (D). Fluorescence and CD data are normalised to native values. DLS data are normalised to the actual hydrodynamic radius for the dimeric β-Lg at neutral pH (3 nm).

For the least concentrated system (1 mg/ml of protein) presented in Figure 4.1 (A) it was observed that the hydrodynamic radius increases immediately after adding ethanol, reaching a value of about 7 times higher than that of the native dimer and then it stays approximately constant thereafter. We also note that at this low protein concentration there was pre-aggregation observed even in aqueous buffer solutions before any ethanol was added, as discussed above. It can be seen from CD results that just after addition of ethanol, the alpha to beta ratio increases sharply to very high values. This is because the secondary structure switched from its native form (high in β-sheets and low in α-helices) to high α -helix content and very low β–sheet content, while the tertiary structure of the protein was completely destroyed immediately after adding ethanol (Figure 3.12 and Figure 3.13). The low ThT intensity observed is in agreement with the low overall β-sheet content in the secondary structure after addition of ethanol as shown by the far UV CD analysis. The addition of ethanol disrupts the hydration shell of the protein, leading to a major exposure of its hydrophobic ends to the solvent because of the preferential binding of the non-polar residues to ethanol molecules. This destabilization in the non-covalent dispersion forces of the protein molecules reduce the energy barrier required for structural transitions which in return will cause aggregation / oligomerization of β-Lg by hydrophobic binding (Grudzielanek et al., 2005). Pace et al. (2004) reported that the formation of α-helix in ethanol is somehow expected base on thermodynamic studies, since this type of structure would involve the burying of the peptide groups in the inner part of the helix, leaving the non-polar residues exposed to the polar solvent. After the first burst in the secondary structure of the protein towards the high α-helix and low β-sheet conformation, it keeps changing during the next day of incubation towards recovering some β-sheet content and corresponding decrease of the α-helix content. ThT fluorescence intensity evolves in parallel with these secondary structural transitions. A lag period of 1.5 hours (96 minutes) is observed after which an exponential growth indicates nucleated restructuring process involving non-native protein conformations. A plateau in the fluorescence intensity is observed after 3 days of incubation at room temperature. The overall shape of the

structural re-arrangements kinetics for ThT- binding fluorescence is a sigmoidal curve characteristic for the nucleation-based growth mechanism, having a growth rate of 1.5 $hour^{-1}$ (Jarret, Berger and Lansburry, 1993). Very similar sigmoidal curve was obtained for a system with even lower concentration of protein (0.5 mg/ml) as it was shown in Figure 3.11. The lag period for this system was calculated to be 2.6 hours and the growth rate 0.4 $hour^{-1}$. Longer lag period and slower growth rate for lower concentration of protein is consistent with the nucleation dependent structural rearrangement of protein clusters follow by β-Lg at low concentrations (0.5 and 1 mg/ml). The secondary structure of the protein seems to have also attained the equilibrium state after 3 days of incubation at room temperature (no significant changes observed in the alpha to beta ratio after 3 days of incubation). This plateau achieved both for structural and size changes indicates that the 1 mg/ml system has reached a stable state with a new free energy minimum after 3 days of incubation and that there are likely no more free β-Lg molecules left in the system.

For the system containing 4 mg/ml of β-Lg, ThT- binding fluorescence starts to increase immediately after adding ethanol to the native protein solution (Figure 4.1 (B)), while it still seems to have a somewhat sigmoidal shape, suggesting that the lag period for this protein concentration is too short to be caught within the experimental time span in our experiments and, in any case, it is much shorter than those calculated for lower protein concentrations (0.5 and 1 mg/ml). The fluorescence intensity continues increasing for the first 72 hours after addition of ethanol, after which the plateau phase is attained. Cluster growth starts immediately after adding ethanol (Figure 3.23), but it is gradual over the first 1 hour, followed by much slower further increase for up to 1 day of incubation, after which the mean hydrodynamic radius remains almost constant for up to 10 days of incubation. It can be seen that at this protein concentration the kinetics of structure evolution leading to enhance ThT binding is becoming comparable to the cluster growth kinetics as seen from DLS measurements. At 4 mg/ml, the kinetics of structure evolution is also much faster than that for the more dilute system at 1 mg/ml.

When protein concentration increases to 10 mg/ml, ThT binding fluorescence measurements show a drastic increase of fluorescence intensity immediately after adding ethanol (much higher compared to 4 mg/ml system), followed by a rapid growth over the next several minutes, while changing only little afterwards. A burst in protein secondary structure is taking place just after ethanol addition: β-sheet content is strongly reduced and α-helix content increases, but not as much as for the 1 mg/ml system. Therefore, the alpha to beta ratio for 10 mg/ml of protein concentration is significantly lower than that for concentrations of 1 mg/ml. The secondary structure of the protein keeps changing during the incubation period: β-sheet content increases (and α-helix content decreases) after 1 day of incubation. In contrast to the secondary structure evolution, the hydrodynamic radius experiences a sudden increase essentially to its final value, which is 2.6 times higher than that of the native dimer, and does not show any significant changes over time for up to 10 days of incubation at room temperature. We note that this is the same final size as seen for clusters in the 4 mg/ml system (Figure 4.1 (B)). No significant structural or size changes are observed after 1 day of incubation. These results suggest that in 10 mg/ml solutions, protein clusters attains their equilibrium size (in terms of the mean hydrodynamic radius) just after adding ethanol but secondary structural changes (as seen by ThT binding and CD) are still taking place inside the formed clusters. This constitutes a very interesting novel insight into the mechanism of protein aggregation.

The behaviour of the most concentrated system (40 mg/ml) is similar to that at 10 mg/ml of protein concentration, but now kinetics of structural evolution (by ThT binding) appears to be even faster than that of cluster growth (by DLS). ThT binding fluorescence increases immediately after adding ethanol to its final value and remains unchanged during 10 days of incubation at room temperature. The hydrodynamic radius increases immediately after ethanol addition and subsequently grows further around 30% over the next few days of incubation. The same secondary structural burst as before is observed just after ethanol addition, followed by the subsequent increase of the overall β-sheet content and decrease of the α-helix

content for up to 1 day of incubation. As it was observed for the system with 10 mg/ml of protein, structural changes at the secondary structure level are still taking place when the equilibrium cluster size has been reached, although suitably organised β-sheets able to bind the dye must have been present immediately after ethanol addition (as see by ThT binding).

Although the measured hydrodynamic radius of the clusters formed at 40 mg/ml was 24% lower than that found at 10 mg/ml, this is due to repulsive interactions in the more concentrated system, as discussed above. This is also evidenced by the peak present in the SAXS scattering intensities measured for 40 mg/ml solutions (Vogtt et al., 2011), which is almost absent at 10 mg/ml (APPENDIX M). This is because at higher concentrations, clusters rearrange themselves with a certain correlation length by the balance of short range attractive and long range repulsive interactions. The SAXS measurements done by collaborators (Dr. Nadeem Javid and Dr. Karsten Vogtt) on samples studied in this project further confirmed that the native dimeric β-Lg undergoes aggregation and forms small stable clusters upon ethanol addition. By analysis of SAXS and DLS data, it was found that these clusters are fairly compact and sphere-like in shape, rather than fibril like structures. The spherical structure of clusters may be attributed to the mechanism of their formation, where a large content of α-helix formed upon addition of ethanol in the lag phase (before cluster formation) may act as a barrier restricting beta-beta stacking to form fibril like structure. These spherical clusters were found to have an enhanced tendency to bind ThT molecules (compared to the native structure) which is most likely an indication of the presence of suitably organised intermolecular or intramolecular β-sheet structures.

5. CONCLUSIONS

β-Lg undergoes important structural changes once ethanol is present (at 50% v/v) in the solution of native dimeric β-Lg in phosphate buffer at pH 7. The tertiary structure is completely disrupted in the presence of the denaturant solvent and the secondary structure experiences major transitions. While the native state of β-Lg is characterised by majority content of β-sheet, CD analysis indicates that a large content of α-helix and much less of β-sheet is present immediately after adding ethanol. This high content of α-helix seems to be an unstable state for the protein in the given solution environment, since the secondary structure gradually evolves towards an equilibrium state: β-sheet content increases and α-helix decreases over 10 days of incubation at room temperature. However, this new equilibrium state does not correspond to a native-like secondary structure, since the content of β-sheet and α-helix are significantly lower and higher respectively to those in the native protein. Secondary structure transitions seem to indicate the formation of an intermediate with a new free energy minimum in solution, whose secondary structure is predominantly α-helix for every concentration of protein. ATR-IR results also indicate that the content of α-helix increases when ethanol is present in solution, together with a decrease of intramolecular β-sheet.

ThT-binding fluorescence spectroscopy indicates that structure evolution kinetics of β-Lg follows a nucleation-dependent growth mechanism for low protein concentrations (0.5 and 1 mg/ml) where a lag period, an exponential growth and a plateau phase are clearly identified and quantitatively analysed. When protein concentration increases (from 4 to 40 mg/ml), the structural rearrangements as monitored by ThT fluorescence become more rapid with increasing protein concentration and the equilibrium state is achieved faster.

Fluorescence intensity in aqueous buffer solutions of β-Lg is approximately proportional to the total protein concentration in solution as can be seen in Table 3.5

and Figure 3.10. During the 10 days of incubation at room temperature after addition of ethanol (50% v/v) the ThT-binding fluorescence intensity gradually increases to a value about 6 times higher than that recorded for the protein in its native state for every concentration (except for the highest concentration of 40 mg/ml where the increase is somewhat less, about 4.5 times). However, kinetics of this transition depends strongly on protein concentration. At 40 mg/ml, transition to the final value is almost instant, while at 10 mg/ml it takes about 1 hour and at 4 mg/ml it takes about 1 day (Figure 3.11). At the lowest protein concentrations (1 mg/ml and 0.5 mg/ml), this transition has a clear lag period followed by a rapid growth, indicating a nucleated phase transition process.

These results suggest that the overall content of α-helix with respect to β-sheets can be related to the lag-phase of the structural transition process, as the corresponding ThT-binding fluorescence measurements show that a distinct lag-phase is observed for lower protein concentrations, where high α-helical content and low β-sheets content was more pronounced and persisted for longer. A very high amount of α-helical structures as compared to β-sheets at low protein concentration slows down the nucleation phase of structure evolution process (longer times of lag phase), while a higher amount of β-sheet at higher protein concentration decreased the lag period.

SAXS analysis showed the size of β-Lg aggregates formed at 10 mg/ml in 50% (v/v) of ethanol to be in good agreement with DLS results, suggesting the formation of sphere-like clusters instead of fibril-like aggregates. This could be due to a high content of α-helix formed upon ethanol addition, which could prevent the beta-beta stacking and therefore the fibril-like aggregates formation. Nevertheless, structural arrangement of β-sheets in these clusters should be such that it allows the enhanced binding of the ThT dye, as evidenced by its fluorescence intensity at 477 nm. It was thought that even though the final shape of aggregates seems to be close to spherical, the internal structure of these spheres must have some suitably organised β-sheet structures able to bind ThT to explain the fluorescence intensity enhancement.

The analysis of DLS results presented in Chapter 4 implies that the real (geometrical) final size of clusters seen at 40 mg/ml is the same as at 10 mg/ml and 4 mg/ml. This indicates that these clusters are nucleated intermediate species in a local free energy minimum, and are thermodynamically controlled intermediates on the aggregation pathway of β–Lg in the mixed buffer-solvent system studied here. The lower apparent higher hydrodynamic radius observed for 40 mg/ml system (Figure 4.1) could be attributed to stronger repulsion interactions in higher concentrated samples; this idea is also supported by SAXS results. These intermediates were remarkably stable, and only after long periods of incubation time (months) at the laboratory temperature, transparent gels were found in samples of 40 mg/ml in 50% (v/v) ethanol.

The broad range of analytical techniques used to study the system considered in this project has led to a complex matrix of results, allowing comprehensive characterisation of the process leading to the formation of nucleated clusters upon the addition of ethanol to aqueous solutions of β-Lg. Identification of these clusters as a novel intermediate on an aggregation pathway of β-Lg, together with the conjugation of spectroscopic and scattering techniques applied to their study, constitutes the main contribution of this work to the scientific knowledge. Nevertheless, further research of this subject would be needed in order to deepen and extend our understanding. Potential future steps in this study are presented in the next and final chapter of this thesis.

6. FUTURE WORK

Future work on this topic could proceed in several directions. First, it would be useful to further investigate and confirm results obtained for 1 mg/ml and 4mg/ml systems measured over short periods of time after ethanol addition by DLS and CD. This would allow developing better understanding of the peculiar behaviour found for low protein concentrations. A better insight in the physical meaning of the long decay time tails of the autocorrelation functions obtained by DLS will also be of major interest, in order to better understand possible large scale clustering phenomena apparently present in relatively diluted protein solutions.

Further characterisation of the clusters found in the system could be done by applying techniques such as atomic force microscopy (AFM) or electron microscopy (EM) with the aim of obtaining images of the clusters with and without sample drying. Dry samples (e.g., using freeze drying) could be subjected to structural analysis techniques such as powder X-ray diffraction (PXRD) or calorimetric techniques such as differential scanning calorimetry (DSC) or differential thermal analysis (DTA).

Analysis of the influence that high pressures would have on the size and structure of β-Lg clusters could also be an interesting future direction for this topic. The use of different solvents such as propanol or TFE, or even different denaturant solutions such as urea or guanidine hydrochloride (GdnHCl) under the same conditions, would likely lead to a much wider range of properties of β-Lg clusters along various aggregation pathways. An additional interesting approach would be the study of possible induction of denaturation when β-Lg clusters are transplanted into aqueous solution of the protein in its native state.

New intermediate species were found in this project, without the formation of mature amyloid fibrils during the incubation time. However, it is very likely that a decrease

in the pH and/or incubation of the system at higher temperatures in the presence of ethanol would lead to the formation of more specific aggregates base in previous studies done with the same protein (Krebs et al., 2009) thanks to the great influence that the solution environment of the protein has on its aggregation pathways (Wang, Nema, and Teagarden, 2010). This would be interesting because the extreme importance that the process of formation of amyloid fibrils has in neudegenerative diseases. Despite the wide research in the field, the detailed mechanisms and pathways of the formation of these structures and their precursors are not fully understood yet; hence, the utility of using model proteins to better understand this process (Lasagna-Reeves et al., 2011; Goldschmidt et al., 2010).

Recent studies have shown the high structural heterogeneity in amyloid fibrils, probably due to molecular level variations in the cross-β sheet motif. Understanding the initial and intermediate species in the amyloid fibril formation as well as possible mechanisms of polymerization such as nucleation and elongation can provide an insight in the final structure of amyloid fibrils (Kumar and Udgaonkar, 2011).

Initially formed amyloid fibrils have been found to evolve further to the formation of supra-molecular-self-assembled structures, such as gels, plaques or spherical assemblies (Juarez et al., 2009). Further analysis of the β-Lg gels found in this project after long periods of time could be also a valuable future work direction. A deeper control of the long-term aggregation found for this system would be important to distinguish different intermediates that may be formed until the formation of gels together with possible off-pathway mechanisms (Gosal el al., 2005; Bhattacharya et al., 2011). It has been found that β-Lg gels have very good swelling behaviour (Reddy, Lavenant, Lefebvre, and Renard, 2006) that can also be applied in the area of drugs delivery. Improvement of the properties of these gels for novel biomaterials could have interesting applications in different areas such as drug delivery or bioengineering.

Despite the uncertain specific function that β-Lg has in the organisms, its ability to bind other compounds is well known. β-Lg complexes have been studied greatly in two major fields: food and pharmacy. Thermally induced protein-nanovehicles have been studied to encapsulate sensitive food compounds that can be degraded at ambient conditions (Gunasekaran, Ko, and Xiao, 2007; Jones, Decker, and McClements, 2009; Shipigel et al, 2010; Jones, Decker, and McClements, 2010). Properties of β-Lg and WPI thermal aggregates have a crucial influence in the final characteristic of emulsions and foams, coating and filming, encapsulation, and cold-gelation, being relevant mainly for the food industry (Nicolai, Britten, and Schmitt, 2011). In the pharmacy area, β-Lg nanoparticles and its complexes have been used for the delivery and controlled release of drugs (Livney, 2010). Most of the stable biopolymers particles obtained were created by thermal treatment of the protein. Despite using organic solvents for food applications is highly questionable nowadays (Nicolai et al., 2011), their use as the denaturation step to start protein aggregation in order to provide possible faster formation of suitably structured biopolymers could be further investigated for different applications.

All these future paths of research regarding β-Lg aggregation may open possible new fields in protein engineering with applications in medicine, pharmacy and food domains.

REFERENCES

Abkevich, V.I., & Shakhnovich, E.I. (2000). What can disulfide bonds tell us about protein energetic, function and folding: Simulations and bioinformatics analysis. *Journal of Molecular Biology, 300,* 975-985.

Acharya, K.R., Ren, J., Stuart, D.I., Phillips, D.C., & Fenna, R.E. (1991). Crystal structure of human alpha-Lactalbumin at 1.7 Å resolution. *Journal of Molecular Biology., 221*, 571-581.

Adamcik, J., Jung, J.M., Flakowski, J., De Los Rios, P., Dietler, G., & Mezzenga, R. (2010). Understanding amyloid aggregation by statistical analysis of atomic force microscopy images. *Nature Nanotechnology, 5*, 423-428.

Aouzelleg, A. (2004). *Molecular and functional investigations of pressure/ temperature induced changes in bovine β-Lactoglobulin and the effect of dextran sulphate*. Published Ph.D. thesis, University of Strathclyde.

Arnaudov, L.N., de Vries, R., Ippel, H., & van Mierlo, C.P.M. (2003). Multiple stops during the formation of β-Lactoglobulin fibrils. *Biomacromolecules, 4*, 1614-1622.

Arnaudov, L.N., & de Vries, R. (2006). Strong impact of ionic strength on the kinetics of fibrilar aggregation of bovine β-Lactoglobulin. *Biomacromolecules, 7,* 3490-3498.

Arnaudov, L.N., & de Vries, R. (2007). Theoretical modelling of the kinetics of fibrilar aggregation of bovine β-Lactoglobulin at pH 2. *The Journal of Chemical Physics, 126*, 145106.

Arnstein, H.R.V., & Cox, R.A. (1992). Protein Biosynthesis (In focus). IRL PRESS at Oxford University Press.

Aymard P., Durand, D., & Nicolai, T. (1996). The effect of temperature and ionic strength on the dimerisation of β-Lactoglobulin. *International Journal of Biological Macromolecules, 19*, 213-221.

Aymard, P., Nicolai, T., & Durand, D. (1999). Static and dynamic light scattering of β-Lactoglobulin aggregates formed after heat-induced denaturation at pH 2. *Macromolecules*, *32*, 2542-2552.

Balakrishnan, S., Javid, N., Weingaertner, H., & Winter, R. (2008). Small-angle X-ray scattering and near–infrared vibrational spectroscopy of water confined in aerosol-OT reverse micelles. *ChemPhysChem, 9*(18), 2794-2801.

Barrow, C.J., Yasuda, A., Kenny, P.T.M., & Zagorski M.G. (1992). Solution conformations and aggregational properties of synthetic amyloid beta-peptides of Alzheimer's disease - analysis of circular-dichroism spectra. *Journal of Molecular Biology*, *225*(4), 1075-1093.

Buxbaum, E. (2007). Fundamentals of protein structure and function. *Part 1: Protein Structure*. Springer, pp. 13-15.

Beretta, S., Chirico, G., & Baldini, G. (2000). Short-range interaction of globular proteins at high ionic strengths. *Macromolecules, 33*, 8663-8670.

Bhattacharjee, C., & Das, K.P. (2000). Thermal unfolding and refolding of β-Lactoglobulin. *European Journal of Biochemistry, 13*, 3957-3964.

Bhattacharya, M., Jain, N., & Mukhopadhyay, S. (2011). Insights into the mechanism of aggregation and fibril formation from bovine serum albumin. *The Journal of Physical Chemistry, 115*, 4195-4205.

Blake, C., & Serpell, L. (1996). Synchrotron X-ray studies suggest that the core of the transthyretin amyloid fibril is continuous β-sheet helix. *Structure, 4*(8), 989-998.

Bolder, S.G., Sagis, L.M.C., Venema, P., & van der Linden, E. (2007). Effect of stirring and seeding on the whey protein fibril formation. *Journal of Agricultural and Food Chemistry, 55*, 5661-5669.

Bhattacharjee, S., Saha, S., Biswas, A., Kundu, M., Ghosh, L., & Das, K.P. (2005). Structural changes of β-Lactoglobulin during thermal unfolding and refolding – an FT-IR and circular dichroism study. *The protein Journal 24*, No.1, 27-35.

Bolder, S.G., Vasbinder, A.J., Sagis, L.M.C., & van der Linden, E. (2007). Heat-induded whey protein isolate fibrils: conversion, hydrolysis and disulphide bond formation. *Internacional Diary Journal*, *17*, 846-853.

Bouchard, M., Zurdo, J., Nettleton, E.J., Dobson, C.M., & Robinson, C.V. (2002). Formation of insulin amyloid fibrils followed by FTIR simultaneously with CD and electron microscopy. *Protein Science, 9*(10), 1960-1967.

Bromley, E.H.C., Krebs, M.R.H., & Donald, A.M. (2005). Aggregation across the length-scales in β-Lactoglobulin. *Faraday Discussions*, 128, 13-27.

Brownlow, S., Cabral, J.H.M., Cooper, R., Flower, D.R., Ywdall, S.J., Polikarpov, I., North, A.C., & Sawyer, L. (1997). Bovine β-Lactoglobulin at 1.8 Å resolution – still an enigmatic lipocalin. *Structure, 5*(4), 481-495.

Bull, L.A. (1998). *Combined effects of thermal and pressure processing on food protein structure.* Published Ph.D. thesis, University of Strathclyde.

Byler, D.M., & Susi, H. (1986). Examination of the secondary structure of proteins by deconvoluted FTIR spectra. *Biopolymers, 25*(3), 469-487.

Carter, D.C., He, X.H., Munson, S.H., Twigg, P.D., Gernet, K.M., Broom, M.B., & Miller, T. (1989). Three-dimensional structure of human serum albumin. *Science, 244*, 1195-1198.

Chen, J., & Stites, W.E. (2001). Packing is a key selection factor in the evolution of protein hydrophobic cores. *Biochemistry, 40*, 15280-15289.

Chi, E.Y., Krishnan, S., Randolph. T.W., & Carpenter, J.F. (2003). Physical stability of proteins in aqueous solution: mechanism and driving forces on non-native protein aggregation. *Pharmaceutical Research, 20*(9), 1325- 1336.

Chiti, F., Webster, P., Taddei, N., Clark, A., Stefani, M., Ramponi, G., & Dobson, C.M. (1999). Designing conditions for *in vitro* formation of amyloid protofilaments and fibrils. . *Proceedings of the National Academy of Science of the United States of America, 96*, 3590-3594.

Chiti, F., & Dobson, C.M. (2006). Protein misfolding, functional amyloid and human disease. *Annual Review of Biochemistry, 75*, 333-66.

Coen, C.J., Blanch, H.W., & Prausnitz, J.M. (1995). Salting out of aqueous proteins: Phase equilibria and intermolecular potentials. *AIChE Journal, 41*, 996-1004.

Come, J.H., Fraser, P.E., & Lansbury, P.T. (1993). A kinetic model for amyloid formation in the prion diseases: importance of seeding. *Proceedings of the National Academy of Science of the United States of America, 90*, 5959-5963.

Cooper, A. (1999). Thermodynamics of protein folding and stability. Published in *Protein: a comprehensive treatise.* JAI Press Inc., *Vol. 2*, pp. 217-270.

Creighton, T.E. (1990). Protein folding. *Biochemistry Journal, 270*, 1-16.

Cromwell, M.E.M., Hilario, E., & Jacobson, F. (2006). Protein aggregation and bioprocessing. *The AAPS Journal, 8*(3), Article 66.

Curtis, R.A., Ulrich, J., Montaser, A., Prausnitz, C.M., & Blanch, H.O.W. (2002). Protein-protein interactions in concentrated electrolyte solutions – Hofmeister-series effect. *Biotechnology and Bioengineering*, *79*(4), 367-380.

D'Alfonso, L., Collini, M., & Baldini, G. (2002). Does β-Lactoglobulin denaturation occur via an intermediate state? *Biochemistry, 41*, 326-333.

Desiraju, R.G., & Steiner, T. (1999). *The weak hydrogen bond - in structural chemistry and biology* (2nd Ed.). Oxford University Press.

Dill, K.P. (1990). Dominant forces in protein folding. *Biochemistry, 29*(31), 7133-7155.

Dong, A., Huang, P., & Caughey, W.A.S. (1990). Protein secondary structures in water from second-derivative amide I infrared spectra. *Biochemistry, 29*, 3303-3308.

Dong, A., Matsuura, S., Allison, S.D., Chrisman, E., C. Manning. C.M., & Carpenter, J.F. (1996). Infrared and circular dichroism spectroscopic characterization of structural differences between β-Lactoglobulin A and B. *Biochemistry, 35*, 1450-1457.

Dong, A., Matsuura, J., Manning, M.A.C., & Carpenter, J.F. (1998). Intermolecular β-sheet results from trifluoroethanol-induced nonnative α-helical structure in β-sheet predominant proteins: infrared and circular dichroism spectroscopy study. *Archives of Biochemistry and Biophysics, 355*(2), 275-281.

Dufour, E., & Haertlé, T. (1990). Alcohol-induced changes of β-Lactoglobulin – retinol-binding stoichiometry. *Protein Engineering, 4*, N°. 2, 185-190.

Dufour, E., Genot, C., & Haertlé, T. (1994). β-Lactoglobulin binding properties during its folding changes studied by fluorescence spectroscopy. *Biochimica et Biophysica Acta, 1205*, 105-112.

Dufour, E., Robert, P., Bertrand, D., & Haertlé, T. (1994). Conformations changes of β-Lactoglobulin: an ATR infrared spectroscopic study of the effect of pH and ethanol. *Journal of Protein Chemistry, 13*, No.2.

Dufour, E., Robert, P., Renard, D., & Llamas, G. (1998). Investigation of β-Lactoglobulin gelation in water/ethanol solutions. *International Dairy Journal, 8*, 87-93.

Dzwolak, W., Kato, M., Shimizu, A., & Taniguchi, Y. (1999). Fourier-transform infrared spectroscopy study of the pressure-induced changes in the structure of the bovine α-Lactalbumin: the stabilizing role of the calcium ion. *Biochimica et Biophysica Acta, 1433*, 45-55.

Dzwolak, W., Grudzielanek, S., Smirnovas, V., Ravindra, R., Nicolini, C., Jansen, R., Loksztejn, A., (…), & Winter, R. (2005). Ethanol-perturbed amyloidogenic self-assembly on insulin: looking for origins of amyloid strains. *Biochemistry, 44*(25), 8948-8958.

Elliot, A., & Malcolm, B.R. (1956). Infrared studies of polypeptides related to silk. *Transactions of the Faraday Society, 52,* 528-536.

Euston, S.R., Ur-Rehman, S., & Costello, G. (2007). Denaturation and aggregation of β-Lactoglobulin – a preliminary molecular dynamics study. *Food Hydrocolloids, 21,* 1081-1091.

Fang Yang, Jr., Zhang, M., Zhou, B-R., Chen, J., & Liang, Y. (2006). Oleic acid inhibits amyloid formation of the intermediate of α-Lactalbumin at moderately acidic pH. *Journal of Molecular Biology, 362,* 821-834.

Ferrone F. (1999). Analysis of protein aggregation kinetics. *Amyloid, prions and other proteins aggregates, 17,* 256-274.

Finsk, A.L., Calciano, L.J., Goto, Y., Kurotsu, T., & Palleros, D.R. (1994). Classification of acid denaturation of proteins: Intermediates and unfolded states. *Biochemistry, 33,* 12504-12511.

Finsy, R. (1994). Particle sizing by quasi-elastic light scattering. *Advances in Colloid and interface Science, 52,* 79-143.

Gaeta, A., & Hider, R.C. (2005). The crucial role of metal ions in neurodegeneration : the basis for a promising therapeutic strategy. *British Journal of Fharmacology, 146,* 1041-1059.

Giese, A., Levin, J, Bertch, U., & Kretzshmar, H. (2004). Effect of metal ions on the novo aggregation of full-length prion protein. *Biochemical and Biophysical Research Communications, 320,* 1240-1246.

Gimel, J.C., Durand, D., & Nicolai, T. (1993). Structure and distribution of aggregates formed after heat-induced denaturation of globular proteins. *Macromolecules, 57,* 583-589.

Goldschmidt, L., Teng, P.K., Riek, R., & Eisenberg, D. (2010) Identifying the amylome, proteins capable of forming amyloid-like fibrils. *PNAS,* vol. *107,* no. *8,* 3487-3492.

Gosal, W.S., Clark, A.H., & Ross-Murphy, S.B. (2004a). Fibrilar β-Lactoglobulin gels: Part 1. Fibril formation and structure. *Biomacromolecules, 5,* 2408-2419.

Gosal, W.S., Clark, A.H., & Ross-Murphy, S.B. (2004b). Fibrilar β-Lactoglobulin gels: Part 2. Dynamic mechanical characterization of heat-set systems. *Biomacromolecules, 5,* 2420-2429.

Gosal, W.S., Clark, A.H., & Ross-Murphy, S.B. (2004c). Fibrilar β-Lactoglobulin gels: Part 3. Dynamic mechanical characterisation of the solvent-induced systems. *Biomacromolecules, 5,* 2430-2438.

Gosal, W.S., Morten, I.J., Hewitt. E.W., Smith, D.A., Thomson, N.H., & Radford, S.E. (2005). Competing pathways determine fibril morphology in the self-assembly of β_2-microglobulin into amyloid. *Journal of Molecular Biology, 351,* 850-864.

Grudzielanek, S., Janse, R., & Winter, R. (2005). Solvational tuning of the unfolding, aggregation and amyloidogenesis of insulin. *Journal of Molecular Biology, 351,* 879-894.

Guijarro, J.I., Sunde M., Jones J.A., Campbell I.D., & Dobson C.M. (1998). Amyloid fibril formation by an SH3 domain. *Proceedings of the National Academy of Science of the United States of America, 95,* 4224-4228.

Gunasekaran, S., Ko, S., & Xiao, L. (2007). Use of whey proteins for encapsulation and controlled delivery applications. *Journal of Food Engineering, 83*, 31-40.

Haass, C., & Selkoe, D.J. (2007). Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β-peptide. *Nature Reviews – Molecular Cell Biology, 8*, 101-112.

Hamada, D., & Dobson, C.M. (2002). A kinetic study of β-Lactoglobulin amyloid fibril formation promoted by urea. *Protein Science, 11*, 2417-2426.

Harrick, N.J. (1967). Internal Reflection Spectroscopy. John Wiley & Sons Inc. p. 342.

Herberhold, H., Royer, C.A., & Winter, R. (2004). Effect of chaotropic and kosmotropic co-solvents on the pressure-induced unfolding and denaturation of proteins: a FT-IR study on Staphylococcal nuclease. *Biochemistry, 43*, 3336-3345.

Hiemenz, P.C., & Rajagopalan, R. (1997). *Principles of colloid and surface chemistry* (3rd Ed.). Marcel Dekker, Inc.

Holmberg, K., & Jönsson, B. *Surfactants and polymers in aqueous solution* (2nd Ed.). John Wiley and Sons.

Hummer, G., Garde, S., Garcia, A.E., Paulaitis, M.E., & Pratt, L.R. (1998). The pressure dependence of hydrophobic interactions is consistent with the observed pressure denaturation of proteins. *Proceedings of the National Academy of Science of the United States of America, 95*, 1552-1555.

Ikeda, S., & Morris, V.J. (2002). Fine-stranded and particulate aggregates of heat-denatures whey proteins visualized by atomic force microscopy. *Biomacromolecules 3*, 382-389.

IUPAC Compendium of Macromolecules Nomenclature (1991). Blackell Scientific Publications, Oxford, p. 48.
Available online at: http://goldbook.iupac.org/R05121.html

Jackson, M., & Mantsch, H.H. (1991). Protein secondary structure from FT-IR spectroscopy: correlation with dihedral angles from three-dimensional Ramachandran plots. *Canadian Journal of Chemistry, 69*, 1639-1642.

Jahn, T.R., & Radford, S.E. (2008). Folding versus aggregation: polypeptide conformations on competing pathways. *Archives of Biochemistry and Biophysics, 469*, 100-117.

Jarret, J., Berger, E., & Lansburry, P. (1993). The carboxy terminus of β-amyloid protein is critical for the seeding of amyloid formation: implications for the pathogenesis of Alzheimer's disease. *Biochemistry, Vol. 32*, No.18.

Javid, N., Vogtt, K., Krywka, C., Tolan, M., & Winter, R. (2007). Capturing the interaction potential of amyloidogenic proteins. *Physical Review Letters, 99*, 028101.

Jones, O.G., Decker, E.A., & McClements, D.J. (2009). Formation of biopolymer particles by thermal treatment of β-Lactoglobulin-pectin complexes. *Food Hydrodolloids*, *23*, 1312-1321.

Jones,O., Decker, E.A., & McClements, D.J. (2010). Thermal analysis of β-Lactoglobulin complexes with pectins or carrageenan for production of stable biopolymer particles. *Food Hydrocolloids, 24*, 239-248.

Juárez, J., Taboada, P., Goy-Lopez, S., Cambon, A., Madec, M-B., Yeates, S.G., & Mosquera. V. (2009). Additional supra-self assembly of human serum albumin under amyloid-like-forming solution conditions. *Journal of Physical Chemistry, 113*, 12391-12399.

Kabsh, W., & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers, Vol. 22*(12), 2577-2637.

Kanjilal, S., Taulier, N., Le Huerou, J.Y., Gindre, M., Urbach, W., & Waks., M. (2003). Ultrasonic studies of the alcohol-induced transconformation in β-Lactoglobulin: the intermediate state. *Biophysical Journal, 85*, 3928-3934.

Kavanagh, G.M., Clark A.H., & Ross-Murphy, S.B. (2000). Heat-induced gelation of globular proteins: part 3. Molecular studies on low pH β-Lactoglobulin gels. *International Journal of Biological Macromolecules, 28*, 41-50.

Kella, N.K.D., & Kinsella, J.E. (1988). Enhanced thermodynamic stability of β-Lactoglobulin at low pH. A possible mechanism. *Biochemistry Journal, 255*(1), 113-118.

Kelly, S.M., Jess, T.J., & Price, N.C. (2005). How to study proteins by circular dichroism. *Biochimica and Biophysica Acta, 1751*, 119-139.

Khare, S.D., & Dokholyan, N.V. (2007). Molecular mechanisms of polypeptide aggregation in human diseases. *Current Protein and Peptide Science, 8*, 573-579.

Khurana, R., Coleman, C., Ionescu-Zanetti, C., Carter, S.A., Krishna, V., Grover, R.K., Raja Roy, R., & Singh, S. (2005). Mechanism of Thioflavin T binding to amyloid fibrils. *Journal of Structural Biology, 151*, 229-238.

Kokalj S.J., Stoka, V., Kenig, M., Guncar, G., Turk, D., & Zerovnik, E. (2004). A central role for protein aggregation in neurodegenerative disease; mechanistic and structural studies of human stefins. *Acta Chimica Slovenica., 52*, 27-33.

Krebs, M.R.H., Bromley, E.H.C., & Donald, A.M. (2005). The binding of Thioflavin-T to amyloid fibrils: location and implications. *Journal of Structural Biology, 149,* 30-37.

Krebs, M.R.H., Devlin, G.L., & Donald, A.M. (2009). Amyloid fibril-like structure underlies the aggregate structure across the pH range for β-Lactoglobulin. *Biophysical Journal, 96,* 5013-5019.

Kumar, S., & Udgaonkar, J.B. (2011). Mechanisms of amyloid fibril formation by proteins. *Current Science, 98,* 639-656.

Lakowicz, J.R. (2006). *Principles of fluorescence spectroscopy* (3rd Ed.). Springer.

Lasagna-Reeves, C.A., Glabe, C.G., & Kayed, R. (2011). Amyloid-β annular protofibrils evade fibrilar fate in Alzheimer's Disease brain. *The Journal of Biological Chemistry*. http://www.jbc.org/cgi/doi/10.1074/jbc.M111.236257

LeVine, H. (1999). Quantification of beta-sheet amyloid fibrils structures with Thioflavin T. *Amyloid, Prions and other protein aggregates*. *Methods in enzymology series, 309,* 274-284.

LeVine, H. III (1995). Thioflavin T interaction with amyloid-sheet structures. *Amyloid, 2*(1), 1-6.

Le Bon, C., Nicolai, T., Kuil, M.E., & Hollander, J.G. (1999). Self-diffusion and cooperative diffusion of globular proteins in solution. *Journal of Physical Chemistry, 103*, 10294-10299.

Lins, L., & Brasseur, R. (1995). The hydrophobic effect in protein folding. *The FASEB Journal, 9,* 535-540.

Livney, Y.D. (2010). Milk proteins as vehicles for bioactives. *Current opinion in colloid and interface science, 15,* 73-83.

Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., & Darnell, J. (2000). Molecular Cell Biology (4th Ed.). New York: W.H. Freemen.

Lomakin, A., Chung, D.S., Benedek, G., Kirschner D.A., & Teplow, D.B. (1996). On the nucleation and growth of amyloid β-protein fibrils: detection of nuclei and quantification of rate constants. *Proceeding of the National Academy of Science of the United States of America, 93,* 1125-1129.

Loupiac, C., Bonetti, M., Pin, S., & Calmettes, P. (2006). β-Lactoglobulin under high pressure studied by small-angle neutron scattering. *Biochimica et Biophysica Acta, 1764,* 211-216.

Luthi-Peng, Q.Q., & Puhan, Z. (1999). Determination of protein and casein in milk by the fourth derivative UV spectrophotometry. *Analytical Chimica Acta, 393*(1-3), 227-234.

Maa, Y-F., & Hsu, C.C. (1997). Protein denaturation by combined effect of shear and air-liquid interface. *Biotechnology and Bioengineering, 54*(6), 503-512.

Majhi, P.R., Ganta, R.R., Vanam, P.P., Seyrek, E., Giger, K., & Dubin, P.L. (2006). Electrostatically driven protein aggregation: β-Lactoglobulin at low ionic strength. *Langmuir, 22,* 9150-9159.

Makhatadze, G.I., & Privalov, P.L. (1992). Protein interactions with urea and guanidinium chloride: A calorimetric study. *Journal of Molecular Biology, 226*(2), 491-505.

Manno, M., Caparro, E.F., Podesta, A., Bulone, D., Carrotta, R., Mortorana, V., Tiana, G., & San Biagio, P.L. (2007). Kinetics of different processes in human insulin amyloid formation. *Journal of Molecular Biology*, *366*, 258-274.

Masson, P., & Cléry, C. (1996). Pressure-induced molten globule states of proteins. *High pressure Bioscience and Biotechnology*, *13,* 117-126.

Wedemeyer, W.J., Welker, E., Narayan, M., & Scheraga, H.A. (2000). Disulfide bonds and protein folding. *Biochemistry, 39*(15), 4207- 4216.

Meyers, R.A. (1995). Molecular biology and biotechnology: a comprehensive desk reference (Wiley-VCH Editions). *Circular dichroism in protein analysis* by Price, N.C. (pp. 179-185).

Monahan, F.J., German, J.B., & Kinsella, J.E. (1995). Effect of pH and temperature on the protein unfolding and thiol/disulfide interchange reactions during heat-induced gelation of whey proteins. *Journal of Agricultural and Food Chemistry, 43*, 46-52.

Monera, O.D., Kay, C.M., & Robert, S.H. (1994). Protein denaturation with guanidine hydrochloride or urea provides a different estimate of stability depending on the contributions of electrostatic interactions. *Protein Science, 3*, 1984-1991.

Mousavi, S., Chobert, J.M., Bordbar, A.K., & Haertlé T. (2008). Ethanol effect of the structure of β-Lactoglobulin B and its ligand binding. *Journal of Agricultural and Food Chemistry, 56*, 8680-8684.

Murphy, R.M. (2007). Kinetics of amyloid formation and membrane interaction with amyloidogenic proteins. *Biochimica et Biophysica Acta, 1768,* 1923-1934.

Naiki, H., Higuchi, K., Hosokawa, M., & Takeda, T. (1989). Fluorometric determination of amyloid fibrils in vitro using the fluorescent dye Thioflavin T. *Analitical Biochemistry, 177*(2), 244-249.

Necula, M., Kayed, R., Milton, S., & Glabe, C.G. (2007). Small molecule inhibitors of aggregation indicate that amyloid β oligomerization and fibrillation pathways are independent and distinct. *The Journal of Biological Chemistry, 282*(14), 10311-10324.

Nilson, M.R. (2004). Techniques to study amyloid fibril formation in vitro. *Methods 34*, 151-160.

Nielsen, L., Khurana, R., Coats, A., Frokjaer, S., Brage, J., Vyas, S., Uversky, V.N., & Fink, A.L. (2001). Effect of environmental factors on the kinetics of insulin fibrils formation: elucidation of the molecular mechanism. *Biochemistry, 40*, 6036-6046.

Pace, C.N., Trevino, S., Prabhakaran, E., & Sholtz, J.M. (2004). Protein structure, stability and solubility in water and other solvents. *The Royal Society, 359*, 1225-1235.

Pain R.H. (1994) Mechanisms of Protein Folding (1st Ed.). *The contribution of the molten globule model* (Chapter 3). Oxford University Press.

Panick, G., Malessa, R., & Winter, R. (1999) Differences between the pressure and temperature induced denaturation of β-Lactoglobulin A, B and AB monitored by FT-IR spectroscopy and small-angle X-ray scattering. *Biochemistry, 38*(20), 6512-6519.

Parker, R., Noel, T.R., Brownsey, G., Laos, K., & Ring, S.G. (2005). The non-equilibrium phase and glass transition behaviour of β-Lactoglobulin. *Biophysical Journal, 89*, 1227-1236.

Pedersen, J.S., & Otzen, D.E. (2008). Amyloid – a state in many guises: survival of the fittest fibril fold. *Protein Science, 17*, 2-10.

PerkinElmer – Life and Analytical Science (2005). Technical Note for FT-IR Spectroscopy-Attenuated Total reflectance (ATR). Retrieved from www.perkinelmer.com

Permyakov, E.A., Morozova, L.A., & Burstein, E.A. (1985). Cation binding effects on the pH, thermal and urea denaturation transitions in α-Lactalbumin. *Biophysical Chemistry, 21*(1), 21-31.

Permyakov, E.A., Shnyrov, V.L., Kalinichenko, L.P., Kuchar, A. Reyzer, I.L., & Berliner, L.J. (1991). Binding of Zn (II) ions to α-Lactalbumin. *Journal of Protein Chemistry, 10*(6), 577-584.

Permyakov, E.A., & Berliner, L.J. (2000). α-Lactalbumin: structure and function. *FEBS Letters, 473*, 269-274.

Pessen, H., Kumosinski, T.F., & Timasheff, S.N. Small-angle X-ray scattering. *Methods in Enzimology, 27*, 151-209.

Petsko, G.A., & Ringe, D. (2004). Protein structure and function. *Chapter 1: From sequence to structure*. New Science Press Ltd.

Price, N.C., & Nairn, J. (2009). Exploring Proteins: a student's guide to experimental skills and methods (1st Ed.). *Chapter 1: Conceptual toolkit: the molecular principles for understanding proteins*. Oxford University Press.

Provencher, S.W., & Glockner, J. (1981). Estimation of globular proteins secondary structure from circular dichroism. *Biochemistry, 20*, 33-37.

Ptitsyn, O.B., Bychkova, V.E., & Uversky V.N. (1995). Kinetic and equilibrium of folding intermediates. *Philosophical Transactions of the Royal Society of London. Series B-Biological Sciences, 348*(1323), 35-41.

Ptitsyn, O.B., Pain, R.H., Semisotnov, G.V, Zerovnik, E., & Razgulyaez, O.I. (1990). Evidence of a molten globule state as a general intermediate in protein folding. *FEBS Letters, 262*(1), 20-24.

Pusey, P.N. (2002). Neutrons, X-Rays and Light. P. Linder and Th. Zemb (Eds.). *Chapter 1: Introduction to Scattering Experiments* (pp. 3-21) and *Chapter 9: Dynamic light scattering* (pp. 203-220).

Qin, B.Y., Bewley, M.C., Creamer, L.K., Baker, H.M., Baker, E.N., & Jameson, G.B. (1998). Structural basis of the Tanford transition of bovine β-Lactoglobulin. *Biochemistry, 37*, 14014-14023.

Rajendran, S., & Prakash, V. (1988). Association-dissociation and denaturation-renaturation of high-molecular-weight protein: Carmin from safflower seed (*Carhamus tinctorius* L.) in alkaline solution. *Journal of Protein Chemistry, 7*(6), 689-711.

Ramos, C.H.I., & Baldwin, R.L. (2002). Sulfate anion stabilization of the native ribonuclease A both by anion binding and by the Hofmeister effect. *Protein Science, 11*, 1771-1778.

Reddy, T.T., Lavenant, L, Lefebvre, J., & Renard, D. (2006). Swelling behaviour and controlled release of theophylline and sulfamethoxazole drugs in β-Lactoglobulin protein gels obtained by phase separation in water/ethanol mixture. *Biomacromolecules, 7*, 323-330.

Redecke, L., von Bergen, M., Clos, J., Konarev, P.V., Svergun, D.I., Fitteschen U.E.A., Broekaert, J.A.C., (…), & Betzel, C. (2007). Structural characterization of β-sheeted oligomers formed on the pathway of oxidative prion protein in vitro. *Journal of Structural Biology, 157*, 308-320.

Renard, D., Lefebvre, J., Robert, P., Llamas, G., & Dufour, E. (1999). Structural investigation of β-Lactoglobulin gelation in ethanol/water solutions. *International Journal of Biological Macromolecules*, *26*, 35-44.

Renard, D., Robert, P., Garnier, C., Dufour, E., & Lefebvre, J. (2000). Gelation by phase separation in the whey protein system: in-situ kinetics of aggregation. *Journal of Biotechnology*, *79*, 231-244.

Robert, P., Lavenant, L., & Renard, D. (2002). Infrared two-dimensional correlation and principal component analyses applied to β-Lactoglobulin aggregation in water-ethanol solution. *Applied Spectroscopy*. *Vol. 56*. N^{o} 9.

Roth, C.M., Neal, B.L., & Lenhoff, A.M. (1996). Van der Waals interactions involving proteins. *Biophysical Journal, 70,* 977-987.

Sakuno, M.M., Matsumoto, S., Kawai S., Taihei, K., & Matsumura, Y. (2008). Adsorption and structural changes of β-Lactoglobulin at the diacylglycerol-water surface. *Langmuir, 24*, 11483-11488.

Sasahara, K., Yagi, H., Sakai, M., Naike, H., & Goto, Y. (2008). Amyloid nucleation triggered by agitation of β_2-Microglobulin under acidic and neutral pH conditions. *Biochemistry, 47*, 2650-2660.

Sawyer, L., & Kontopidis, G. (2000). The core lipocalin, bovine β-Lactoglobulin. *Biochimica and Biophysica Acta, 1482*, 136-148.

Semenyuk, A.V., & Svergun, D.I. (1991). GNOM – a program Packaged for small-angle scattering data processing. *Journal of Applied Crystallography, 24*, 537-540.

Schokker, E.P., Singh, H., & Creamer, L.K. (2000). Heat-induced aggregation of β-lactoglobulin A and B with α-lactalbumin. *International Dairy Journal*, *10*, 843-853.

Shpigelman, A., Israeli, G., & Livney, Y.D. (2010). Thermally induced protein-polyphenol co-assemblies: β-Lactoglobulin-based nanocomplexes as protective nanovehicles for EGCG. *Food Hydrocolloids, 24*, 735-743.

Sipe, J.D., & Cohen, A.S. (2000). Review: history of the amyloid fibril. *Journal of Structural Biology, 130*, 88-98.

Smith, C.A., & Wood, E.J. (1992). Molecular and cell biochemistry – Biological molecules (2nd Ed.). Chapman and Hall, p. 61.

Song, K.B., & Damodaran, S. (1991). Influence of electrostatic forces on the adsorption of succinylated β-Lactoglobulin at the air-water interface. *Langmuir, 7*, 2737-2742.

Stadtman, E.R.. & Levine, R.L. (2003). Free radical-mediated oxidation of free amino acids and amino acid residues in proteins. *Amino Acids, 25*, 207-218.

Surewicz, W.K., & Mantsch, H.H. (1988). New insight into protein secondary structure from resolution-enhanced infrared-spectra. *Biochimica and Biophysica Acta*, *952*(2), 115-130.

Svergun, D. I., Petoukhov, M. V., & Koch, M.H.J. (2001). Determination of domain structure of proteins from X-ray solution scattering. *Biophysical Journal*, *80*, 2946-2953.

Takata, S., Norisuye, T., Tanaka, N., & Shibayama, M. (2000). Heat-induced gelation of β-Lactoglobulin. 1. Time-resolved dynamic light scattering. *Macromolecules, 33*(15), 5470-5475.

Tayyab, S., Sharma, N., & Khan, M.M. (2000). Use of domain specific ligands to study urea-induced unfolding of bovine serum albumin. *Biochemical and Biophysical Research Communications, 277*, 83-88.

Tian, Z.M., Wan, M.X., Wang, S.P., & Kang, J.Q. (2004). Effects of ultrasound and additives on the function and structure of trypsin. *Ultrasonics Sonochemistry, 11*, 399-404.

Trexler, A.J., & Nilsson, M.R. (2007). The formation of amyloid fibrils from proteins in the lyzozyme family. *Current Protein and Peptide Science, 8*(6), 537-557.

Uversky, V., Narizhneva, N., Kirschstein, S.O., Winter, S., & Löber G. (1997). Conformational transitions provoked by organic solvents in β-Lactoglobulin: Can a molten globule like intermediate be induced by the decrease in dielectric constant? *Folding and design, 2*, 163-172.

Van Stokkum, I.H.M., Spoelder, H.J.W., Bloemendal, M., Van Grondelle, R., & Groen, F.C.A. (1990). Estimation of protein secondary structure and error analysis from circular dichroism spectre. *Analytical Biochemistry, 191*(15), 110-118.

Venkatachalam, C.M. (1968). Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of the three linked peptide units. *Biopolymers, 6*(10), 1425-1436.

Veprintsev, D.B., Permyakov, S.E., Permyakov, E.A., Rogue, V.V., Catherin, K.M., & Berliner, L.J. (1997). Cooperative thermal transitions of bovine and human apo-α-Lactalbumins: evidence for a new intermediate state. *FEBS Letter, 412*, 625-628.

Verheul, M., Pedersen, J.S., Roefs, S.P.F.M., & Fruif, K.G. (1999). Association behaviour of native β-Lactoglobulin. *Biopolymers, 49*, 11-20. 1999.

Vernaglia, B.A., Huang, J., & Clark, E.D. (2004). Guanidine hydrochloride can induce amyloid fibril formation from egg-white lyzozyme. *Biomacromolecules, 5*, 1362-1370.

Vogtt K., Javid N., Álvarez E, Sefcik J., & Bellissent-Funel M.C. (2011). Tracing nucleation pathways in protein aggreagtion by using small angle sscattering methiods. DOI: 10.1039/c0sm00978d.

Volkov, V.V., & Svergun, D.I. (2003). Uniqueness of ab initio shape determination in small-angle scattering. *Journal of Applied Crystallography, 36*, 860-864.

Walsh, D.M., & Selkoe, D.J (2004). Oligomers on the brain: the emerging role of soluble protein aggregates in neurodegeneration. *Protein and Peptide Letters, 11(3)*, 213-228(16).

Wang, W., Nema, S., & Teagarden, D. (2010). Protein aggregation – pathways and influencing factors. *International Journal of Pharmaceutics, 390*, 89-99.

Whittaker, S.B., Spence, G.R., Gunter Grossman, J., Radford, S.E., & Moore, G.R. (2007). NMR analysis of the conformational properties of the trapped on-pathway folding intermediate of the bacterial immunity Protein Im7. *Journal of Molecular Biology, 366*, 1001-1015.

Wong, K-P., & Tanford, C. (1973). Denaturation of bovine carbonic anhydrase B by guanidine hydrochloride. *The Journal of Biological Chemistry, 248*(24), 8518-8523.

Wood, S.J., Wypych, J., Steavenson, S., Louis, J-C., Citron, M., & Biere, A.L. (1999). α–Synuclein fibrilogenesis is nucleation dependent. *The Journal of Biological Chemistry, 274*(28), 19509-19512.

Yamin, G., Munishkina, L.A., Karymov, M.A., Lyubchenko, Y.L., Uversky, V. N., & Fink, A.L. (2005). Forcing nonamyloidogenic β-Synuclein to fibrillate. *Biochemistry, 44*, 9096-9107.

Yang, J., Power, J.R., Clark, S., Dunker, A.K., & Swanson, B.G. (2001). β-Lactoglobulin molten globule induced by high pressure. *Journal of Agricultural and Food Chemistry, 49*, 3236-3243.

Yoshida, K., Yamaguchi, T., Osaka, N., Endo, H., & Shibayama, M. (2010). A study of alcohol-induced gelation of β-Lactoglobulin with small-angle neutron scattering, neutron spin echo and dynamic light scattering. *Physical Chemistry Chemical Physics, 12*, 3260-3269.

Zandomeneghi G., Krebs, M.R.H., Mccammon M.G., & Fandrich M. (2004). FTIR reveals structural differences between native β-sheets proteins and amyloid fibrils. *Protein Science, 13*, 3314-3321.

APPENDIX

In the following pages, the reader will be able to find the complete appendices set mentioned through the present work; from APPENDIX A to APPENDIX M. Theoretical explanations and mathematical treatment of result are included here. Tables and graphs in this block collect all those results that were not included in the main core of the thesis either because its extension or because they were not conclusive to be part of the main discussion. Table I shows a brief relation of the appendices content following hereafter for better guide.

Table I – Appendix content.

APPENDIX	CONTENT
APPENDIX A, pp. 173	Online server **DICHROWEB** – fitting of CD spectra in the far UV region
APPENDIX B, pp. 179	**DLS** – fit and error of the fit
APPENDIX C, pp. 183	**ThT- binding fluorescence** spectra for aqueous solutions of β-Lg
APPENDIX D, pp. 191	**Apparent R_h** and normalised **ACF** for aqueous solution of β-Lg
APPENDIX E, pp. 202	Normalised **ACF** for samples containing different percentages (by volume) of ethanol and the **apparent R_h**
APPENDIX F, pp. 208	Aggregation kinetics of β-Lg induced by ethanol – **ThT-binding fluorescence spectra vs. incubation time**
APPENDIX G, pp. 219	**UV absorbance**
APPENDIX H, pp. 220	**CD** – near UV and far UV spectra. Quantitative analysis of far UV spectra. Analysis in the far UV for shorter periods of time of samples of 1 mg/ml of β-Lg
APPENDIX I, pp. 230	**ATR-IR** - spectra and quantitative analysis
APPENDIX J, pp. 235	**Apparent R_h** calculated for every repetition of 10 days-systems in the presence of ethanol
APPENDIX K, pp. 238	**ACF** for every repetition of 10 days-systems in the presence of ethanol
APPENDIX L, pp. 242	**ACF** obtained for 1 mg/ml and 4 mg/ml of protein in the presence of ethanol measured during short periods of time and **apparent R_h**
APPENDIX M, pp. 248	**SAXS** results

APPENDIX A

DICHROWEB is the online server used for the quantitative analysis of far UV spectra obtained from CD measurements. DICHROWEB is hosted at Birkbeck Collage, University of London, U.K. and it can be reached in the following link: http://dichroweb.cryst.bbk.ac.uk/html/process.shtml having been online since 2001. The server facilitates analyses using five popular algorithms (CDSSTR, CONTIN, SELCON3, VARSLC and K2d) and seven different reference protein-data bases. Far UV spectra need to be converted to text files before being introduced in DICHROWEB input data page. This can be done when the file is saved using the software in Jasco-810 instrument used for CD measurements (see section 2.2.3.2 for full detail of this instrument).

The software package chosen to deconvolute far UV spectra was CONTIN, a Fortran IV program for inverting noisy linear operation equations developed by Provencher and Glockner in 1982 and improved by Van Stokkum later (Provencher and Glockner, 1981; Van Stokkum et al., 1990). This software analyses the far UV spectrum of the globular protein under study as a linear combination of the CD spectra of the reference proteins contained in the reference dataset. The secondary structure of these proteins is already known by X-ray analysis. Thus, they are used as reference. Different datasets can be also chosen, depending on the protein under study and on the range of input data. Set 3 (optimised for 185-240 nm) was the reference data set used in this project. This set contains 38 proteins whose far UV spectrum is similar to that for the native state of β-Lg. Results output page in DICHROWEB using the above analysis programme and reference set, gives results in a variety of formats: a short summary of the best deconvolution (listing only the secondary structure composition and total content), a longer report including the deconvolution details and the normalised root mean square deviation (NRMSD) fit parameters, and finally a graph of the experimental and fitted spectra, with the difference graph superimposed.

The NRMSD represents the goodness-of-the-fit defined as the square root of the sum of the difference-square between the experimental and calculated ellipticities (θ_{exp} and θ_{cal} respectively) for all wavelengths analysed over the sum of the square experimental ellipticities:

$$\left[\frac{\sum(\theta_{\exp}-\theta_{cal})^2}{\sum(\theta_{\exp})^2}\right]^{1/2}$$

This parameter is important to compare the experimental and calculated data from the software and can be used to judge the quality of results. Generally, it is admitted that a low value of NRMSD (< 0.1) is a necessary (although not sufficient) condition to confirm that good results has been obtained.

The following protocol indicates the steps followed from the far UV spectrum measurement of a random sample up to the calculation of the percentages of α-helix, β-sheet, turns and random coil present in the protein under study:

1. CD measurement in the far UV region using a Jasco-810. Manipulation of the spectra with the aims of subtracting the blank, smoothing, correcting to concentration and obtaining final units as molar ellipticity (deg.cm^2.dmol^{-1}). Save the final file both as *.jws* (software in the instrument) and as text *(.txt*).

   E.g.: *SM_blg_ethanol_1mgml_DAY1_far_minusblank_molE.jws*
   SM_blg_ethanol_1mgml_DAY1_far_minusblank_molE.txt

2. Input data:

   Go to http://dichroweb.cryst.bbk.ac.uk/html/process.shtml (online) and to "*Input data*" on the menu in the left hand side of the website. Introduce the userID and password (obtained upon completion of an application form) and the rest of the fields until total completion. When browsing the data to be analysed, it is important to remember that the file has to be introduced as text,

i.e., the one with *.txt* extension (no that with *.jws* extension). The parameters used are always the same. For each analysis, only the "*Protein name*" and the "*File location*" may be changed.

E.g.: **Input File Detail**

Protein name: *BLG*

File location:

*SM_blg_ethanol_1mgml_DAY1_far_minusblank_molE.**txt***

About the Data File

File Format: *JASCO 1.50 (with preview)*

Input Unit: *Mean residue ellipticity*

Initial Wavelength (nm): *185*

Final Wavelength (nm): *260*

Wavelength step (nm): *0.2*

Lowest data point to use in the analysis: *185*

Choice of Methods

Analysis Programme: *CONTIN*

Reference Set: Set *3 (Optimised for 185-240)*

Advanced Options

Optional Scaling Factor: *1.0*

Output Options

Output Units: *Mean residue ellipticity*

3. Output data:

Click on "submit" at the bottom of the input data page. In the next displayed page, check the parameters and click on "continue" if everything is okay. Click on "continue" in the next page and select the option of showing the results. For the calculation of the secondary structure content of the protein

only the best deconvolution ("**Compact Results**") is used. However, the goodness-of-fit appears in the "**Extended Results**" together with the deconvolution details. The output data shown as a "**Plot**" is also useful to see the correspondence between the experimental and calculated spectra from the fit and to judge the quality of results. The output data used for further calculation is shown in the example.

E.g.: (i) **Extended results** - showing the compact results table and NRMSD value as the relevant information required from the fitting.

BLG

Provencher & Glockner Method: Reference set 3

NRMSD:0.036

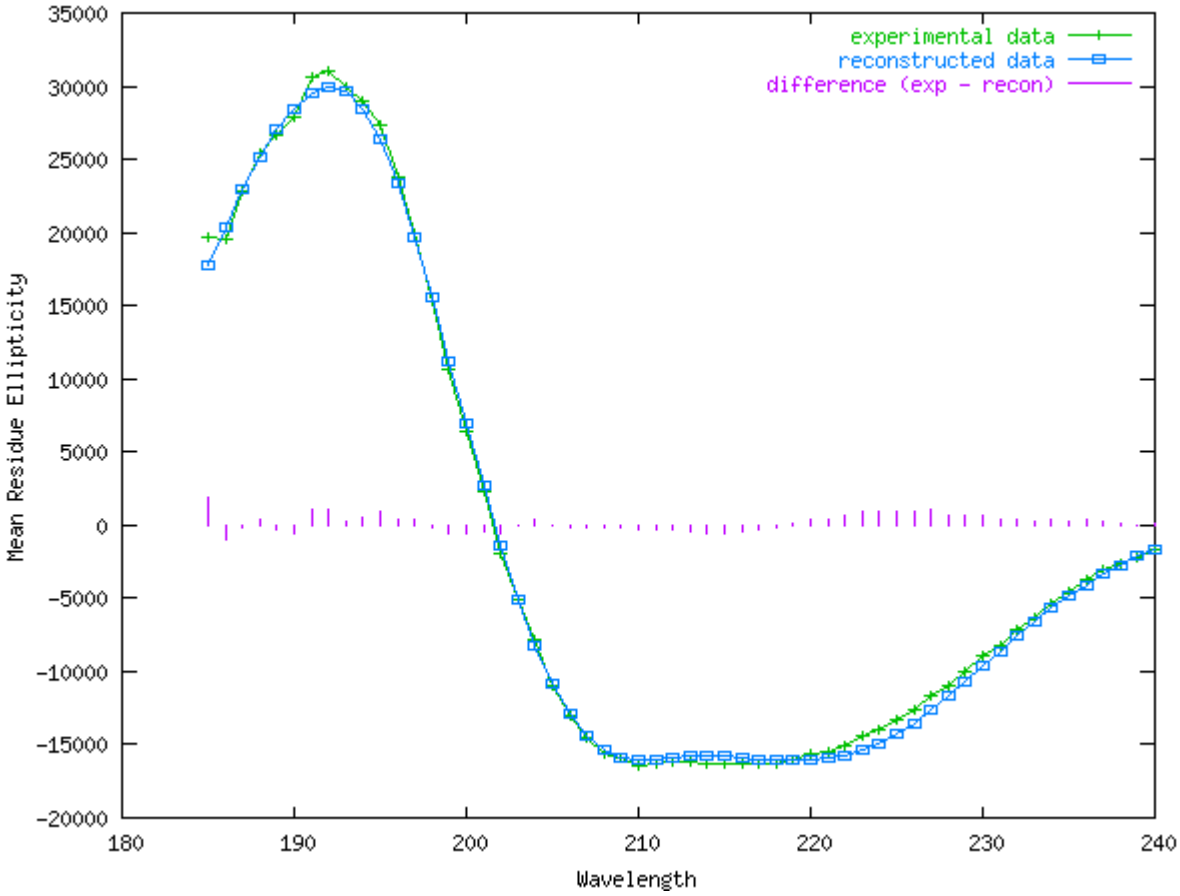
Helix segments per 100 residues: 4.056 **Strand segments per 100 residues: 3.003**

Ave helix length per segment: 10.976 **Ave strand length per segment: 4.370**

Result	Helix1	Helix2	Strand1	Strand2	Turns	Unordered	Total
1	0.289	0.163	0.079	0.063	0.154	0.253	1.001
2	0.283	0.162	0.071	0.060	0.164	0.260	1

1. Closest matching solution with all proteins
2. Average of all matching solutions

(iii) **Plot**



4. Percentage calculation of α-helix, β-sheet, turns and random coil (unordered):

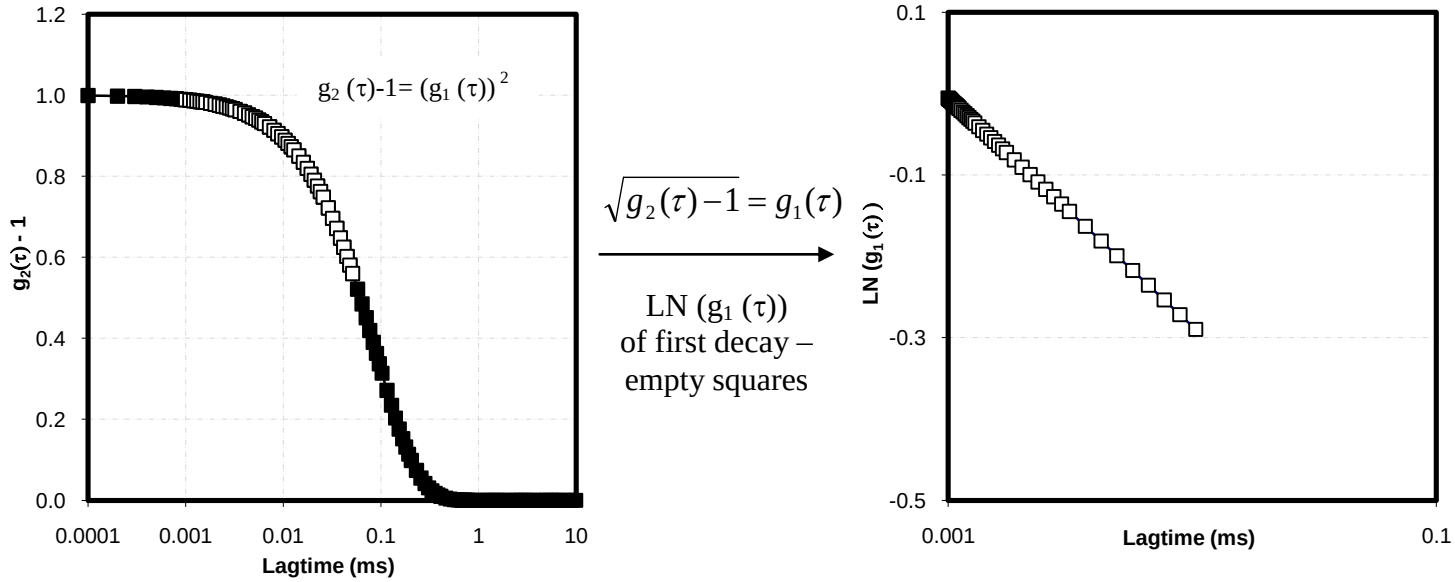
   Helix1 and Helix2 values in the compact results table were added for results 1 and 2 given a single value for "helix" for each result. These values were again added and divided by two to calculate the arithmetic mean. This gives the fraction of α-helix in the protein. By multiplying this fraction by 100 the percentage of α-helix is obtained. The same calculation was done for the "strand" values. For turns and random coil, the calculation is shorter since the output data is only one value for each result 1 and 2. Table A.1 shows this calculation.

Table A.1 - Calculation of the percentages of the different types of secondary structure present in the protein from far UV quantitative analysis.

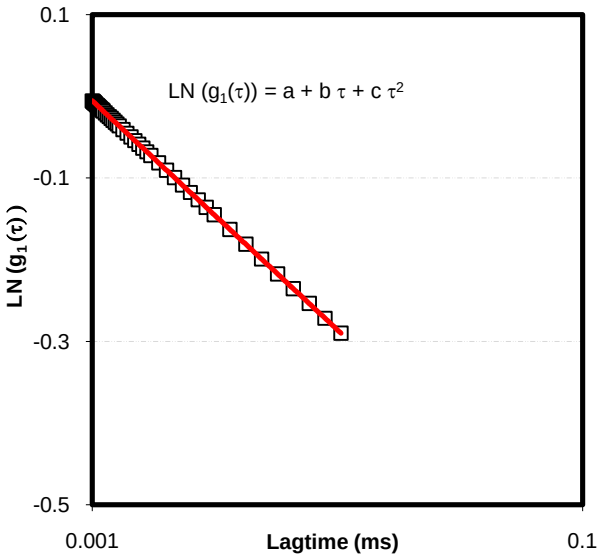
Result	**Helix1**	**Helix2**	**Helix1 + Helix2**	**Mean of $(Helix1+Helix2)_1$ + $(Helix1+Helix2)_2$**	**% α-Helix**
1	0.289	0.163	0.452	0.4485	44.85
2	0.283	0.162	0.445		
Result	**Strand1**	**Stran2**	**Stran1 + Strand2**	**Mean of $(Strand1+Strand2)_1$ + $(Strand1+Strand2)_2$**	**% β-Sheet**
1	0.079	0.063	0.142	0.1365	13.65
2	0.071	0.060	0.131		
Result	**Turns**			**Mean of $(Turns)_1 + (Turns)_2$**	**% Turns**
1	0.154			0.159	15.90
2	0.164				
Result	**Unordered**			**Mean of $(Unordered)_1 + (Unordered)_2$**	**% Unordered (random coil)**
1	0.253			0.2565	25.65
2	0.260				

APPENDIX B

To calculate the apparent hydrodynamic radius from the autocorrelation function obtained by DLS analysis the cumulant method is applied. The output signal given by ALV / CG3 is $g_2(\tau)$-1 which can be transformed into $g_1(\tau)$ according to DLS theory.



The first decay of the autocorrelation function (as LN (g_1 (τ)) was fitted to a second order polynomial function using Microsoft Excel package:



From the first order coefficient (b) the decay rate (Γ) is obtained, whose units are 1/ms. The error in Γ (σ_Γ) is giving by the formula:

$$\sigma_\Gamma = \sigma \cdot \sqrt{\frac{1}{\sum(\tau_i - \bar{\tau})^2}}$$

where

- σ is the mean error of the fitting
- τ_i are the individual values of the lagtime (in ms)
- $\bar{\tau}$ is the average of every lagtime (in ms)

The mean error of the fitting (σ) is calculated with Microsoft Excel package by:

$$\sigma = \left(\frac{1}{N-2} \sqrt{\sum (y - y_{pred})^2} \right)^{\frac{1}{2}}$$

where

- y corresponds with the experimental data – in this case LN ($g_1(\tau)$) obtained from the raw data $g_2(\tau)$-1 for each lagtime
- y_{pred} is the LN($g_1(\tau)$) values predicted by the second order polynomial function obtained by the fit. The difference y-y_{pred} is usually called "the residuals"
- N is the number of points, in this case, the number of points fitted corresponding with values of lagtime for the first decay of the autocorrelation function.

The average of the lagtime ($\bar{\tau}$) is calculated by the general formula:

$$\bar{\tau} = \frac{\sum \tau_i}{N}$$

where τ_i and N has the same meaning as above. Therefore by the second order polynomial fitting the decay rate and its error is obtained, both in units (1/ms):

$$\Gamma \pm \sigma_\Gamma \; [=] \; (1/ms)$$

The decay rate is related to the diffusion coefficient (*D*) of the particles in solution (Brownian motion) by the equation

$$\Gamma = D \cdot q^2$$

where *q* is the scattering vector characteristic for each experiment since it only depends on physical parameters:

$$q = \frac{4 \cdot \pi \cdot \eta}{\lambda} \cdot \sin\left(\frac{\theta}{2}\right)$$

where

- is the refractive index of the solvent
- λ is the wavelength of the laser
- θ is the angle at which the scattered light is measured

The hydrodynamic radius is then calculated using the Stoke-Einstein equation:

$$R_{hi} = \frac{k_B \cdot T}{6 \cdot \pi \cdot \mu \cdot D}$$

Where

- k_B is the Boltzmann constant
- *T* is the absolute temperature
- μ is the solvent viscosity at the given temperature

- *D* is the diffusion coefficient as above

The standard error in R_{hi} (σ_{Rhi}) is calculated by applying the error propagation formula since R_h and Γ are related:

$$\sigma_{R_{hi}} = \left(\frac{\partial(R_{hi})}{\partial \Gamma} \right) \cdot \sigma_{\Gamma}$$

Where:

- $\left(\frac{\partial(R_{hi})}{\partial \Gamma} \right)$ is the derivative of Stoke-Einstein's equation respect to Γ
- σ_{Γ} is the standard error in Γ calculated as indicated above

The hydrodynamic radius for every single measurement (*i*) is given then by:

$$R_{hi} \pm \sigma_{R_{hi}} \ (nm)$$

In this work, index *i* corresponds to different repetitions named as A, B and C thorough this thesis. Since experiments were done in triplicate, the final hydrodynamic radius for all repetitions under the same physical conditions is given as the arithmetic mean of all R_{hi} values obtained, together with the standard deviation due to different repetitions:

$$R_h \pm \sigma_{R_h} \ (nm)$$

APPENDIX C

ThT- binding fluorescence intensities at 477 nm for the systems under study (1, 4, 10 and 40 mg/ml of protein in 50% (v/v) ethanol and 10 mM phosphate buffer – pH 7) incubated during 10 days at room temperature (including results for the native protein) are shown in Tables C.1, C.2, C.3 and C.4 (for 1, 4, 10 and 40 mg/ml of β-Lg respectively). Different repetitions are named with an alphabet letter followed by the protein concentration of the system; hence "system A1" indicates repetition "A" of the system with "1 mg/ml of β-Lg." Variability among repetitions is shown by the standard deviation (SD) of the three values of fluorescence intensity for the three repetitions aforementioned. ThT emission spectra were recorded just after adding ethanol to the protein (labelled as "Day 0") and after 1, 3, 7 and 10 days of incubation at 22ºC.

Table C.1 - ThT-binding fluorescence intensities [a.u] at 477 nm for every 10 days-system with 1 mg/ml of protein concentration.

System:	**A1**	**B1**	**C1**	**SD**
Native protein	6.11	6.83	5.96	0.46
Day 0	5.48	7.29	5.97	0.93
Day 1	23.4	29.6	24.0	3.47
Day 3	36.5	38.7	34.0	2.36
Day 7	38.8	43.5	37.0	3.38
Day 10	40.0	42.1	34.9	3.70

Table C.2 - ThT-binding fluorescence intensities at 477 nm for every 10 days-system with 4 mg/ml of protein concentration.

System:	**A4**	**B4**	**C4**	**SD**
Native protein	33.6	34.6	35.3	0.87
Day 0	40.2	41.0	44.9	2.50
Day 1	180.4	209.9	211.9	17.6
Day 3	197.6	212.2	227.8	15.1
Day 7	241.9	214.4	225.3	13.6
Day 10	217.0	200.8	203.6	8.65

Table C.3 - ThT-binding fluorescence intensities at 477 nm for every 10 days-system with 10 mg/ml of protein concentration.

System:	**A10**	**B10**	**C10**	**SD**
Native protein	60.4	59.5	58.6	0.92
Day 0	175.6	194.7	149.6	22.6
Day 1	346.1	369.6	333.9	18.2
Day 3	376.6	385.0	367.7	8.68
Day 7	365.8	403.7	359.6	23.9
Day 10	383.5	423.9	377.0	25.4

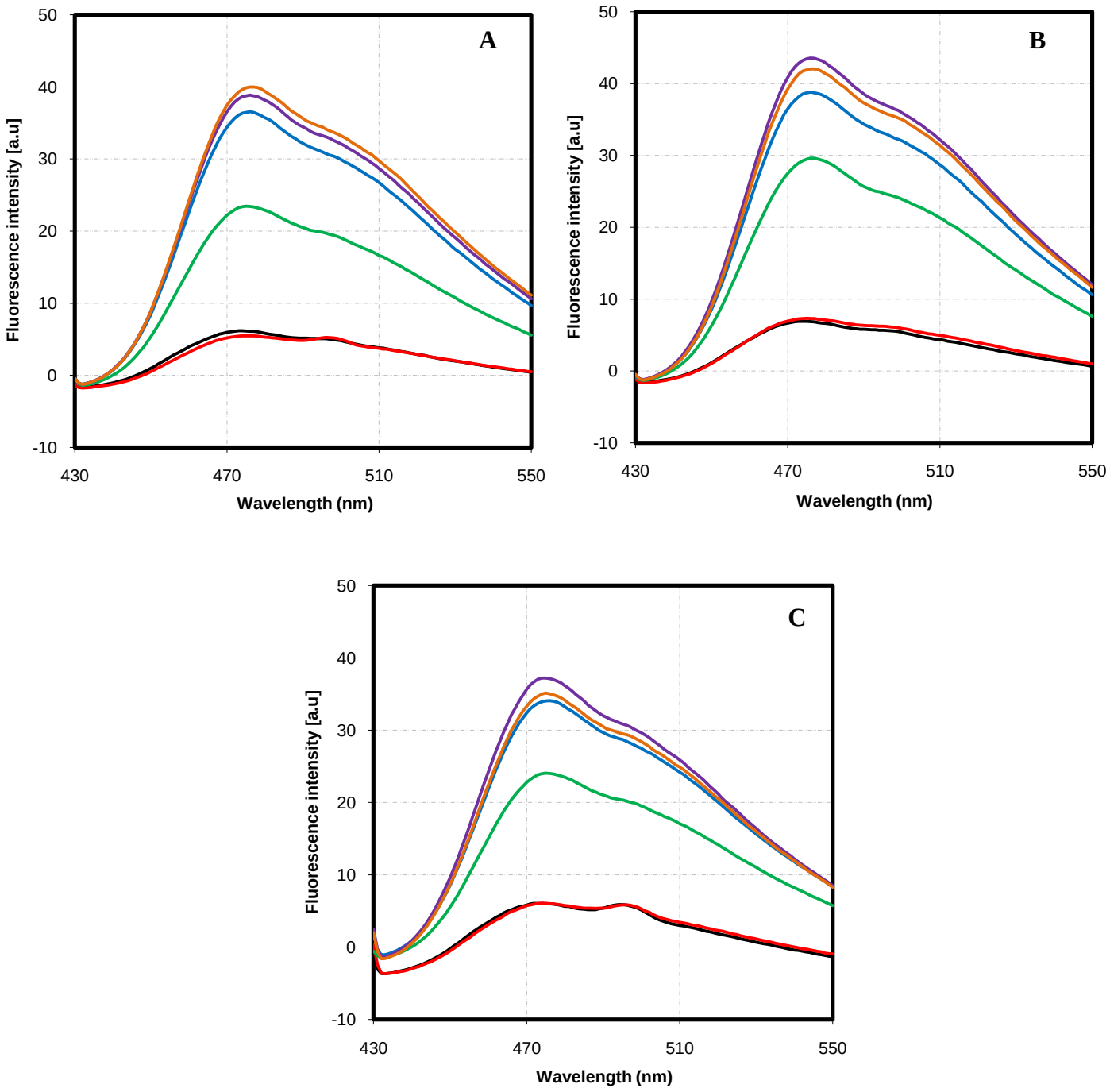
Table C.4 - ThT-binding fluorescence intensities at 477 nm for every 10 days-system with 40 mg/ml of protein concentration.

System:	**A40**	**B40**	**C40**	**SD**
Native protein	205.0	212.5	229.0	12.3
Day 0	810.3	861.4	859.5	28.9
Day 1	820.1	959.9	973.7	84.9
Day 3	880.1	965.9	924.0	42.8
Day 7	877.6	922.0	999.9	61.9
Day 10	888.6	965.1	918.9	38.5

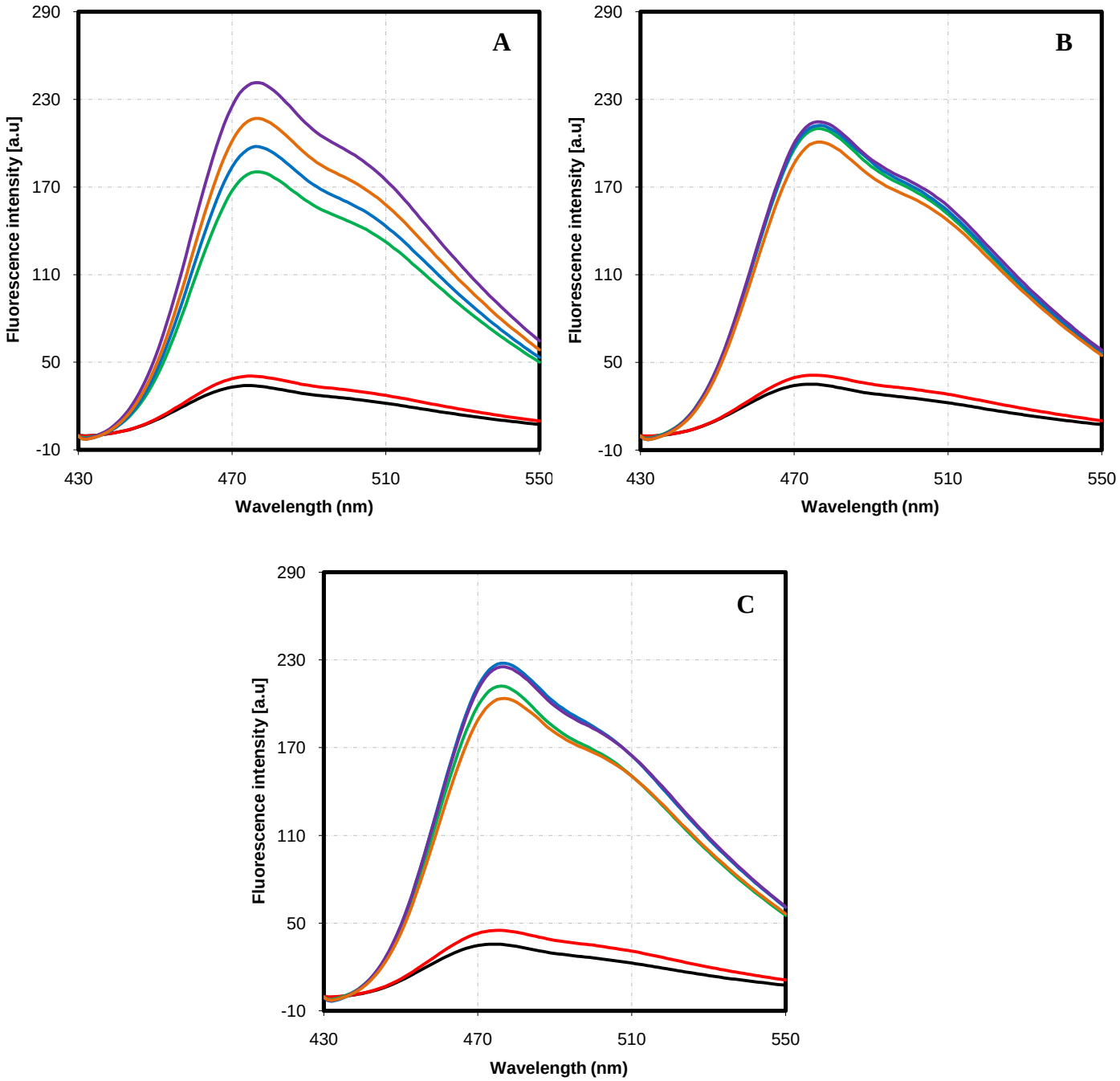
The following graphs represent the fluorescence spectra obtained for every concentration of protein and repetition during 10 days of incubation at room temperature (22°C) including the fluorescence spectra of the aqueous solutions in 10 mM buffer - native state of the protein. The legend in Graphs from C.1 to C.5 is as follows:

- Black line = native state of the protein – aqueous solution of β-Lg in 10 mM buffer.
- Red line = day 0 – after addition of ethanol.
- Green line = after 1 day of incubation at 22°C
- Blue line = after 3 days of incubation at 22°C
- Purple line = after 7 days of incubation at 22°C
- Orange line = after 10 days of incubation at 22°C

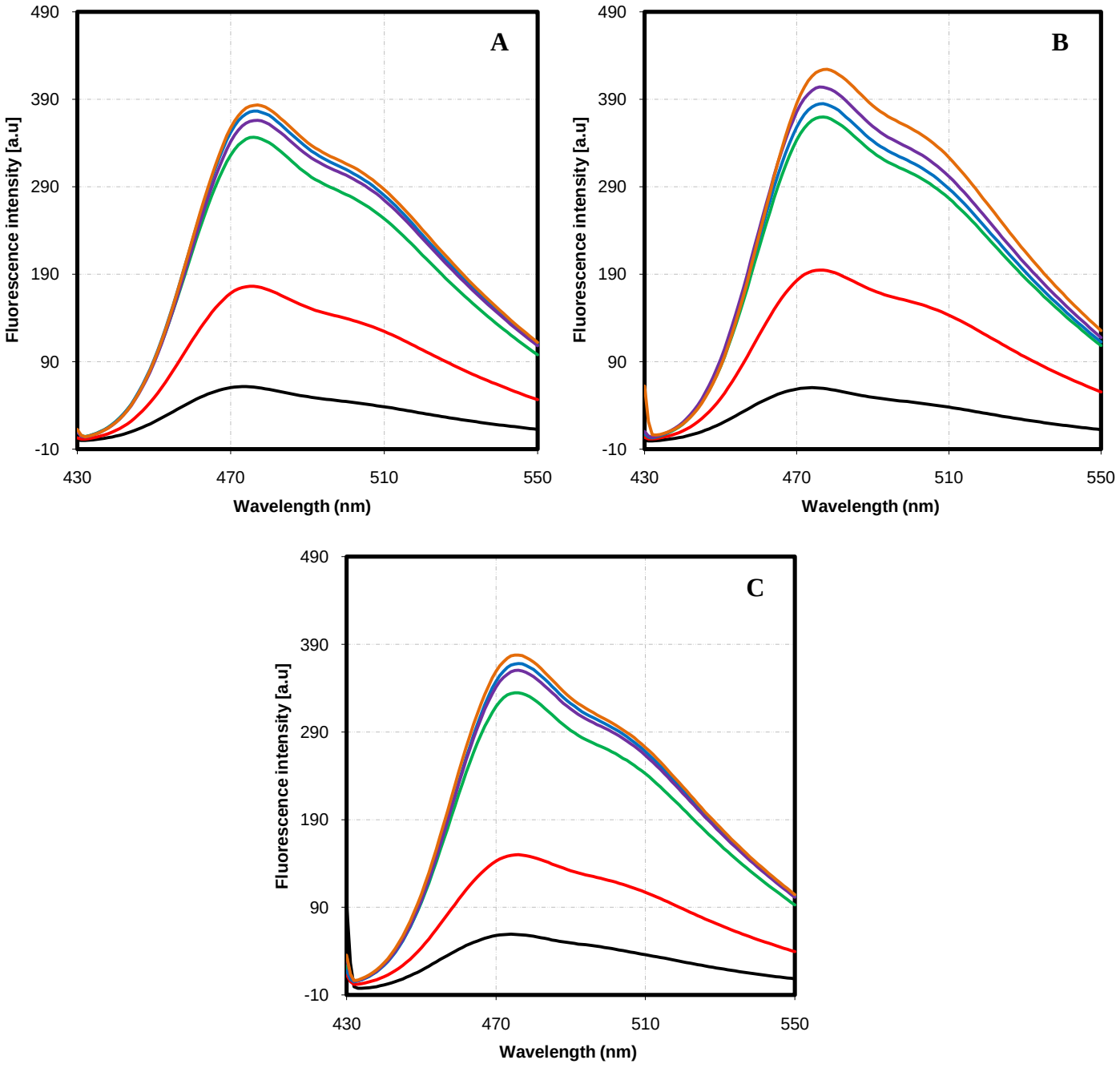
Graphs C.1 – ThT-binding fluorescence intensity spectra obtained for every repetition (A, B and C) of the 10 days-system with 1 mg/ml of β-Lg.



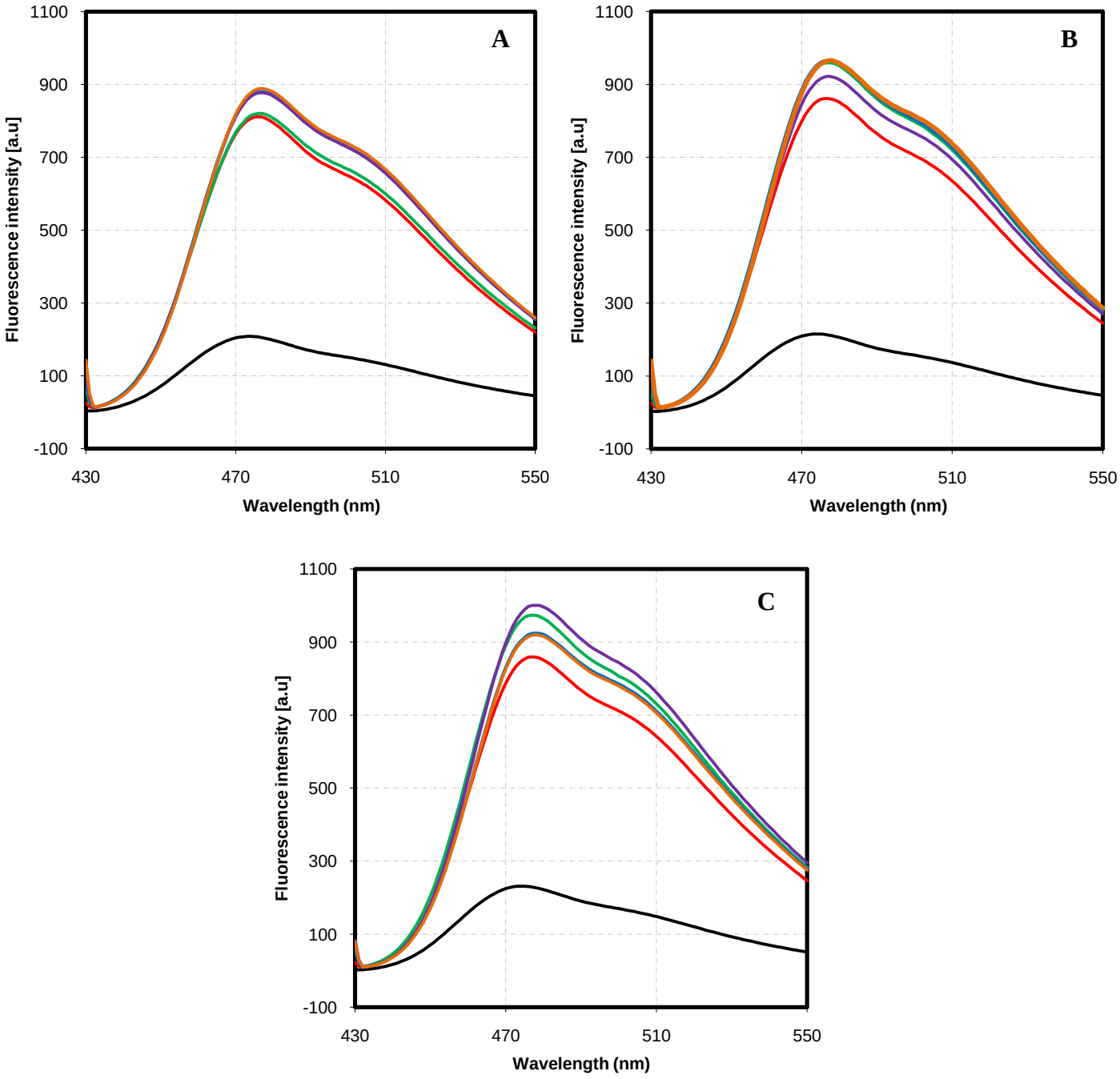
Graphs C.2 – ThT-binding fluorescence intensity spectra obtained for every repetition (A, B and C) of the 10 days-system with 4 mg/ml of β-Lg.



Graphs C.3 – ThT-binging fluorescence intensity spectra obtained for every repetition (A, B and C) of the 10 days-system with 10 mg/ml of β-Lg.

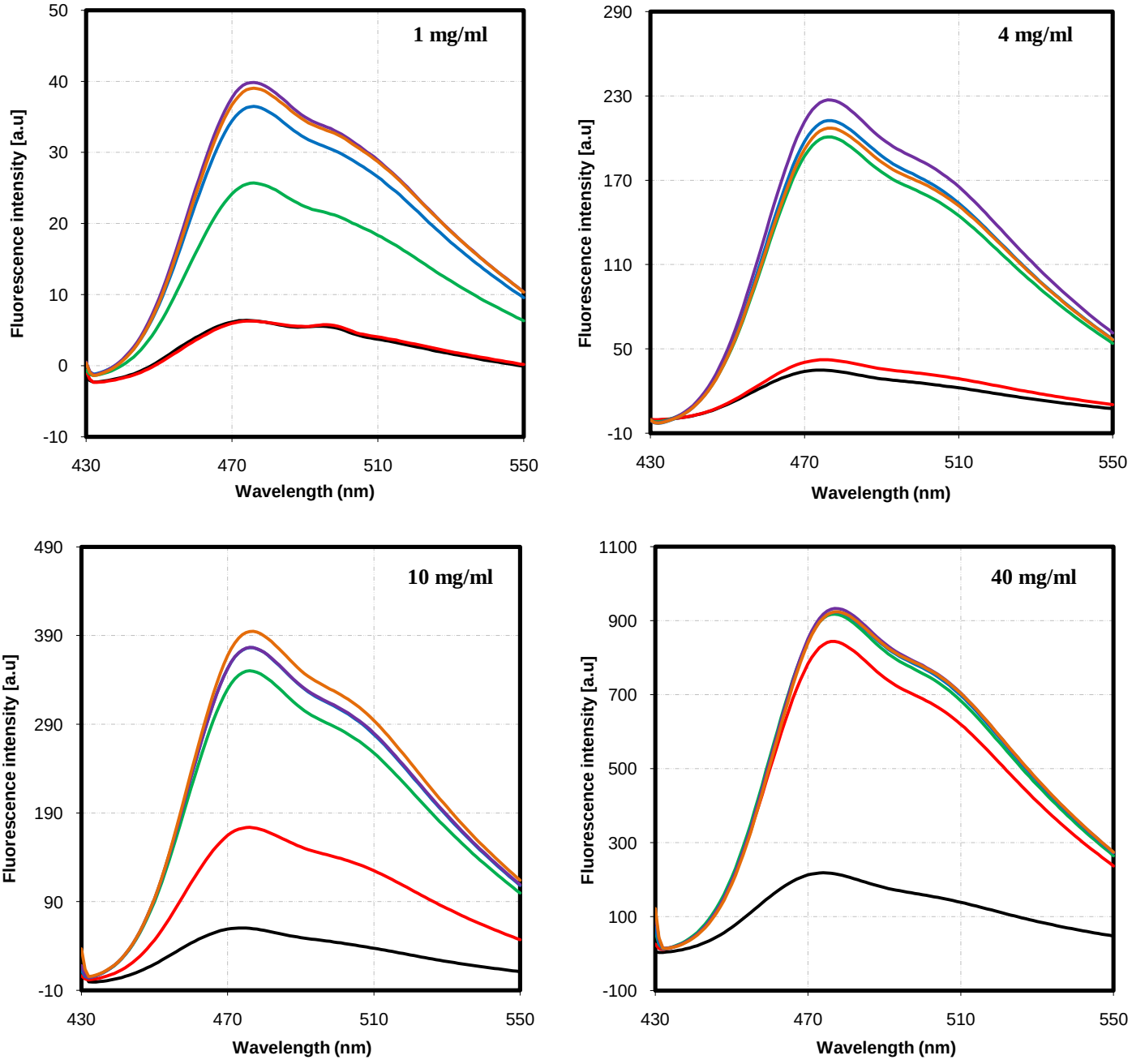


Graphs C.4 – ThT-binding fluorescence intensity spectra obtained for every repetition (A, B and C) of the 10 days-system with 40 mg/ml of β-Lg.



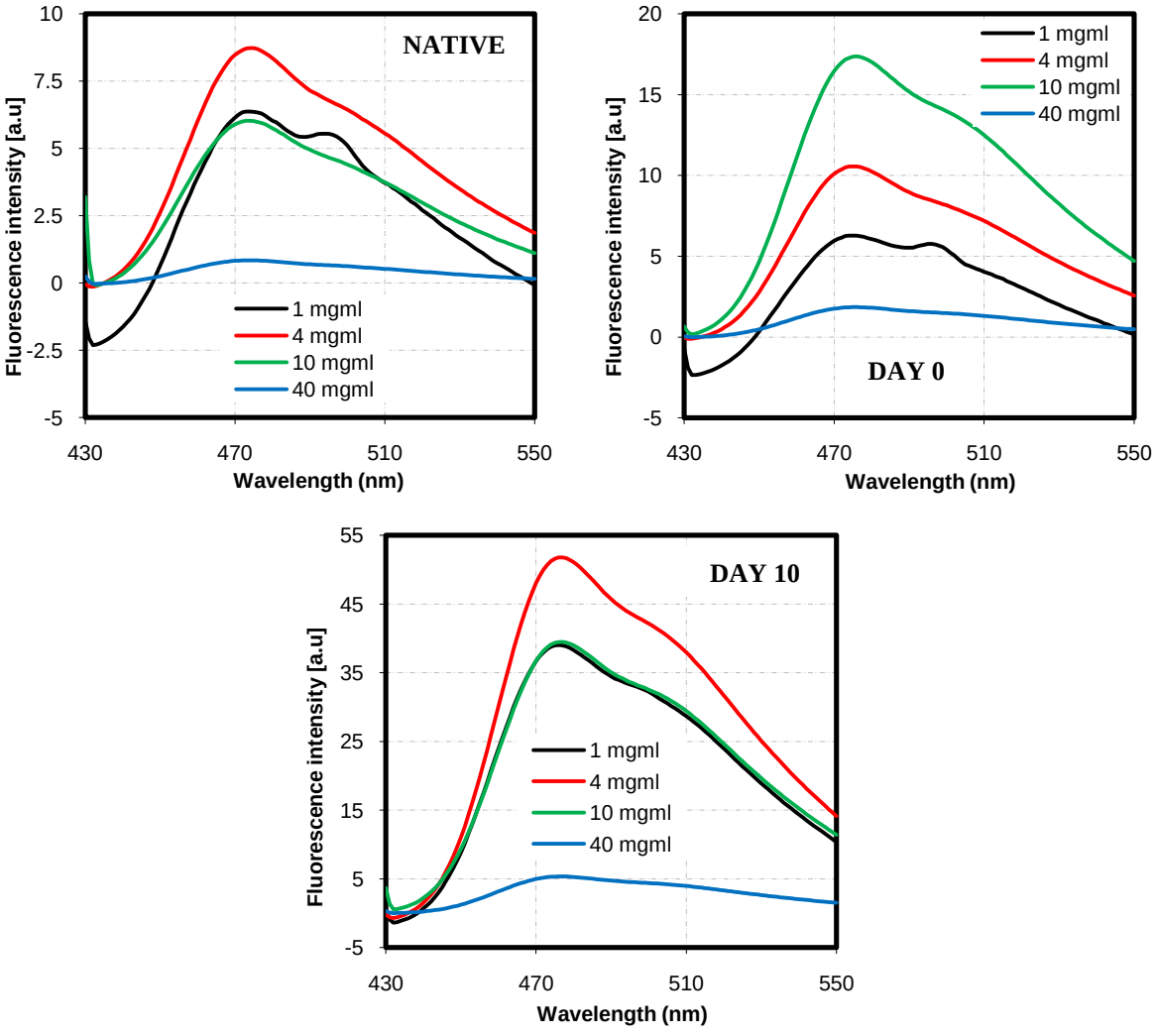
The average ThT-binding fluorescence intensity spectra obtained from the repetitions shown above for each concentration of protein are shown in Graphs C.5.

Graphs C.5 – Average ThT-binding fluorescence intensity spectra obtained as the mean of the fluorescence spectra obtained from the three repetitions done for each protein concentration during 10 days of incubation at 22°C.



Average fluorescence intensity spectra shown in Graphs C.5 were normalised to concentration (dividing by the corresponding protein concentration) for the protein in its native state, after addition of ethanol (Day 0) and after 10 days of incubation at 22ºC. The objective was to analyse in further detail the shape of the spectra to see if the concentration of the protein had more influence in the results than only the different intensity of the fluorescence signal (approximately proportional to concentration). However no further significant information was found as mentioned in section 3.3.1. Normalised fluorescence intensity spectra are shown in Graphs C.6 for each incubation time, together with the native state of the protein.

Graphs C.6 –ThT-binding fluorescence spectra normalised to concentration for the aqueous solution (native protein) and each incubation time (Day 0 and Day 10).



APPENDIX D

Table D.1 collects the individual apparent hydrodynamic radius (R_{hi}) calculated for each concentration of protein (1, 4, 10 and 40 mg/ml) dissolved in distilled water, phosphate buffer 10 mM and phosphate buffer 20 mM for every repetition together with the error of the fit associated with each individual hydrodynamic radius (see APPENDIX B for details in the calculation). Different repetitions are named with an alphabet letter followed by the protein concentration of the system; hence "system A1" indicates repetition "A" of the system with "1 mg/ml of β-Lg."

Table D.1 - Apparent hydrodynamic radii, R_{hi} (nm) for every repetition done for aqueous solutions of β-Lg (protein in its native state) dissolved in distilled water and phosphate buffer with concentrations 10 mM and 20 mM (pH 7) at 22°C.

R_{hi} (nm)	1 mg/ml			4 mg/ml		
Repetition:	**A1**	**B1**	**C1**	**A4**	**B4**	**C4**
Water	$2.1 \pm 2.3 \cdot 10^{-2}$	$2.8 \pm 1.3 \cdot 10^{-2}$	$4.4 \pm 3.2 \cdot 10^{-2}$	$1.9 \pm 3.1 \cdot 10^{-2}$	$2.8 \pm 9.8 \cdot 10^{-2}$	$2.1 \pm 1.4 \cdot 10^{-2}$
Buffer 10 mM	$4.9 \pm 6.4 \cdot 10^{-2}$	$3.9 \pm 4.8 \cdot 10^{-2}$	$3.3 \pm 4.00 \cdot 10^{-2}$	$3.3 \pm 1.1 \cdot 10^{-2}$	$3.0 \pm 7.2 \cdot 10^{-3}$	$3.2 \pm 9.7 \cdot 10^{-3}$
Buffer 20 mM	$19 \pm 1.8 \cdot 10^{-1}$	$15 \pm 4.9 \cdot 10^{-1}$	$12 \pm 1.1 \cdot 10^{-1}$	$3.0 \pm 1.1 \cdot 10^{-2}$	$3.4 \pm 1.2 \cdot 10^{-2}$	$3.2 \pm 9.6 \cdot 10^{-3}$
R_{hi} (nm)	10 mg/ml			40 mg/ml		
Repetition:	**A10**	**B10**	**C10**	**A40**	**B40**	**C40**
Water	$1.9 \pm 5.4 \cdot 10^{-3}$	$2.2 \pm 1.1 \cdot 10^{-2}$	$1.5 \pm 7.6 \cdot 10^{-3}$	$1.8 \pm 5.2 \cdot 10^{-3}$	$1.9 \pm 9.4 \cdot 10^{-3}$	$1.6 \pm 2.5 \cdot 10^{-3}$
Buffer 10 mM	$3.3 \pm 9.7 \cdot 10^{-3}$	$2.9 \pm 9.5 \cdot 10^{-3}$	$3.2 \pm 7.1 \cdot 10^{-3}$	$2.3 \pm 6.3 \cdot 10^{-3}$	$2.5 \pm 4.2 \cdot 10^{-3}$	$2.4 \pm 5.4 \cdot 10^{-3}$
Buffer 20 mM	$3.1 \pm 2.9 \cdot 10^{-3}$	$3.3 \pm 3.9 \cdot 10^{-3}$	$3.5 \pm 5.5 \cdot 10^{-3}$	$2.8 \pm 1.2 \cdot 10^{-2}$	$3.2 \pm 2.3 \cdot 10^{-3}$	$3.0 \pm 3.1 \cdot 10^{-3}$

Normalised autocorrelation functions for samples of the protein in its native state consisting in 1 mg/ml of β-Lg dissolved in distilled water, phosphate buffer 10 mM and phosphate buffer 20 mM at pH 7 and 22°C are shown in Graphs D.1 to D.3. Firstly, the normalised autocorrelation functions for each repetition are plotted with the corresponding theoretical autocorrelation function calculated from the hydrodynamic radius (R_{hi}) obtained from the fitting of the raw data for that repetition. This shows the goodness of the fit. Lastly, the autocorrelation functions for each repetition are plotted together with the theoretical autocorrelation function obtained from the mean apparent hydrodynamic radius (R_h) calculated from different repetitions and shown in Table 3.3. The theoretical autocorrelation function (thick black line) was calculated as follows:

R_{hi} (from repetition) or R_h (mean of the three repetitions)

$$\downarrow$$

$$D = \frac{k_B \cdot T}{6 \cdot \pi \cdot \mu \cdot R_h}$$

$$\downarrow$$

$$\Gamma = D \cdot q^2$$

$$\downarrow$$

$$g_1(\tau) = \exp(-\Gamma \cdot \tau)$$

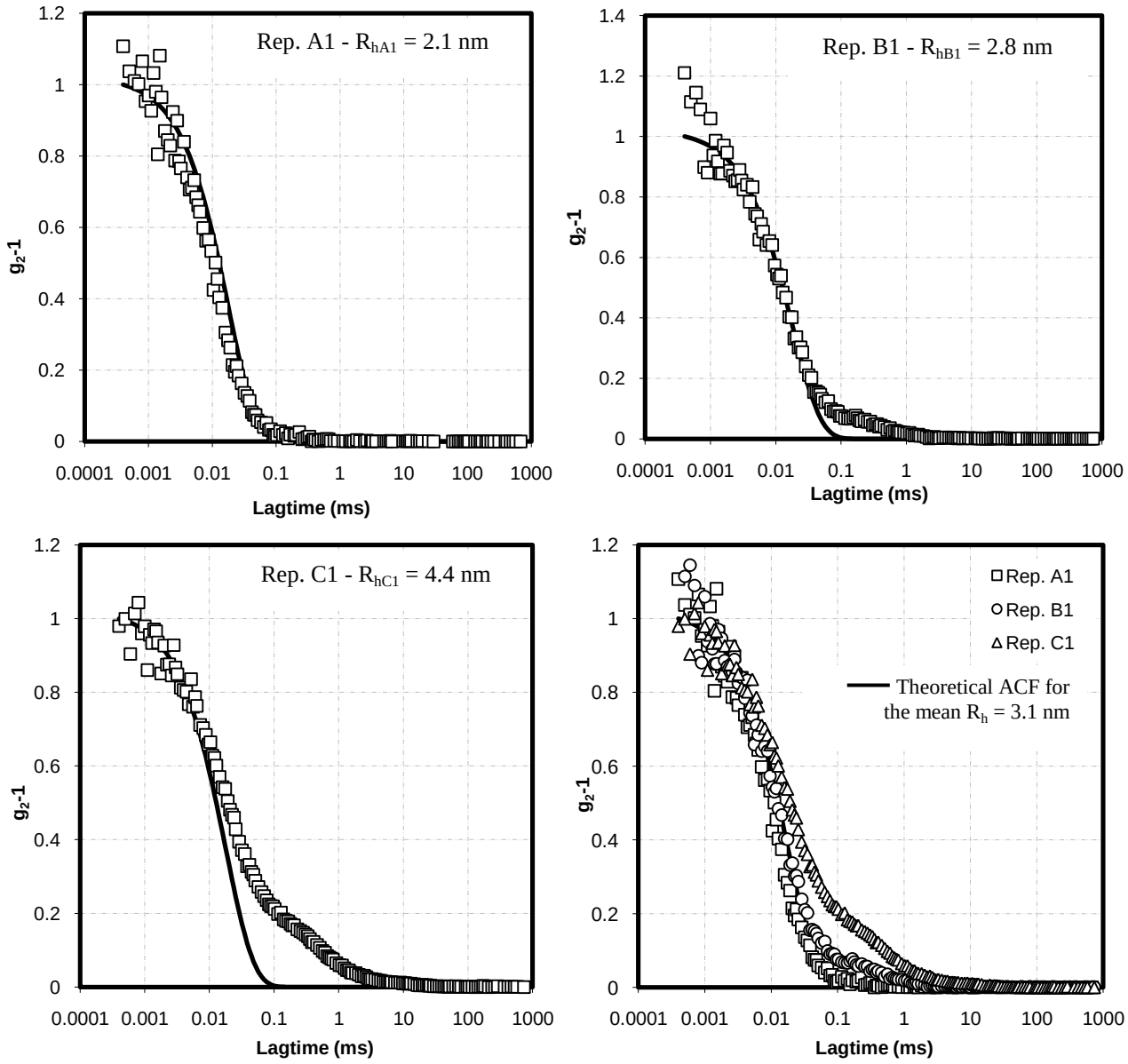
$$\downarrow$$

$$g_2(\tau) - 1 = (g_1(\tau))^2$$

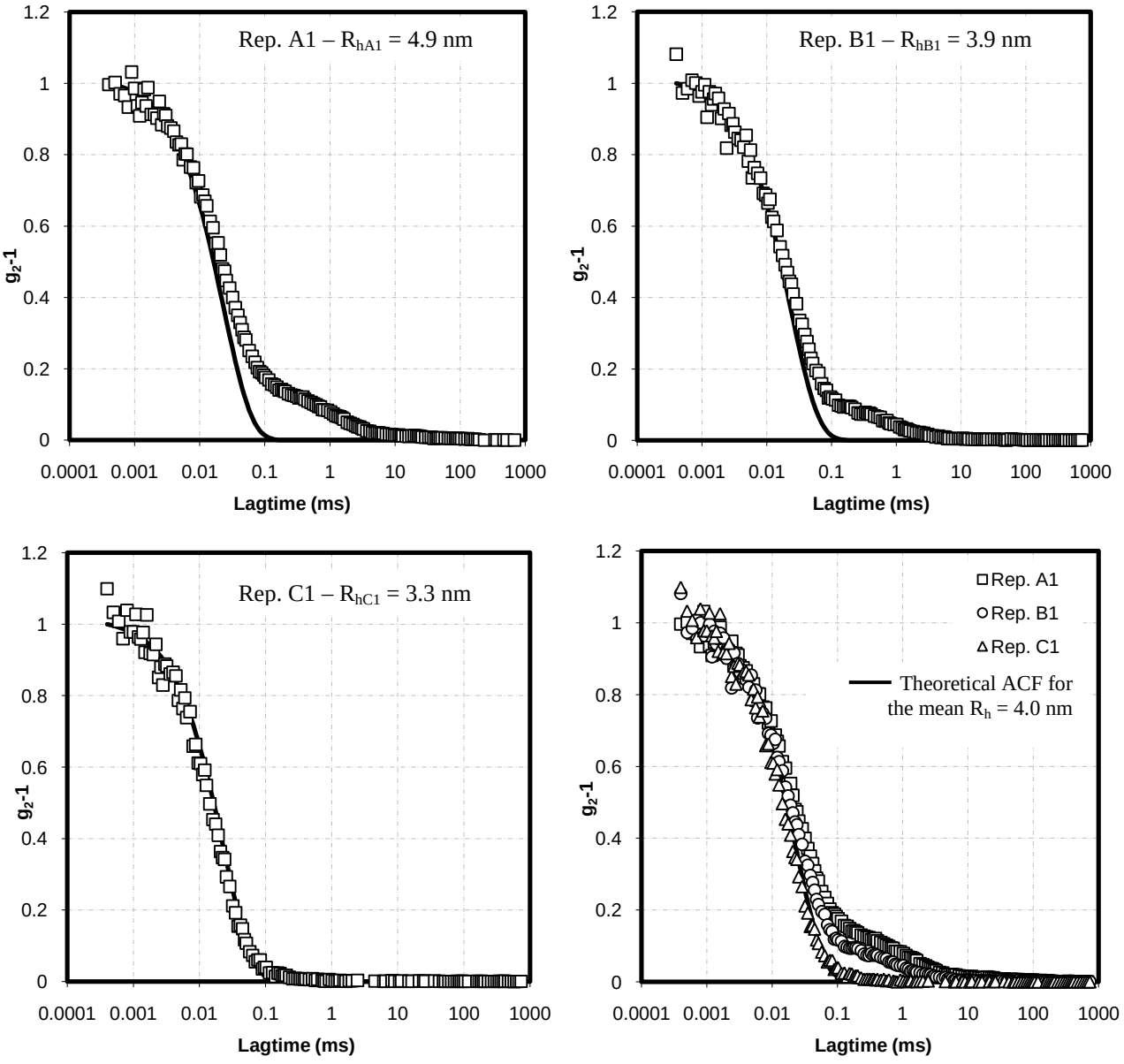
Graphs D.1 to D.3 represent the goodness of the fit of the autocorrelation functions obtained for the lowest concentration of protein in its native state in different solvents. It can be seen that for 1 mg/ml of protein the autocorrelation functions are

less defined and the first decay is not as smooth as it would be required for an accurate fit.

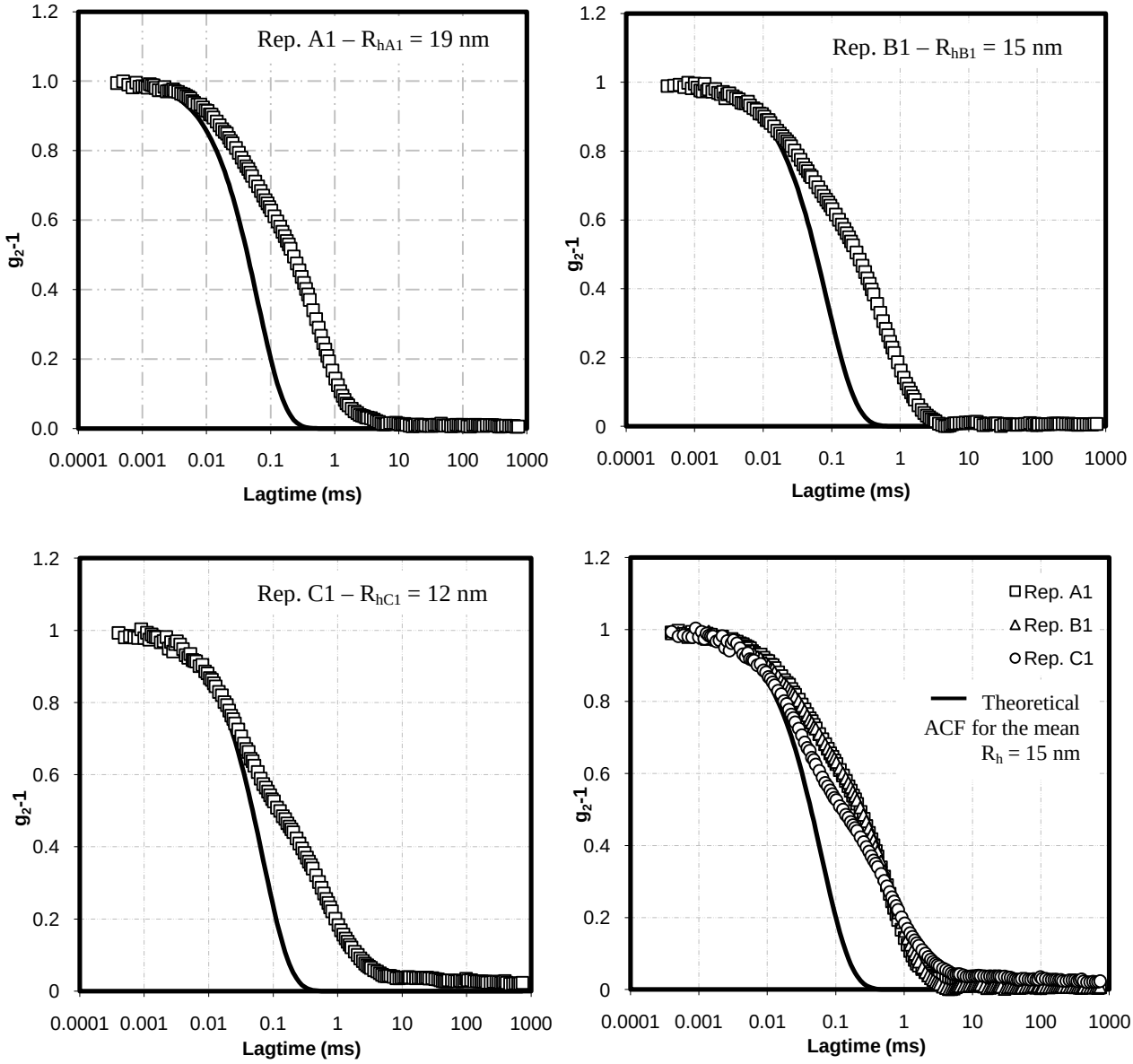
Graphs D.1 – Normalised autocorrelation functions from β-Lg dissolved in distilled water to a concentration of 1 mg/ml.



Graphs D.2 – Normalised autocorrelation functions from β-Lg dissolved in phosphate buffer 10 mM to a concentration of 1 mg/ml.



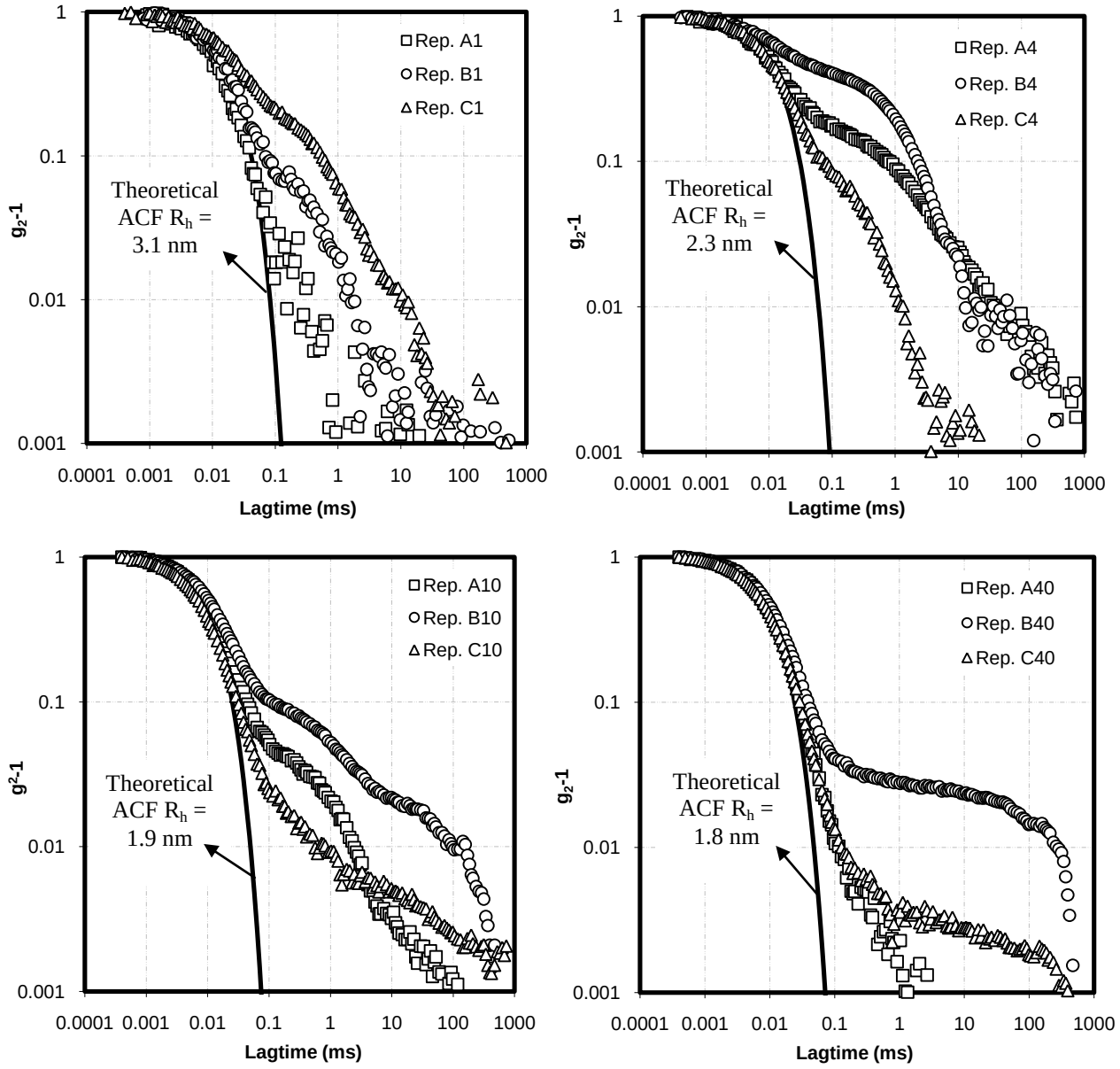
Graphs D.3 – Normalised autocorrelation functions from β-Lg dissolved in phosphate buffer 20 mM to a concentration of 1 mg/ml.



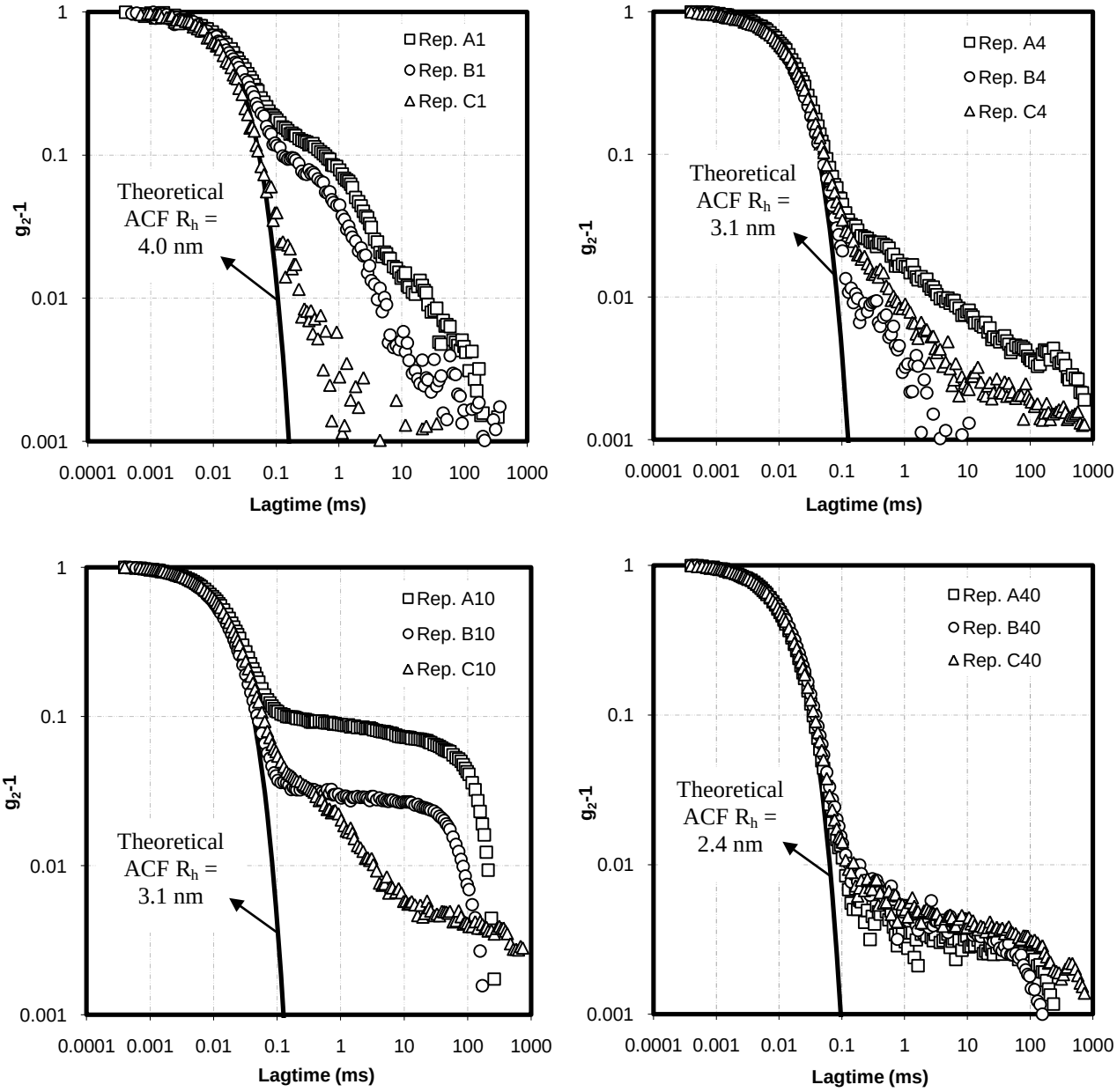
Next graphs (Graphs D.4, D.5 and D.6) represent the tails observed in the autocorrelation functions obtained at long lagtimes for the aqueous solution of β-Lg. The normalised autocorrelation functions obtained for each repetition is shown together with the theoretical autocorrelation function calculated from the mean apparent hydrodynamic radius (R_h) collected in Table 3.3 (and shown again here) for every concentration of protein. Each repetition is labelled with the alphabet letter A to C follow by the protein concentration. Hence, experiment "A1" represents repetition "A" of the system with "1 mg/ml of β-Lg." Logarithmic scale in the vertical axis is required to allow representation of the autocorrelation functions for long lagtimes.

The tails observed for long lagtimes diverging from the theoretical autocorrelation function indicate a rather polydisperse nature of the aqueous solutions of β-Lg analysed by DLS. It is worth mentioning that these tails are observed for each concentration of the native protein and for every solvent varying in the lagtime at which they are more pronounced. As mentioned in section 3.1, the meaning of these tails is not fully understood yet in the scientific community. Therefore, the analysis of the last part of the autocorrelation function can only be made from a qualitative point of view and only up to the point to confirm that the systems under study present a polydisperse character whose level varies depending on protein concentration and solvent.

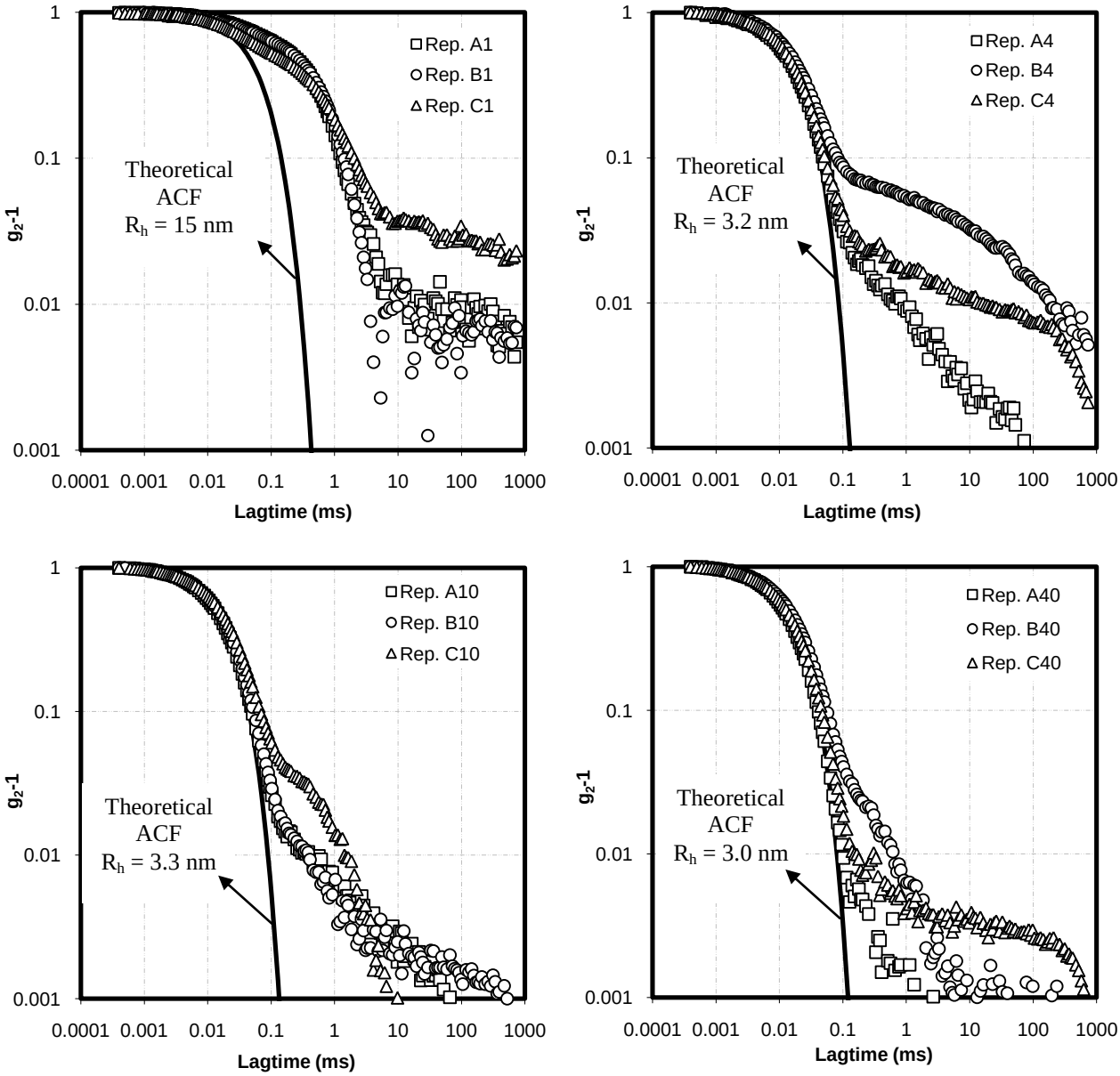
Graphs D.4 – Autocorrelation functions tails for samples of β-Lg dissolved in distilled water.



Graphs D.5 – Autocorrelation functions tails for samples β-Lg dissolved in 10 mM phosphate buffer.

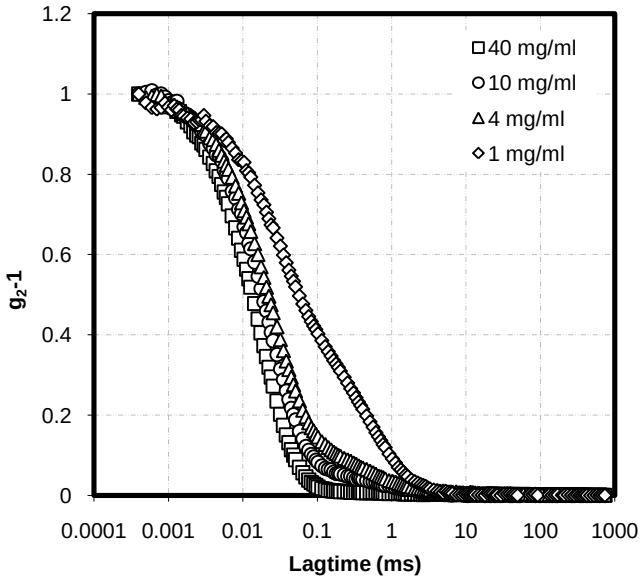


Graphs D.6 – Autocorrelation functions tails for samples β-Lg dissolved in 20 mM phosphate buffer.



Graphs D.7 and D.8 shows the normalized autocorrelation functions obtained in the dilution experiment mentioned in section 3.1. Graph D.7 shows the autocorrelation functions where the first decay is clearer while graph D.8 highlights the last tails observed. Table D.2 shows the hydrodynamic radius obtained from the fitting of these autocorrelation functions.

Graph D.7 – Normalised autocorrelation functions highlighting the first decay for all concentration of protein obtained in the dilution experiment (section 3.1).



Graph D.8 - Normalised autocorrelation functions highlighting the tails at long lagtimes, for all concentrations of protein obtained in the dilution experiment (section 3.1).

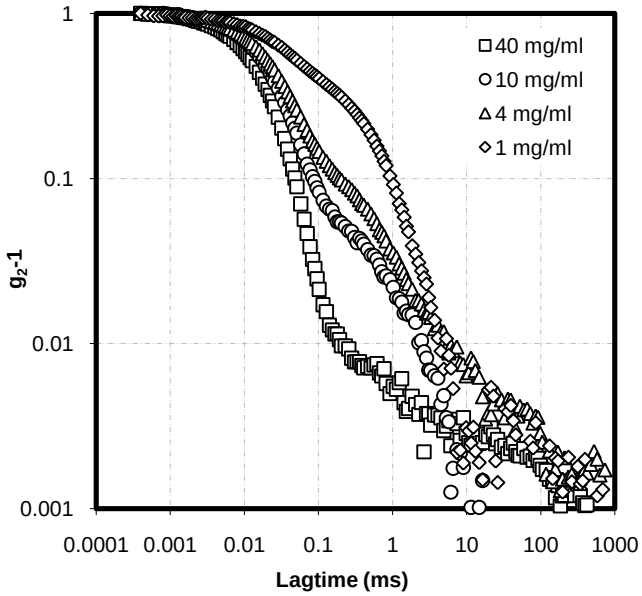


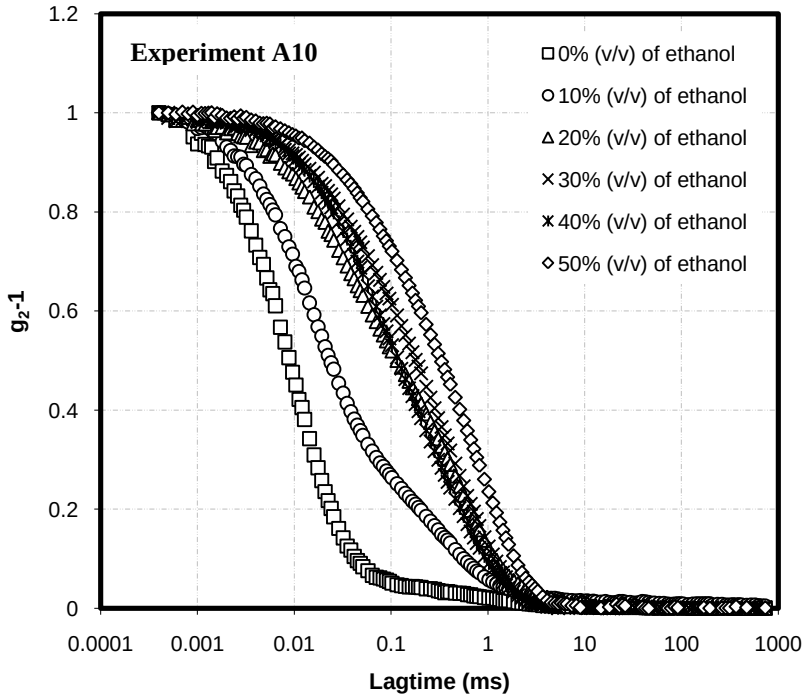
Table D.2 - Apparent hydrodynamic radii, R_{hi} (nm) obtained from a sample of 40 mg/ml of β-Lg diluted subsequently to 10 mg/ml, 4 mg/ml and 1 mg/ml keeping a buffer concentration of 20 mM (pH 7) at 22°C.

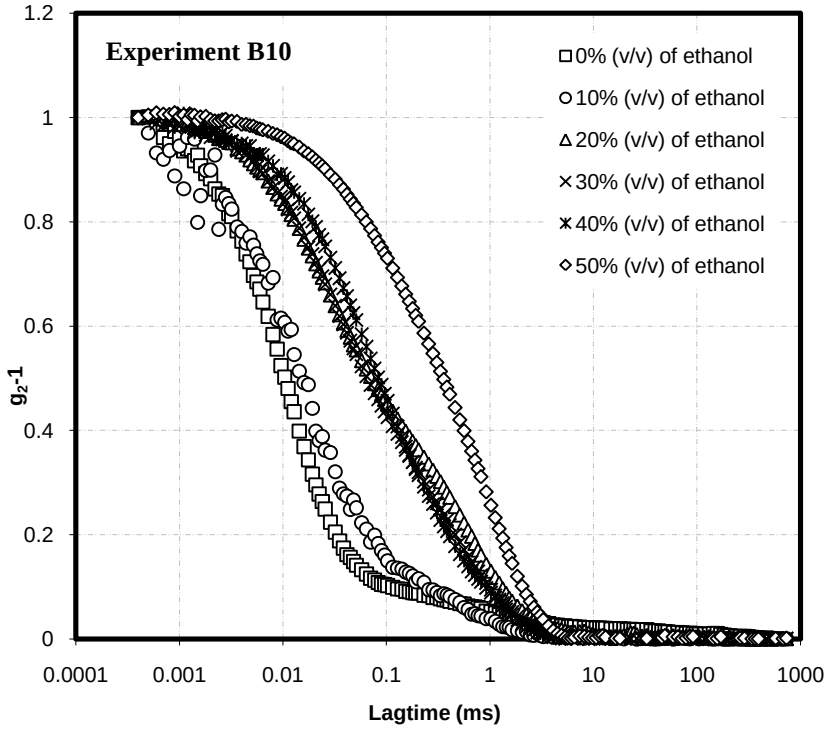
Concentration of protein (mg/ml)	**$R_{hi} \pm \sigma_{hi}$ (nm)**
40	$3.1 \pm 2.9 \cdot 10^{-3}$
10	$4.1 \pm 1.2 \cdot 10^{-2}$
4	$4.8 \pm 1.2 \cdot 10^{-2}$
1	$9.7 \pm 9.0 \cdot 10^{-2}$

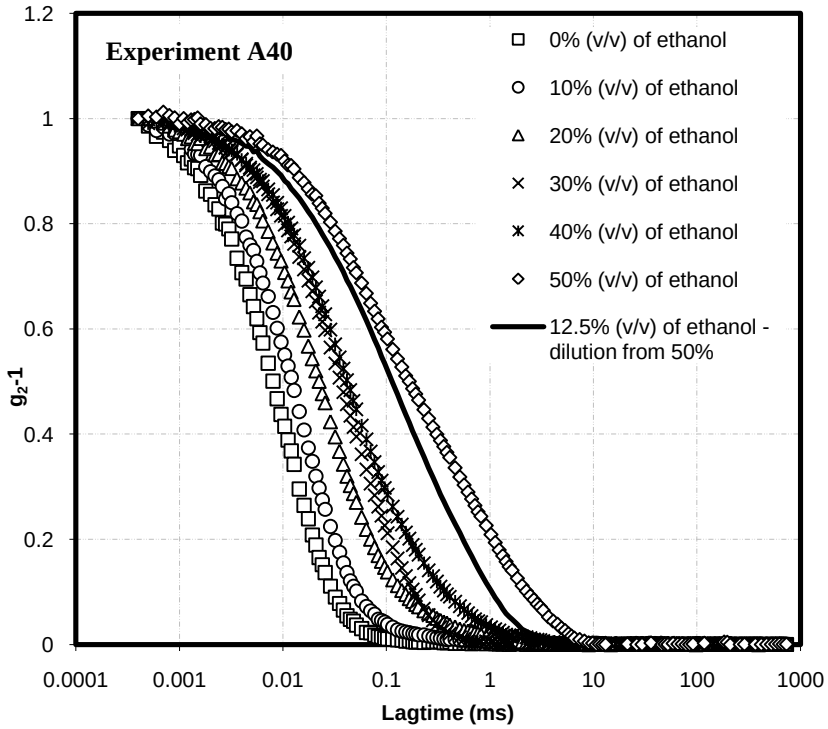
APPENDIX E

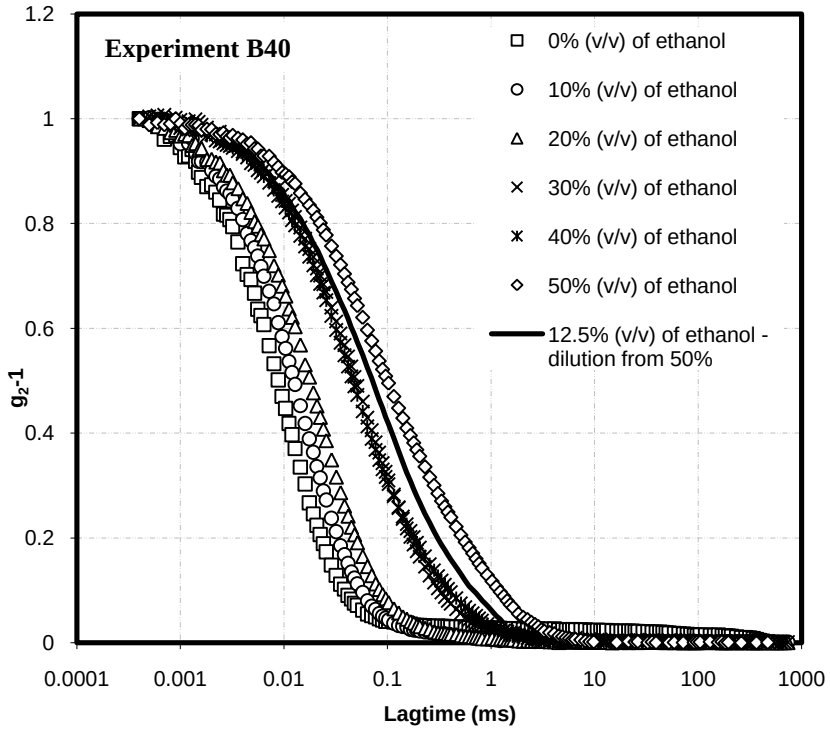
Normalised autocorrelation functions obtained for samples with different percentages of ethanol and protein concentration 10 mg/ml and 40 mg/ml in both distilled water and phosphate buffer 10 mM are presented in Graphs E.1 and E.2. In this case, only two repetitions were done for every concentration of ethanol since the objective of the experiment was rather qualitative. The individual hydrodynamic radius (R_{hi}) and the average values (R_h) presented in Figure 3.8 are shown in Table E.1 and Table E.2 for 10 and 40 mg/ml of protein respectively. Each experiment was labelled with the alphabet letters A and B followed by the protein concentration. Hence, experiment "A10" indicates "repetition A" of a system with "10 mg/ml of β-Lg."

Graphs E.1 – Normalised autocorrelation functions for samples of β-Lg with different percentages of ethanol (v/v) in distilled water.

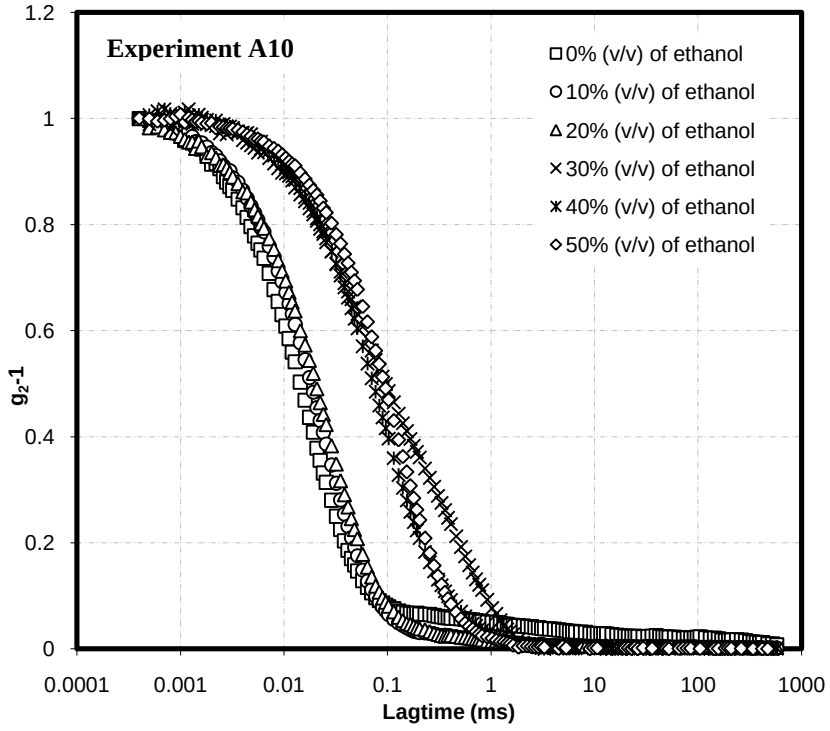


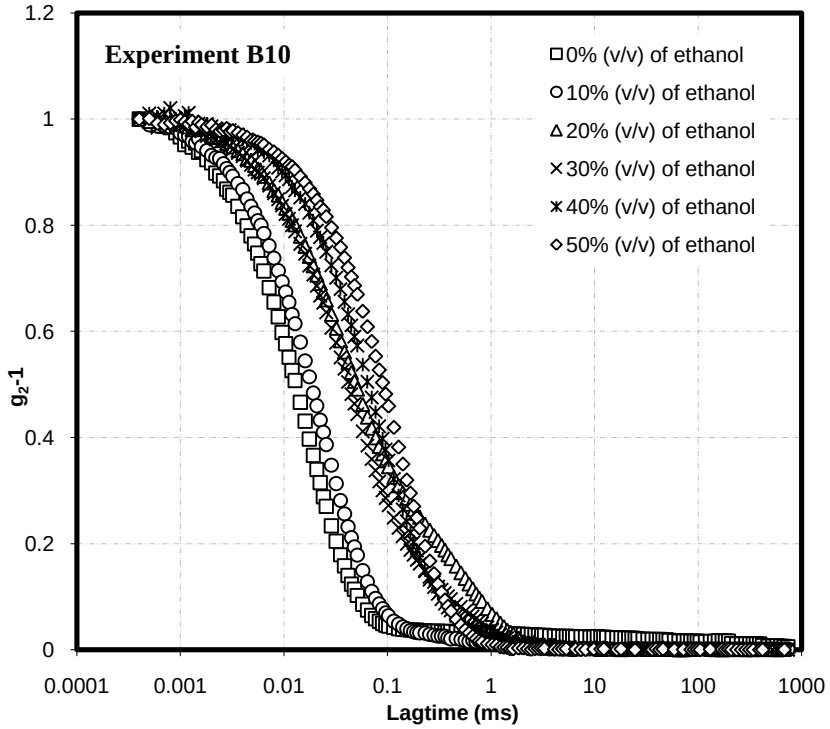

Experiment B10
0% (v/v) of ethanol
10% (v/v) of ethanol
20% (v/v) of ethanol
30% (v/v) of ethanol
40% (v/v) of ethanol
50% (v/v) of ethanol
g2-1
Lagtime (ms)
1.2
1
0.8
0.6
0.4
0.2
0
0.0001
0.001
0.01
0.1
1
10
100
1000

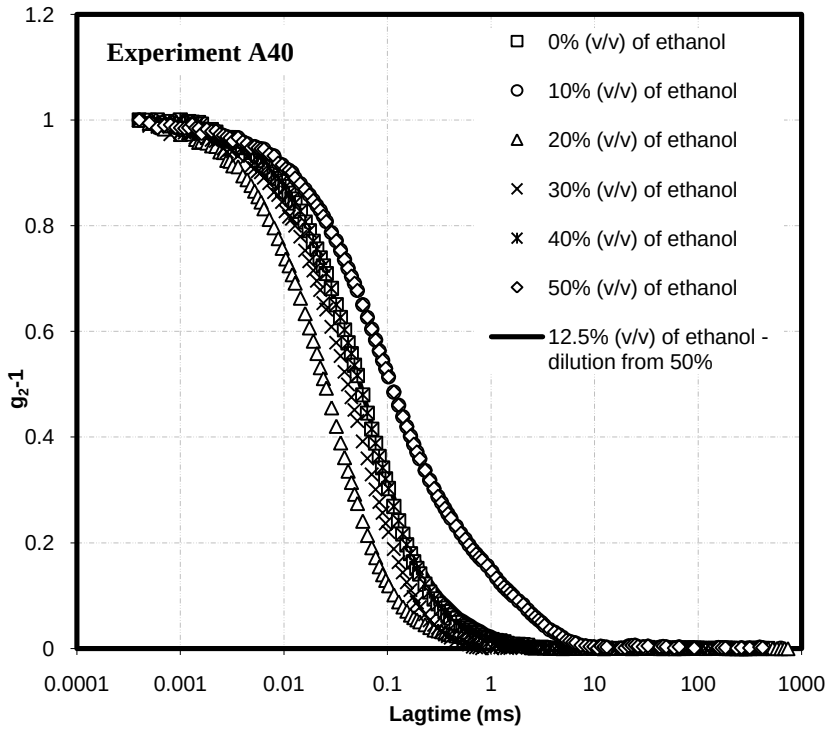

Experiment A40
0% (v/v) of ethanol
10% (v/v) of ethanol
20% (v/v) of ethanol
30% (v/v) of ethanol
40% (v/v) of ethanol
50% (v/v) of ethanol
12.5% (v/v) of ethanol - dilution from 50%
g2-1
Lagtime (ms)
1.2
1
0.8
0.6
0.4
0.2
0
0.0001
0.001
0.01
0.1
1
10
100
1000



Graphs E.2 – Normalised autocorrelation functions for samples of β-Lg with different percentages of ethanol (v/v) in 10 mM phosphate buffer.




Experiment B10
0% (v/v) of ethanol
10% (v/v) of ethanol
20% (v/v) of ethanol
30% (v/v) of ethanol
40% (v/v) of ethanol
50% (v/v) of ethanol
g2-1
Lagtime (ms)
1.2
1
0.8
0.6
0.4
0.2
0
0.0001
0.001
0.01
0.1
1
10
100
1000


Experiment A40
0% (v/v) of ethanol
10% (v/v) of ethanol
20% (v/v) of ethanol
30% (v/v) of ethanol
40% (v/v) of ethanol
50% (v/v) of ethanol
12.5% (v/v) of ethanol - dilution from 50%
g2-1
Lagtime (ms)
1.2
1
0.8
0.6
0.4
0.2
0
0.0001
0.001
0.01
0.1
1
10
100
1000

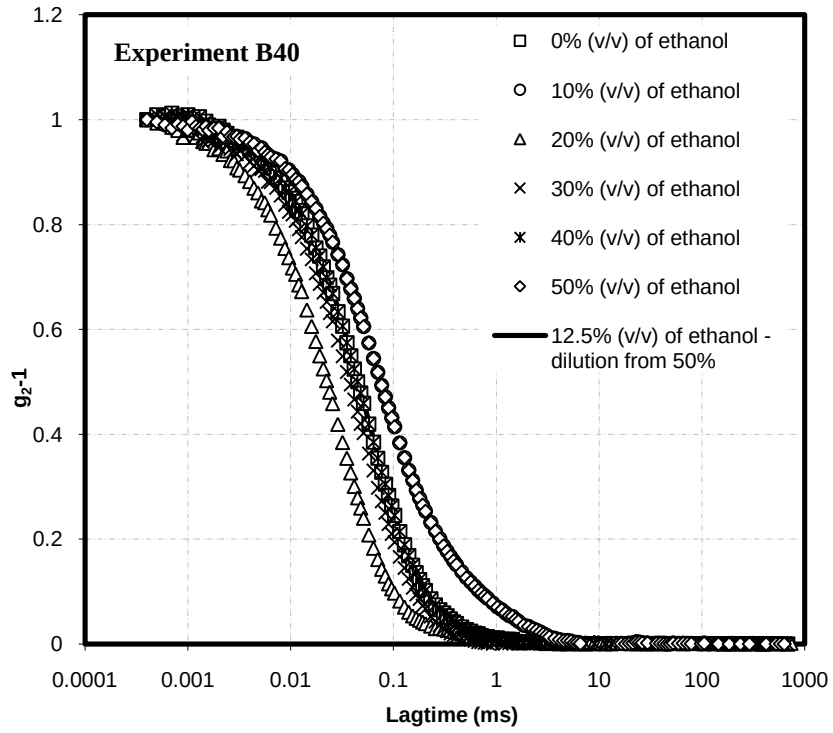


Table E.1 – Apparent hydrodynamic radius R_{hi} (nm) calculated from autocorrelation functions in Graphs E.1 and E.2 for every repetition of the system containing 10 mg/ml of β-Lg. Mean hydrodynamic radius R_h (nm) also shown.

R_{hi} (nm)	**10 mg/ml in water**		**10 mg/ml in 10 mM buffer**		**Mean R_h for 10 mg/ml (presented in Figure 3.8)**	
Repetition:	**A10**	**B10**	**A10**	**B10**	**In water**	**In 10mM buffer**
0 % (v/v) of ethanol	$1.9 \pm 5.4 \cdot 10^{-3}$	$2.2 \pm 1.1 \cdot 10^{-2}$	$3.2 \pm 4.1 \cdot 10^{-3}$	$3.0 \pm 3.7 \cdot 10^{-3}$	2.0	3.1
10 % (v/v) of ethanol	$2.8 \pm 2.4 \cdot 10^{-2}$	$2.3 \pm 1.2 \cdot 10^{-1}$	$3.1 \pm 3.0 \cdot 10^{-3}$	$3.1 \pm 5.7 \cdot 10^{-3}$	2.6	3.1
20 % (v/v) of ethanol	$6.3 \pm 6.6 \cdot 10^{-2}$	$4.3 \pm 2.0 \cdot 10^{-2}$	$2.6 \pm 3.2 \cdot 10^{-3}$	$4.7 \pm 1.7 \cdot 10^{-2}$	5.3	3.7
30 % (v/v) ff ethanol	$8.3 \pm 8.2 \cdot 10^{-2}$	$3.6 \pm 2.8 \cdot 10^{-2}$	$6.0 \pm 5.1 \cdot 10^{-2}$	$3.7 \pm 3.7 \cdot 10^{-2}$	5.9	4.8
40 % (v/v) of ethanol	$6.4 \pm 4.6 \cdot 10^{-2}$	$5.4 \pm 3.0 \cdot 10^{-2}$	$5.9 \pm 2.8 \cdot 10^{-2}$	$5.3 \pm 3.6 \cdot 10^{-2}$	5.9	5.6
50 % (v/v) of ethanol	$12 \pm 1.1 \cdot 10^{-1}$	$13 \pm 8.1 \cdot 10^{-2}$	$7.9 \pm 1.2 \cdot 10^{-2}$	$7.9 \pm 1.0 \cdot 10^{-2}$	13	7.9

Table E.2 – Apparent hydrodynamic radius R_{hi} (nm) calculated from autocorrelation functions in Graphs E.1 and E.2 for every repetition of the system containing 40 mg/ml of β-Lg. Mean hydrodynamic radius R_h (nm) also shown.

R_{hi} (nm)	**40 mg/ml in water**		**40 mg/ml in 10 mM buffer**		**Mean R_h for 40 mg/ml (presented in Figure 3.8)**	
Repetition:	**A10**	**B10**	**A10**	**B10**	**In water**	**In 10mM buffer**
0 % (v/v) of ethanol	$1.8 \pm 6.3 \cdot 10^{-3}$	$1.9 \pm 9.4 \cdot 10^{-3}$	$2.6 \pm 4.1 \cdot 10^{-3}$	$2.6 \pm 3.0 \cdot 10^{-2}$	1.9	2.6
10 % (v/v) of ethanol	$2.0 \pm 8.7 \cdot 10^{-3}$	$2.1 \pm 7.6 \cdot 10^{-3}$	$2.8 \pm 4.4 \cdot 10^{-3}$	$2.3 \pm 2.9 \cdot 10^{-2}$	2.0	2.6
20 % (v/v) of ethanol	$2.6 \pm 1.1 \cdot 10^{-2}$	$2.1 \pm 1.3 \cdot 10^{-2}$	$3.1 \pm 8.1 \cdot 10^{-3}$	$2.6 \pm 3.9 \cdot 10^{-2}$	2.3	2.9
30 % (v/v) ff ethanol	$3.0 \pm 1.9 \cdot 10^{-2}$	$4.0 \pm 1.7 \cdot 10^{-2}$	$4.2 \pm 2.4 \cdot 10^{-2}$	$3.7 \pm 2.6 \cdot 10^{-2}$	3.5	3.9
40 % (v/v) of ethanol	$3.1 \pm 8.3 \cdot 10^{-3}$	$3.4 \pm 1.1 \cdot 10^{-2}$	$4.6 \pm 2.5 \cdot 10^{-2}$	$3.9 \pm 2.7 \cdot 10^{-2}$	3.2	4.3
50 % (v/v) of ethanol	$6.9 \pm 7.2 \cdot 10^{-2}$	$5.4 \pm 4.8 \cdot 10^{-2}$	$7.2 \pm 3.5 \cdot 10^{-2}$	$5.9 \pm 2.3 \cdot 10^{-2}$	6.2	6.6
12.5 % (v/v) of ethanol – diluted from 50% (v/v) of ethanol	$8.2 \pm 9.1 \cdot 10^{-2}$	$6.4 \pm 5.7 \cdot 10^{-2}$	$8.3 \pm 2.6 \cdot 10^{-2}$	$7.0 \pm 3.9 \cdot 10^{-2}$	7.3	7.6

APPENDIX F

ThT-binding fluorescence intensities for different concentrations of β-Lg monitored over incubation time at 22ºC to calculate the kinetics of β-Lg structure evolution, corresponding with results shown in Figure 3.11 are presented here. The original representation in linear scale of these data (ThT-binding fluorescence intensity vs. incubation time) is also shown in this appendix. In Figure 3.11 this graph was represented in logarithmic scale in both axes for clarity.

Tables F.1 to F.5 show the ThT-binding fluorescence intensity at every time of incubation for each concentration of protein, including the fluorescence intensity observed when the native sample was measured. Time 0 corresponds with the first measurement immediately after adding 50% (v/v) of ethanol to the protein solution. For each measurement, ThT and protein sample were mix in the fluorescence quartz cell as it was detailed in section 2.2.2. Incubation time is given in hours (hr) while the ThT fluorescence intensity has arbitrary units [a.u].

Graph F.1 collects the data en Tables F.1 to F.5 showing the structure evolution kinetics for each concentration from the moment the organic solvent is added to the sample. In this case, linear scales are used to represent the original data. Graph F.2 collect the same data as Graph F.1 but including fluorescence intensities for the aqueous solution of the protein in 10 mM buffer. The different behaviour between the system with 4 mg/ml of protein concentration and higher concentrated systems (10 and 40 mg/ml) explained in section 3.1 can be shown in this graph.

Table F.1 – ThT-binding fluorescence intensity at 477 nm (emission wavelength) (λ_{exc} = 425 nm) for different incubation times at 22°C for the system of **0.5 mg/ml** of β-Lg in the presence of 50% of ethanol (v/v) and 10 mM buffer (pH 7).

PROTEIN CONCENTRATION: **0.5 mg/ml**		
Measure's name	***Incubation time (hr)***	***Fluorescence Intensity [a.u]***
Native state of β-Lg	-	6.05
Time 0 – after adding ethanol	0	5.89
Time 1	0.27	5.64
Time 2	0.95	6.26
Time 3	1.95	5.99
Time 4	2.95	6.07
Time 5	3.95	7.13
Time 6	4.95	6.73
Time 7	8.00	7.93
Time 8	9.00	8.42
Time 9	10.0	8.08
Time 10	11.0	8.17
Time 11 – Day 1	21.2	10.7
Time 12 – Day 3	70.5	15.9
Time 13 – Day 7	165	26.4
Time 14 – Day 10	238	28.7
Time 15 – Day 15	357	24.5
Time 16 – Day 21	504	26.9
Time 17 – Day 27	646	28.9

Table F.2 – ThT-binding fluorescence intensity at 477 nm (emission wavelength) (λ_{exc} = 425 nm) at every incubation time at 22°C for the system of **1 mg/ml** of β-Lg in the presence of 50% of ethanol (v/v) and 10 mM buffer (pH 7).

PROTEIN CONCENTRATION: **1 mg/ml**		
Measure's name	***Incubation time (hr)***	***Fluorescence Intensity [a.u]***
Native state of β-Lg	-	6.30
Time 0 – after adding ethanol	0	6.74
Time 1	0.25	6.30
Time 2	0.50	5.81
Time 3	0.75	6.52
Time 4	1.00	6.88
Time 5	1.25	7.12
Time 6	1.50	6.69
Time 7	1.75	6.55
Time 8	2.00	7.39
Time 9	2.25	8.42
Time 10	2.50	7.72
Time 11	2.75	9.21
Time 12	3.00	8.82
Time 13	4.00	10.6
Time 14	5.00	11.2
Time 15	6.00	12.8
Time 16	7.00	14.8
Time 17	8.00	14.6
Time 18	9.00	16.1
Time 19 – Day 1	25.0	27.8
Time 20 – Day 3	68.7	36.5
Time 21 – Day 7	165	38.8
Time 22 – Day 10	237	40.0

Table F.3 – ThT-binding fluorescence intensity at 477 nm (emission wavelength) (λ_{exc} = 425 nm) at every incubation time at 22°C for the system of **4 mg/ml** of β-Lg in the presence of 50% of ethanol (v/v) and 10 mM buffer (pH 7).

PROTEIN CONCENTRATION: **4 mg/ml**		
Measure's name	***Incubation time (hr)***	***Fluorescence Intensity [a.u]***
Native state of β-Lg	-	34.5
Time 0 – after adding ethanol	0	24.1
Time 1	0.25	51.1
Time 2	0.50	67.2
Time 3	0.75	76.5
Time 4	1.00	80.8
Time 5	1.25	105.8
Time 6	1.50	105.9
Time 7	1.75	113.4
Time 8	2.00	110.3
Time 9	2.25	125.6
Time 10	2.50	121.4
Time 11	2.75	104.7
Time 12	3.00	113.1
Time 13	4.00	130.3
Time 14	5.00	151.5
Time 15	6.00	135.2
Time 16	7.00	104.7
Time 17	8.00	157.6
Time 18 – Day 1	24.1	130.3
Time 19 – Day 3	72.6	212.2
Time 20 – Day 7	169	214.4
Time 21 – Day 10	240	200.8

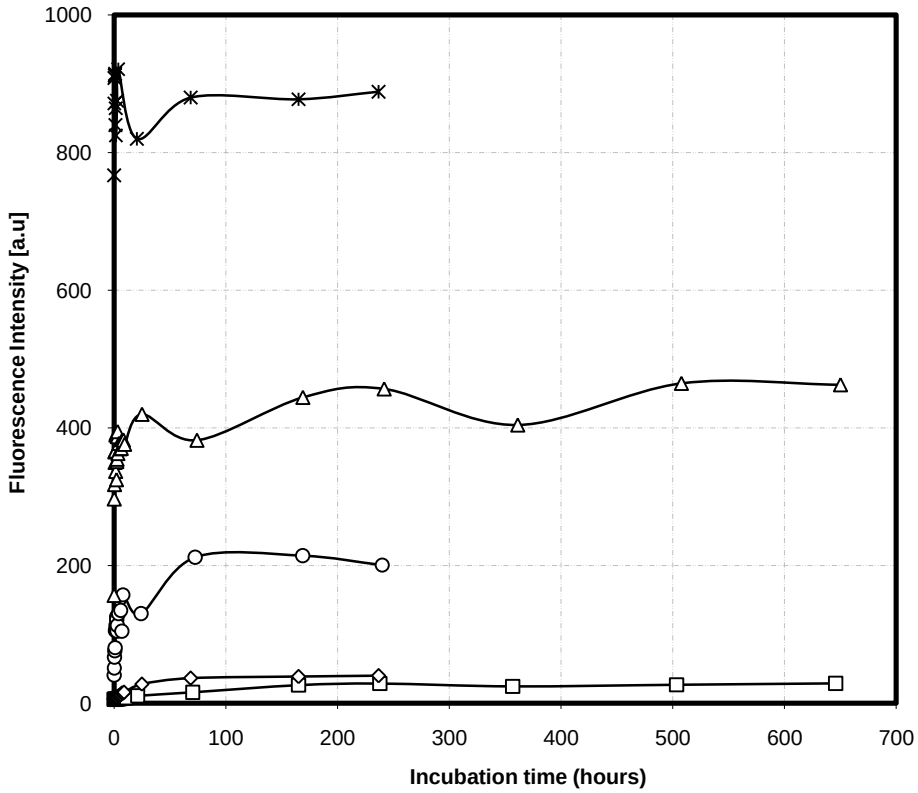
Table F.4 – ThT-binding fluorescence intensity at 477 nm (emission wavelength) (λ_{exc} = 425 nm) at every incubation time at 22°C for the system of **10 mg/ml** of β-Lg in the presence of 50% of ethanol (v/v) and 10 mM buffer (pH 7).

PROTEIN CONCENTRATION: **10 mg/ml**		
Measure's name	***Incubation time (hr)***	***Fluorescence Intensity [a.u]***
Native β-Lg	-	59.5
Time 0 – after adding ethanol	0	156.9
Time 1	0.25	296.9
Time 2	0.50	317.9
Time 3	0.75	365.3
Time 4	1.00	350.0
Time 5	1.25	366.7
Time 6	1.50	336.9
Time 7	1.75	389.8
Time 8	2.00	392.1
Time 9	2.25	324.9
Time 10	2.50	351.4
Time 11	2.75	354.7
Time 12	3.00	387.5
Time 13	3.25	394.8
Time 14	3.50	362.9
Time 15	4.50	377.7
Time 16	5.50	369.8
Time 17	6.50	370.4
Time 18	7.50	382.9
Time 19	8.50	382.3
Time 20	9.50	376.3
Time 21 – Day 1	25.10	419.9
Time 22 – Day 3	74.25	382.2
Time 23 – Day 7	169	444.4
Time 24 – Day 10	242	456.8
Time 25 – Day 15	361	404.4
Time 26 – Day 21	508	464.8
Time 27 – Day 27	650	462.8

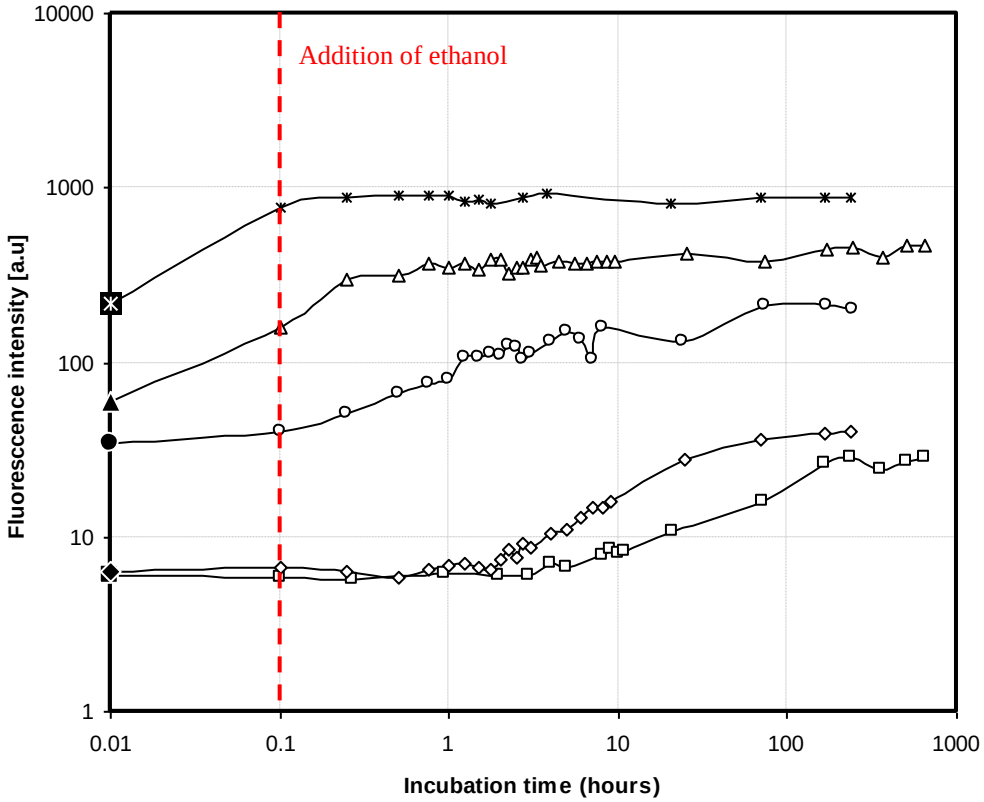
Table F.5 – ThT binding fluorescence intensity at 477 nm (emission wavelength) (λ_{exc} = 425 nm) at every incubation time at 22°C for the system of **40 mg/ml** of β-Lg in the presence of 50% of ethanol (v/v) and 10 mM buffer (pH 7).

PROTEIN CONCENTRATION: **40 mg/ml**		
Measure's name	***Incubation time (hr)***	***Fluorescence Intensity [a.u]***
Native state of β-Lg	-	215.5
Time 0 – after adding ethanol	0	766.9
Time 1	0.25	871.5
Time 2	0.50	908.5
Time 3	0.75	912.1
Time 4	1.00	914.6
Time 5	1.25	840.1
Time 6	1.50	864.5
Time 7	1.75	824.9
Time 8	2.75	875.6
Time 9	3.75	921.2
Time 10 – Day 1	20.5	820.1
Time 11 – Day 3	68.5	880.1
Time 12 – Day 7	165	877.6
Time 13 – Day 10	237	888.6

Graph F.1 - β-Lg structure evolution kinetics in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7) for different concentration of protein: 0.5 mg/ml (squares); 1 mg/ml (diamonds); 4 mg/ml (circles); 10 mg/ml (triangles) and 40 mg/ml (stars). Points intercepting the vertical axis correspond with fluorescence values just after ethanol addition. Both axes are in linear scale (raw data).

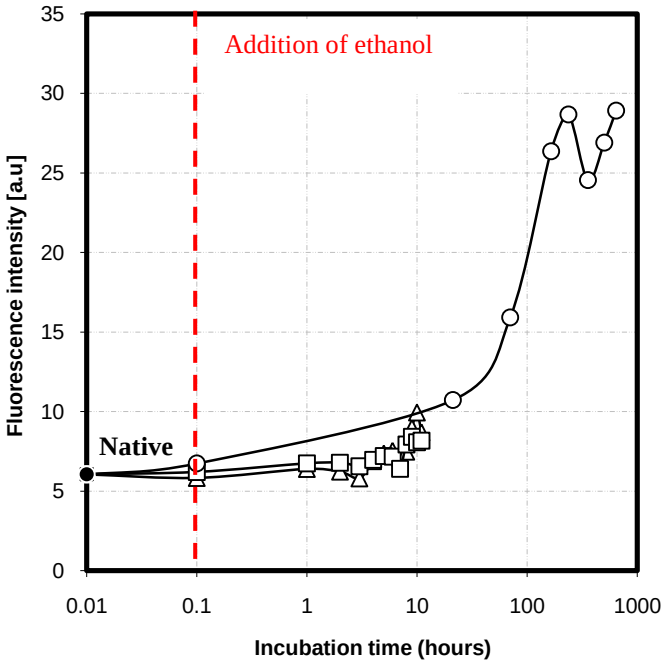


Graph F.2 - β-Lg structure evolution kinetics in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7) for different concentration of protein: 0.5 mg/ml (squares); 1 mg/ml (diamonds); 4 mg/ml (circles); 10 mg/ml (triangles) and 40 mg/ml (stars). Full symbols intercepting the vertical axis correspond with fluorescence values for the aqueous solution of the protein in 10 mM buffer (native state). The dash line indicates the point when ethanol is added to the protein solutiom. Both axes are in log-scale for clarity.

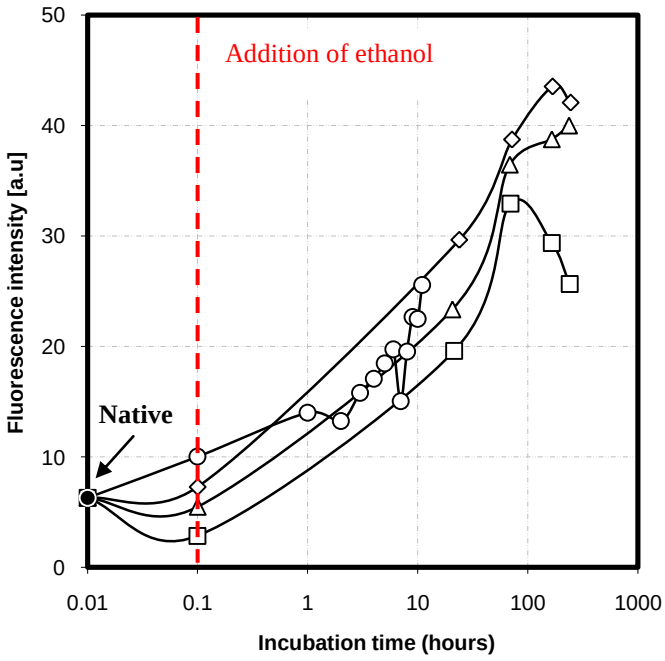


Next graphs (from F.3 to F.7) correspond to the fluorescence intensity obtained from different repetitions done at different concentrations of the protein (from 0.5 mg/ml to 40 mg/ml) plotted vs. the incubation time at 22ºC.

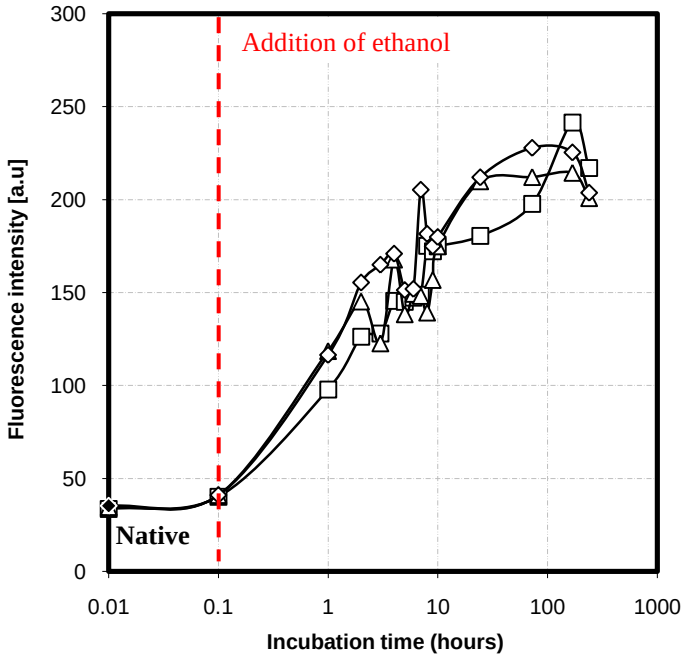
Graph F.3 – ThT-binding fluorescence intensity at 477 nm vs. incubation time for different repetitions done with samples of **0.5 mg/ml of β-Lg** in the presence of 50 % (v/v) of ethanol and 10 mM buffer at room temperature.



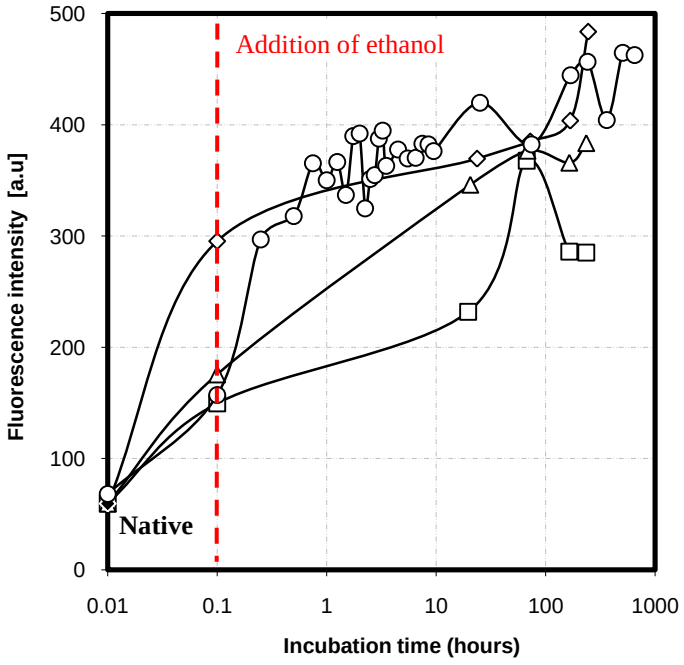
Graph F.4 - ThT-binding fluorescence intensity at 477 nm vs. incubation time for different repetitions made with samples of **1 mg/ml of β-Lg** in the presence of 50 % (v/v) of ethanol and 10 mM buffer at room temperature.



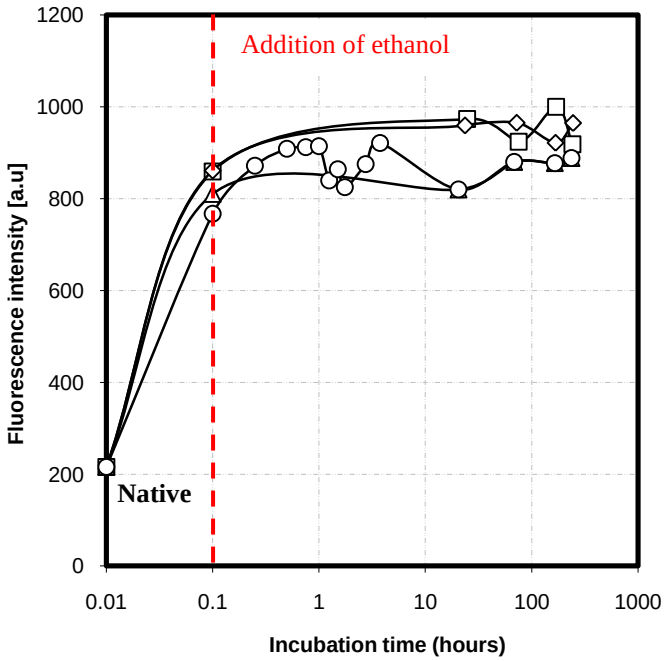
Graph F.5 - ThT-binding fluorescence intensity at 477 nm vs. incubation time for different repetitions made with samples of **4 mg/ml of β-Lg** in the presence of 50 % (v/v) of ethanol and 10 mM buffer at room temperature.



Graph F.6 - ThT-binding fluorescence intensity at 477 nm vs. incubation time for different repetitions made with samples of **10 mg/ml of β-Lg** in the presence of 50 % (v/v) of ethanol and 10 mM buffer at room temperature.



Graph F.7 - ThT-binding fluorescence intensity at 477 nm vs. incubation time for different repetitions made with samples of **40 mg/ml of β-Lg** in the presence of 50 % (v/v) of ethanol and 10 mM buffer at room temperature.



APPENDIX G

Table G.1 collects UV absorbance results used for the normalization of far and near UV CD spectra. Systems with higher protein concentration (10 and 40 mg/ml) were diluted to 1 mg/ml to get suitable far and near UV signals. Protein concentration (mg/ml) was obtained as previously mentioned in section 2.2.1 using Equation 1:

$$C = \frac{A_{280} - A_{310}}{\varepsilon \cdot cell \;\; path\, length}$$

- C is the protein concentration (mg/ml).
- A_{280} and A_{310} are the absorbance values at 280 and 310 nm respectively.
- ξ is the absorption coefficient, which depends on the type of protein. For β-Lg it has a value of 0.96 $ml \cdot mg^{-1} \cdot cm^{-1}$ (Bhattacharjee and Das, 2000).
- Cell path-length is given is cm.

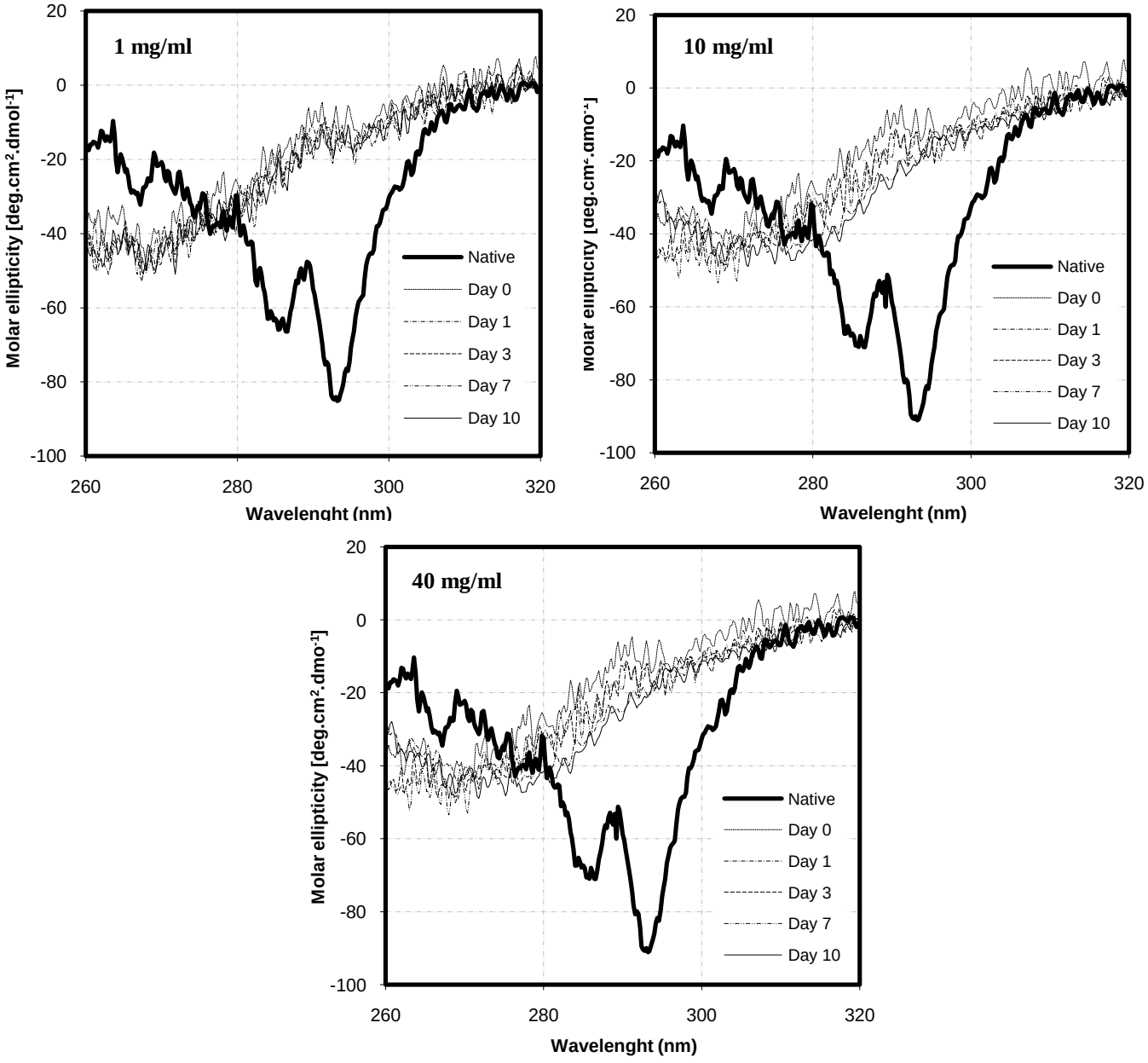
Table G.1 - Protein concentration (mg/ml) calculated from UV absorbance spectroscopy for CD data correction for each repetition and protein concentration.

	1 mg/ml			
Protein concentration (by weigh)	**Dilution required for CD analysis**	**Absorbance at 310 nm A_{310}**	**Absorbance at 280 nm A_{280}**	**UV - Protein concentration (mg/ml)**
Repetition A	No	$7.93 \cdot 10^{-3}$	0.383	0.782
Repetition B		$1.01 \cdot 10^{-2}$	0.399	0.810
Repetition C		$7.96 \cdot 10^{-3}$	0.408	0.832
	10 mg/ml			
Repetition A	Yes – to 1 mg/ml	$8.37 \cdot 10^{-3}$	0.411	0.839
Repetition B		$1.0.5 \cdot 10^{-2}$	0.401	0.8125
Repetition C		$8.18 \cdot 10^{-3}$	0.359	0.730
	40 mg/ml			
Repetition A	Yes – to 1 mg/ml	$8.02 \cdot 10^{-3}$	0.419	0.856
Repetition B		$9.89 \cdot 10^{-3}$	0.411	0.837
Repetition C		$7.92 \cdot 10^{-3}$	0.394	0.805

APPENDIX H

Near UV spectra obtained for samples of 40 mg/ml (Figure 3.12), 10 mg/ml of β-Lg (both diluted to 1 mg/ml for CD measurements) and 1 mg/ml (without further dilution) in the presence of 50% (v/v) ethanol and 10 mM buffer (pH 7) incubated at 22°C during 10 days are shown in Graphs H.1.

Graphs H.1 - Near UV spectra obtained for 1, 10 and 40 mg/ml systems incubated during 10 days at room temperature.



Far UV spectra obtained for three different repetitions for every concentration of protein (1, 10 and 40 mg/ml) is shown in the next graphs. Far UV spectra were recorded during 10 days of incubation as usual. Tables below each graph (H.2, H.3 and H.4) correspond to the quantitative analysis of the far UV spectra using the DICHROWEB server online (see APPENDIX A for further details). Spectrum corresponding to the native state of the protein was also included as a reference.

Graphs H.2 - Far UV spectra obtained for 1 mg/ml system for each repetition: A, B and C.

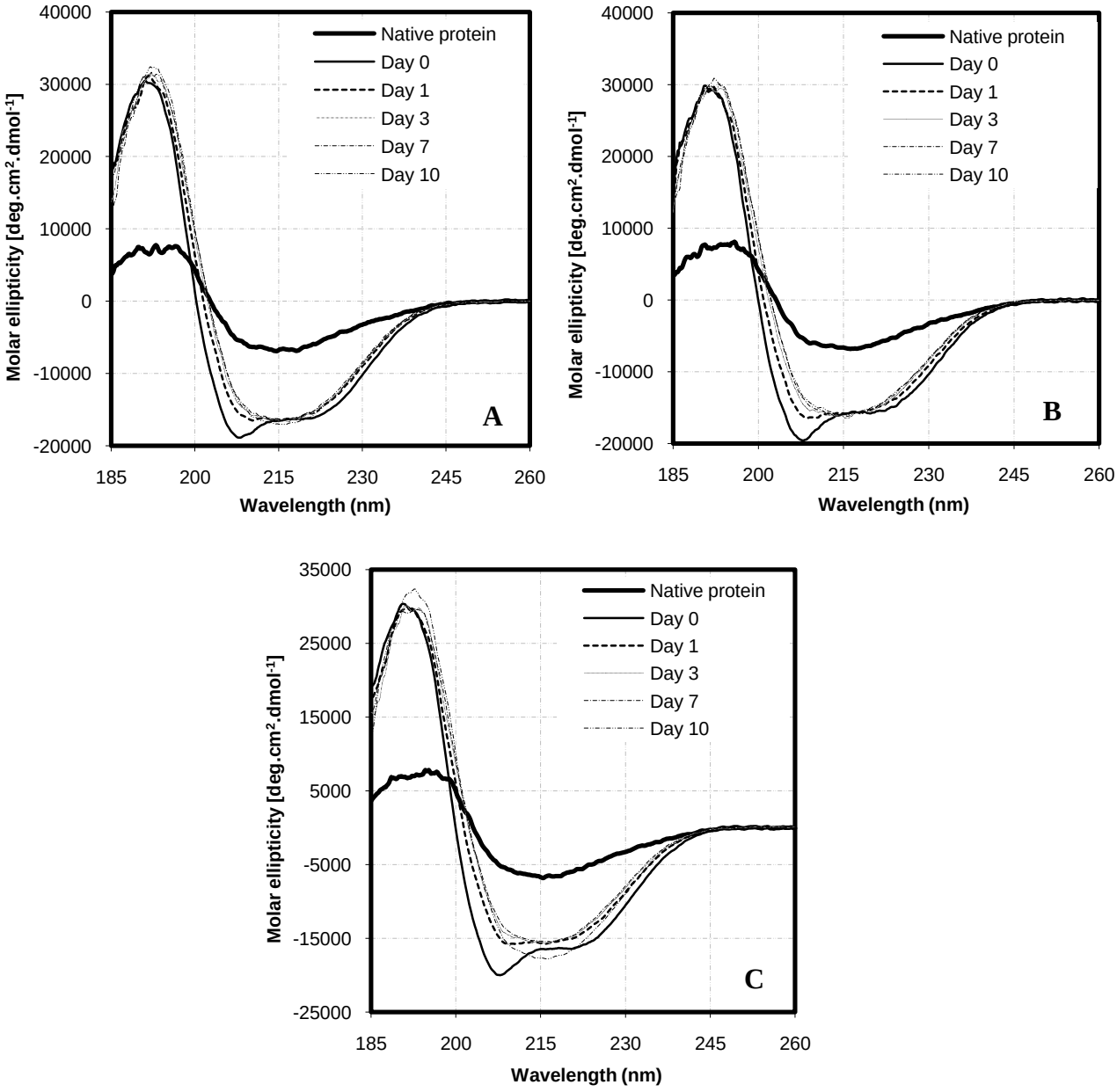
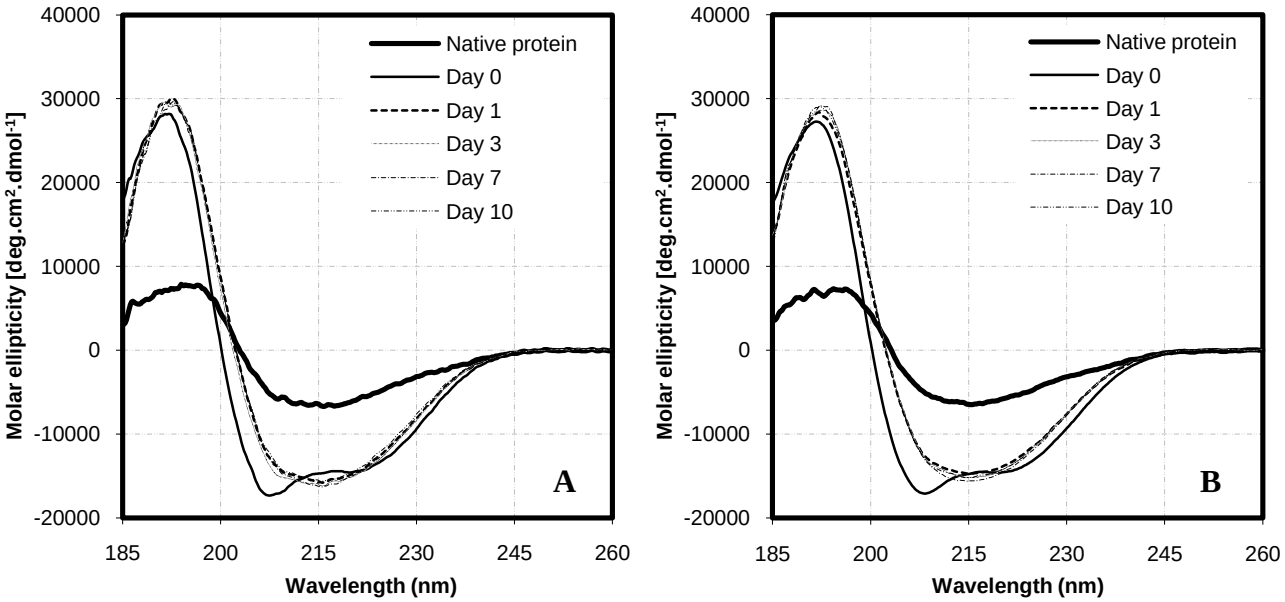


Table H.2 – Percentage content of the different types of secondary structure of β-Lg for 1 mg/ml sample in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7) during 10 days of incubation at 22ºC.

1 mg/ml of protein					
	Repetition A				
	α-helix (%)	**β-sheets (%)**	**Turn (%)**	**Random coil (%)**	**NRMSD**
Day 0	48.9	8.15	18.3	24.7	0.059
Day 1	44.9	13.7	15.9	25.7	0.036
Day 3	37.4	17.3	18.5	26.8	0.045
Day 7	36.3	18.8	18.1	26.9	0.044
Day 10	38.9	18.2	16.6	26.4	0.038
	Repetition B				
	α-helix (%)	**β-sheets (%)**	**Turn (%)**	**Random coil (%)**	**NRMSD**
Day 0	52.2	5.45	18.5	23.9	0.045
Day 1	41.5	15.4	17.0	26.2	0.034
Day 3	38.3	17.9	17.3	26.7	0.047
Day 7	35.4	19.6	17.8	27.3	0.059
Day 10	35.2	18.7	18.9	27.4	0.043
	Repetition C				
	α-helix (%)	**β-sheets (%)**	**Turn (%)**	**Random coil (%)**	**NRMSD**
Day 0	46.7	7.85	21.0	24.6	0.050
Day 1	39.9	14.3	19.4	26.5	0.045
Day 3	35.4	17.8	19.3	27.6	0.053
Day 7	35.4	19.0	18.5	27.3	0.058
Day 10	36.5	18.9	17.9	26.8	0.043

Graphs H.3 - Far UV spectra obtained for 10 mg/ml system for each repetition: A, B and C.



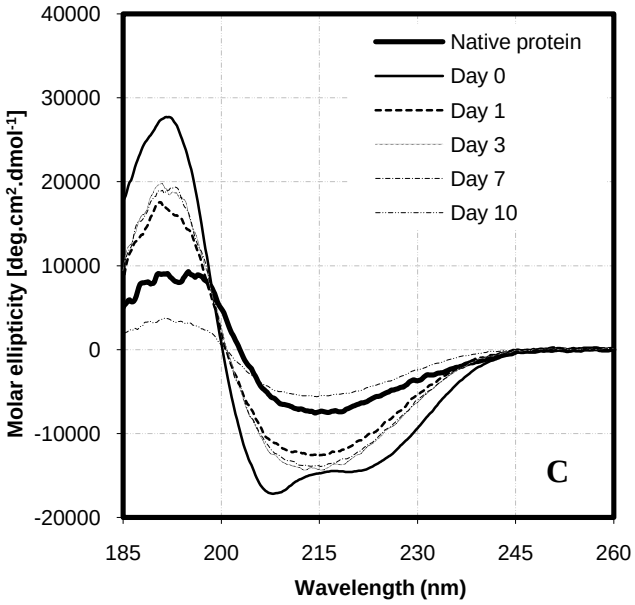


Table H.3 - Percentage content of the different types of secondary structure of β-Lg for 10 mg/ml sample in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7) during 10 days of incubation at 22ºC.

10 mg/ml of protein					
	Repetition A				
	α-helix (%)	**β-sheets (%)**	**Turn (%)**	**Random coil (%)**	**NDMSD**
Day 0	47.8	8.80	18.3	25.2	0.035
Day 1	35.1	19.2	18.5	27.4	0.053
Day 3	35.9	17.9	18.9	27.5	0.049
Day 7	34.3	20.0	18.5	27.3	0.053
Day 10	34.8	19.9	17.9	27.4	0.056
	Repetition B				
	α-helix (%)	**β-sheets (%)**	**Turn (%)**	**Random coil (%)**	**NDMSD**
Day 0	47.2	9.05	18.7	25.0	0.028
Day 1	31.9	21.5	18.6	28.2	0.049
Day 3	33.7	20.5	18.2	27.8	0.052
Day 7	33.2	20.2	18.8	27.9	0.059
Day 10	33.5	20.6	18.2	27.8	0.054
	Repetition C				
	α-helix (%)	**β-sheets (%)**	**Turn (%)**	**Random coil (%)**	**NDMSD**
Day 0	46.8	8.65	18.4	25.3	0.048
Day 1	34.2	20.5	18.5	27.0	0.069
Day 3	32.5	18.4	18.5	27.5	0.059
Day 7	35.2	20.3	18.4	27.3	0.063
Day 10	32.7	19.8	18.1	27.7	0.061

Graphs H.4 - Far UV spectra obtained for 40 mg/ml system for each repetition: A, B and C.

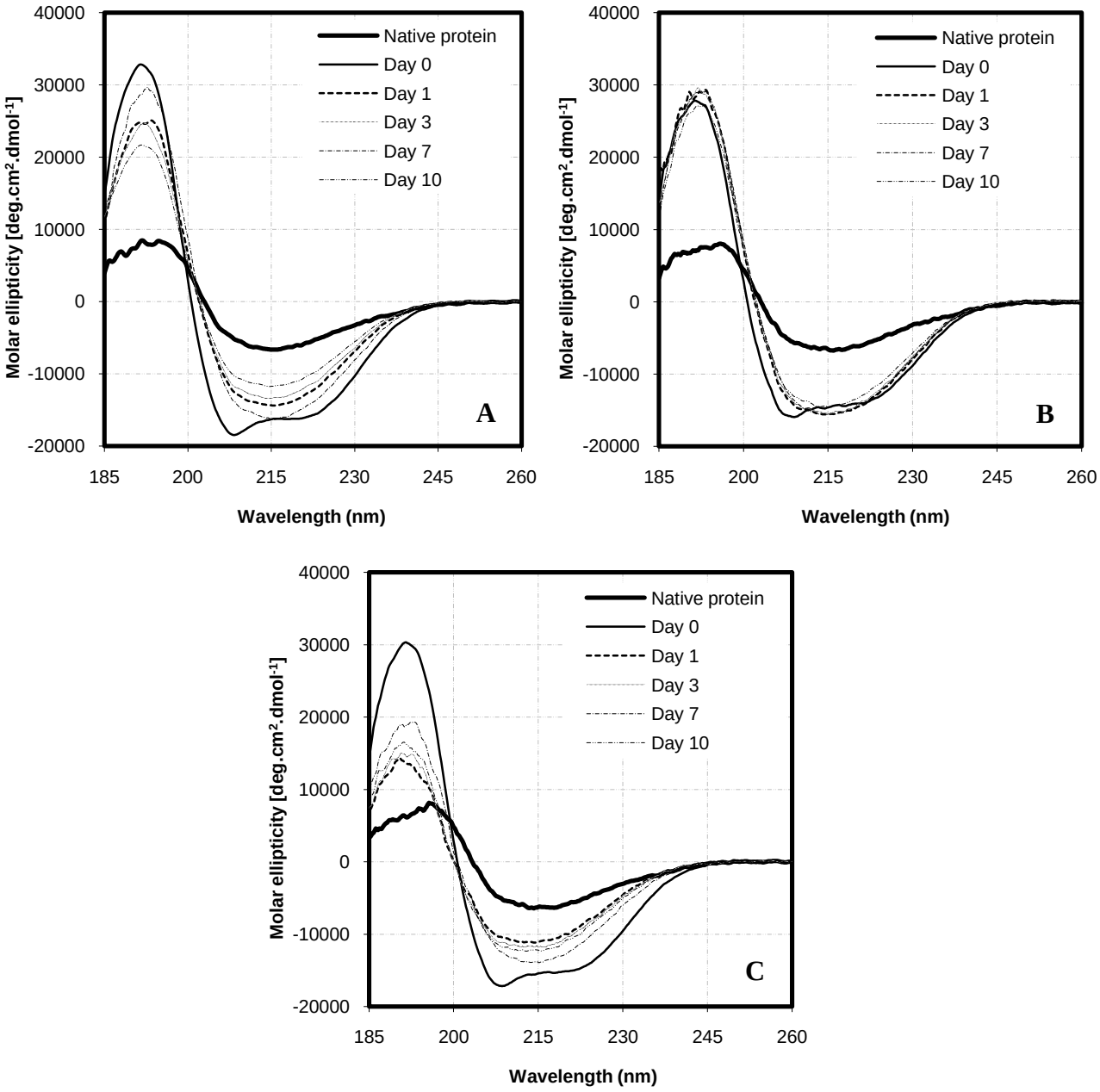
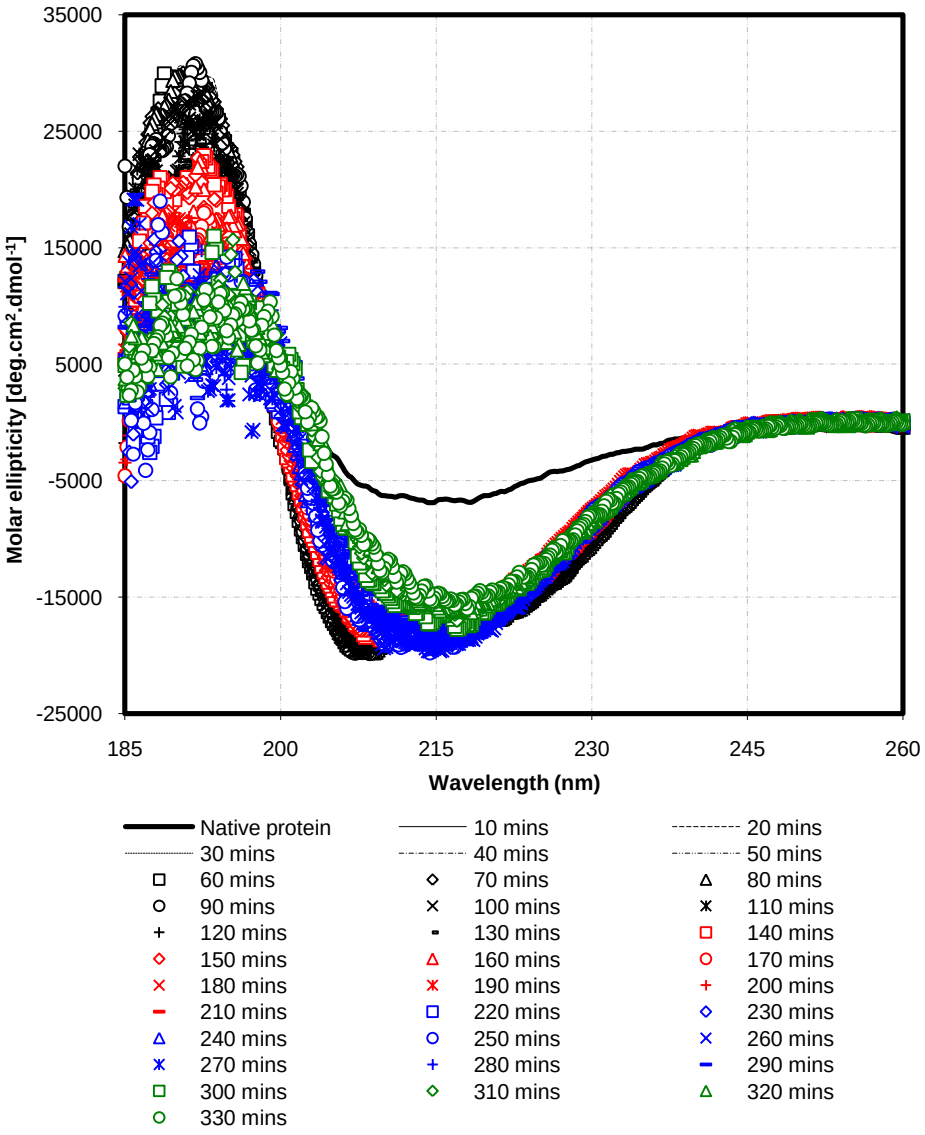


Table H.4 - Percentage content of the different types of secondary structure of β-Lg for 40 mg/ml sample in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7) during 10 days of incubation at 22ºC.

40 mg/ml of protein					
	Repetition A				
	α-helix (%)	**β-sheets (%)**	**Turn (%)**	**Random coil (%)**	**NRMSD**
Day 0	36.2	8.25	25.4	30.2	0.069
Day 1	30.4	21.9	19.3	28.3	0.064
Day 3	30.4	21.9	19.1	28.6	0.059
Day 7	21.5	28.4	20.5	29.7	0.044
Day 10	27.2	24.5	19.1	29.3	0.045
	Repetition B				
	α-helix (%)	**β-sheets (%)**	**Turn (%)**	**Random coil (%)**	**NRMSD**
Day 0	40.7	13.3	19.6	26.5	0.037
Day 1	38.4	18.3	16.3	27.1	0.043
Day 3	34.3	19.6	18.3	27.9	0.056
Day 7	34.7	19.5	18.0	27.8	0.058
Day 10	31.4	21.7	19.0	28.1	0.061
	Repetition C				
	α-helix (%)	**β-sheets (%)**	**Turn (%)**	**Random coil (%)**	**NSMRD**
Day 0	39.2	10.9	24.1	29.0	0.061
Day 1	35.2	19.4	17.3	27.0	0.070
Day 3	32.8	20.7	17.8	28.0	0.062
Day 7	30.7	31.7	19.1	28.2	0.073
Day 10	29.5	22.6	20.1	28.0	0.067

Far UV spectra for a system of 1 mg/ml of β-Lg in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7) recorded every 10 minutes are shown in Graph H.5. A total of 33 spectra were recorded in 5 hours and 20 minutes of incubation time at 22ºC. Some of these results correspond with those presented in Figure 3.17 where the number of spectra shown was reduced for clarity.

Graph H.5 - Far UV spectra for 1 mg/ml of β-Lg in 50 % (v/v) ethanol and 10 mM buffer (pH 7) recorded every 10 minutes and up to 330 minutes at 22ºC.



Graph H.5 shows gradual changes in the shape of far UV spectra going from a characteristic α-helix pattern (black spectra) right after adding ethanol, to a shape more characteristic for β-sheets structure (green spectra) for the longest time of incubation. A more accurate analysis was done by mathematical deconvolution of far UV spectra shown in Graph H.5. However, quantitative analysis only led to accurate

results for the first 10 spectra in Graph H.5 (up to 100 minutes of incubation), with NRMSD values less than 1. Graph H.6 shows these spectra. Quantitative analysis is tabulated in Table H.5, including values reported for the native protein as reference.

Graph H.6 - Far UV spectra for 1 mg/ml of β-Lg in 50 % (v/v) ethanol and 10 mM buffer (pH 7) recorded every 10 minutes and up to 100 minutes of incubation at 22ºC.

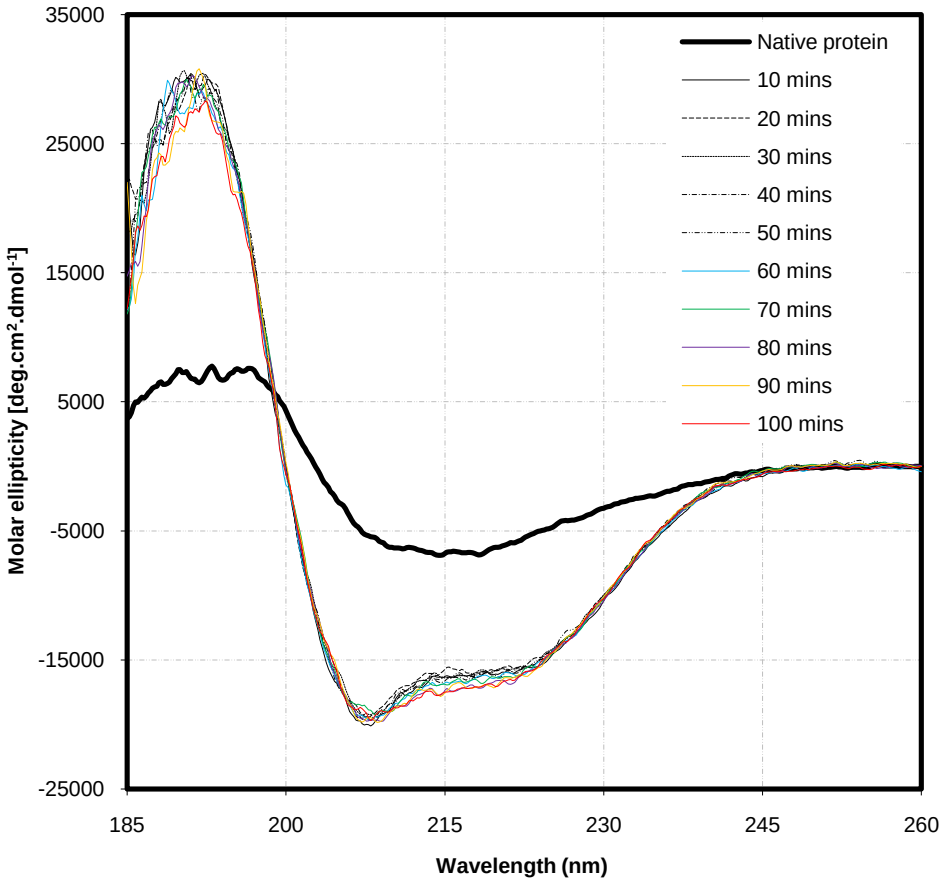


Table H.5 – Percentage content of the different types of secondary structure for a sample of 1 mg/ml of β-Lg in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7) measured every ten minutes during 100 minutes at 22ºC. Values calculated previously (section 3.1) for the native samples also included for reference.

Incubation time (hours)	α-helix (%)	β-sheet (%)	β-turns (%)	Random coil (%)	NRSMD
Native	13.9	37.4	21.8	26.9	0.031
0 – addition of ethanol	51.9	5.55	17.9	24.6	0.063
0.17	56.3	2.10	16.2	23.8	0.057
0.33	50.1	7.05	17.2	25.1	0.056
0.50	50.5	6.80	18.2	24.6	0.050
0.67	41.2	11.3	22.0	25.7	0.073
0.83	42.8	9.20	21.6	26.5	0.065
1.00	41.3	10.4	21.9	26.4	0.063
1.17	43.1	10.6	19.9	20.5	0.059
1.33	43.6	9.15	20.6	26.7	0.053

The rest of the spectra obtained between 2 and 5 hours did not lead to accurate results when fitted (NRSMD > 0.1). Therefore, it cannot be concluded that there is a second transition between percentages of α-helix and β-sheets when the former becomes less than the latest despite the shape of the far UV spectra, more similar to a major content of β-sheet. These results indicate that major changes in the secondary structure of the protein takes place in the first 24 hours after adding ethanol, with a drastic jump from majority of β-sheets in the native state of the protein to a majority of α-helix when ethanol is present in solution. The evolution of both structures goes in opposite ways thereafter (α-helix deceasing and β-sheet increasing). Results suggest that after 3 days of incubation changes in the secondary structure have concluded for every concentration of protein, being the percentages of α-helix and β-sheets different to those in the native protein: the percentages of β-sheet does not longer reach a values as high as in the a native β-Lg dimer (37 %).

APPENDIX I

ATR-IR results obtained for samples with 10 mg/ml and 40 mg/ml of β-Lg measured in short intervals of time are presented in Graphs I.1 and I.2. For a better comprehension it was decided to tabulate absorbance values at the wavenumbers associated to the most relevant information from IR analysis, i.e., intra-molecular β-sheet (absorbance at 1623 cm^{-1}), α-helix (absorbance at 1655 cm^{-1}) and inter-molecular β-sheet (absorbance at 1684 cm^{-1}) (Tables I.1 and I.2). Only some of the spectra obtained are shown for clarity since a total of 56 spectra were collected over a period of 5 hours for each protein concentration.

Graph I.1 - Secondary structure transitions kinetics by ATR-IR spectra recording for a system of 10 mg/ml in the presence of 50 % ethanol (v/v) and 10 mM buffer (pH 7) at 22ºC. Native spectra also included for reference. Arrows indicate the evolution of every type of secondary structure in the first states of incubation time.

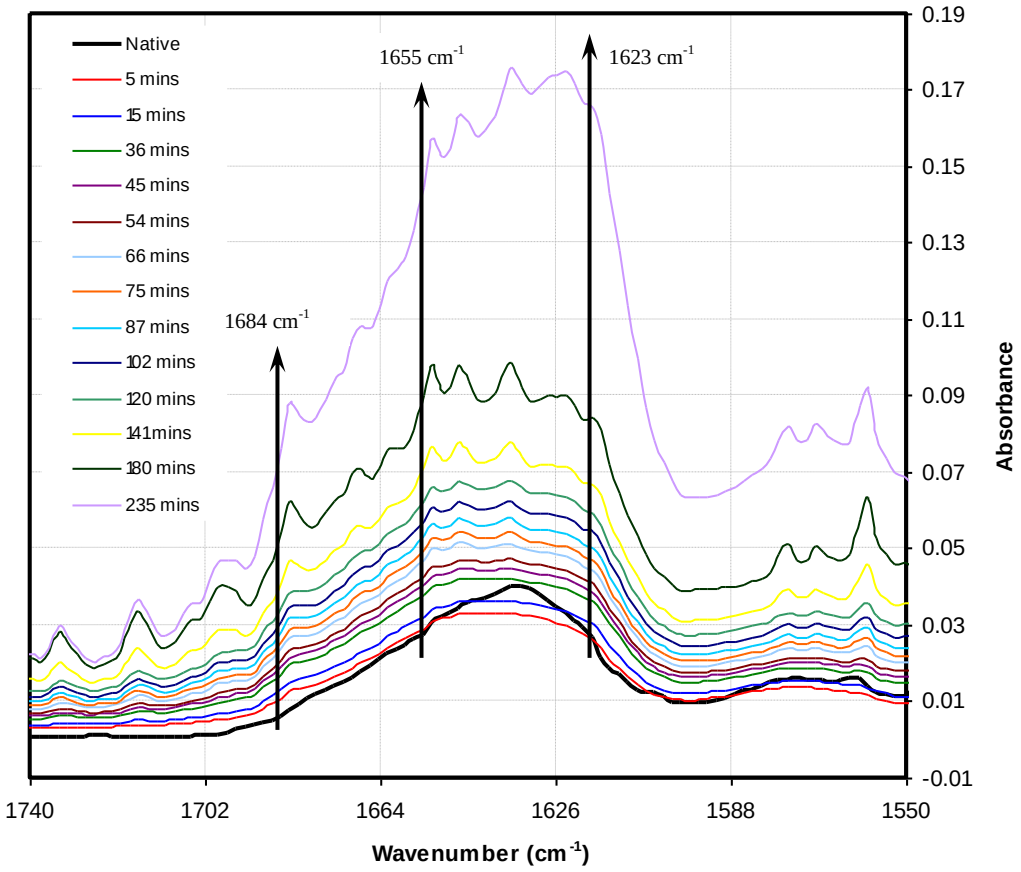


Table I.1 – ATR-IR absorbance data at wavenumbers associated to different secondary structures of the protein vs. the incubation time for the system with 10 mg/ml of protein concentration in the presence of 50% (v/v) ethanol and 10 mM buffer (pH 7) at 22°C.

Incubation time (hours)	Intramolecular β-sheet 1623 cm^{-1}	α-helix 1655 cm^{-1}	Intermolecular β-sheet 1684 cm^{-1}	Incubation time (hours)	Intramolecular β-sheet 1623 cm^{-1}	α-helix 1655 cm^{-1}	Intermolecular β-sheet 1684 cm^{-1}
Native	$3.32 \cdot 10^{-2}$	$2.72 \cdot 10^{-2}$	$7.60 \cdot 10^{-3}$				
0 – Addition of ethanol	$3.19 \cdot 10^{-2}$	$3.20 \cdot 10^{-2}$	$1.63 \cdot 10^{-2}$	1.40	$5.57 \cdot 10^{-2}$	$5.80 \cdot 10^{-2}$	$3.73 \cdot 10^{-2}$
0.05	$3.40 \cdot 10^{-2}$	$3.40 \cdot 10^{-2}$	$1.87 \cdot 10^{-2}$	1.45	$5.63 \cdot 10^{-2}$	$5.85 \cdot 10^{-2}$	$3.77 \cdot 10^{-2}$
0.10	$3.52 \cdot 10^{-2}$	$3.51 \cdot 10^{-2}$	$1.95 \cdot 10^{-2}$	1.50	$5.68 \cdot 10^{-2}$	$5.93 \cdot 10^{-2}$	$3.83 \cdot 10^{-2}$
0.15	$3.65 \cdot 10^{-2}$	$3.63 \cdot 10^{-2}$	$2.08 \cdot 10^{-2}$	1.55	$5.75 \cdot 10^{-2}$	$6.03 \cdot 10^{-2}$	$3.91 \cdot 10^{-2}$
0.20	$3.81 \cdot 10^{-2}$	$3.91 \cdot 10^{-2}$	$2.32 \cdot 10^{-2}$	1.60	$5.8 \cdot 10^{-2}$	$6.12 \cdot 10^{-2}$	$3.99 \cdot 10^{-2}$
0.25	$3.98 \cdot 10^{-2}$	$4.11 \cdot 10^{-2}$	$2.54 \cdot 10^{-2}$	1.65	$5.85 \cdot 10^{-2}$	$6.20 \cdot 10^{-2}$	$4.04 \cdot 10^{-2}$
0.30	$4.09 \cdot 10^{-2}$	$4.24 \cdot 10^{-2}$	$2.6 \cdot 10^{-2}$	1.70	$5.89 \cdot 10^{-2}$	$6.24 \cdot 10^{-2}$	$4.06 \cdot 10^{-2}$
0.35	$4.14 \cdot 10^{-2}$	$4.29 \cdot 10^{-2}$	$2.66 \cdot 10^{-2}$	1.75	$5.94 \cdot 10^{-2}$	$6.29 \cdot 10^{-2}$	$4.07 \cdot 10^{-2}$
0.40	$4.25 \cdot 10^{-2}$	$4.40 \cdot 10^{-2}$	$2.75 \cdot 10^{-2}$	1.80	$6.00 \cdot 10^{-2}$	$6.35 \cdot 10^{-2}$	$4.10 \cdot 10^{-2}$
0.45	$4.34 \cdot 10^{-2}$	$4.51 \cdot 10^{-2}$	$2.83 \cdot 10^{-2}$	1.85	$6.04 \cdot 10^{-2}$	$6.40 \cdot 10^{-2}$	$4.15 \cdot 10^{-2}$
0.50	$4.41 \cdot 10^{-2}$	$4.57 \cdot 10^{-2}$	$2.88 \cdot 10^{-2}$	1.90	$6.09 \cdot 10^{-2}$	$6.46 \cdot 10^{-2}$	$4.17 \cdot 10^{-2}$
0.55	$4.49 \cdot 10^{-2}$	$4.68 \cdot 10^{-2}$	$2.97 \cdot 10^{-2}$	1.95	$6.12 \cdot 10^{-2}$	$6.48 \cdot 10^{-2}$	$4.17 \cdot 10^{-2}$
0.60	$4.56 \cdot 10^{-2}$	$4.76 \cdot 10^{-2}$	$3.04 \cdot 10^{-2}$	2.00	$6.18 \cdot 10^{-2}$	$6.53 \cdot 10^{-2}$	$4.17 \cdot 10^{-2}$
0.65	$4.65 \cdot 10^{-2}$	$4.86 \cdot 10^{-2}$	$3.12 \cdot 10^{-2}$	2.05	$6.23 \cdot 10^{-2}$	$6.59 \cdot 10^{-2}$	$4.21 \cdot 10^{-2}$
0.70	$4.72 \cdot 10^{-2}$	$4.95 \cdot 10^{-2}$	$3.18 \cdot 10^{-2}$	2.10	$6.26 \cdot 10^{-2}$	$6.64 \cdot 10^{-2}$	$4.23 \cdot 10^{-2}$
0.75	$4.77 \cdot 10^{-2}$	$4.99 \cdot 10^{-2}$	$3.21 \cdot 10^{-2}$	2.15	$6.31 \cdot 10^{-2}$	$6.69 \cdot 10^{-2}$	$4.27 \cdot 10^{-2}$
0.80	$4.82 \cdot 10^{-2}$	$5.06 \cdot 10^{-2}$	$3.27 \cdot 10^{-2}$	2.20	$6.37 \cdot 10^{-2}$	$6.74 \cdot 10^{-2}$	$4.29 \cdot 10^{-2}$
0.85	$4.87 \cdot 10^{-2}$	$5.08 \cdot 10^{-2}$	$3.29 \cdot 10^{-2}$	2.25	$6.40 \cdot 10^{-2}$	$6.79 \cdot 10^{-2}$	$4.33 \cdot 10^{-2}$
0.90	$4.91 \cdot 10^{-2}$	$5.04 \cdot 10^{-2}$	$3.32 \cdot 10^{-2}$	2.30	$6.48 \cdot 10^{-2}$	$6.88 \cdot 10^{-2}$	$4.39 \cdot 10^{-2}$
0.95	$4.96 \cdot 10^{-2}$	$5.16 \cdot 10^{-2}$	$3.29 \cdot 10^{-2}$	2.35	$6.55 \cdot 10^{-2}$	$6.98 \cdot 10^{-2}$	$4.46 \cdot 10^{-2}$
1.00	$5.03 \cdot 10^{-2}$	$5.22 \cdot 10^{-2}$	$3.33 \cdot 10^{-2}$	2.40	$6.63 \cdot 10^{-2}$	$7.06 \cdot 10^{-2}$	$4.55 \cdot 10^{-2}$
1.05	$5.09 \cdot 10^{-2}$	$5.28 \cdot 10^{-2}$	$3.36 \cdot 10^{-2}$	2.45	$6.70 \cdot 10^{-2}$	$7.15 \cdot 10^{-2}$	$4.61 \cdot 10^{-2}$
1.10	$5.15 \cdot 10^{-2}$	$5.34 \cdot 10^{-2}$	$3.40 \cdot 10^{-2}$	2.50	$6.78 \cdot 10^{-2}$	$7.24 \cdot 10^{-2}$	$4.69 \cdot 10^{-2}$
1.15	$5.21 \cdot 10^{-2}$	$5.40 \cdot 10^{-2}$	$3.43 \cdot 10^{-2}$	3.17	$2.08 \cdot 10^{-2}$	$1.69 \cdot 10^{-2}$	$2.00 \cdot 10^{-3}$
1.20	$5.27 \cdot 10^{-2}$	$5.46 \cdot 10^{-2}$	$3.47 \cdot 10^{-2}$	3.62	$6.13 \cdot 10^{-2}$	$4.59 \cdot 10^{-2}$	$1.54 \cdot 10^{-2}$
1.25	$5.35 \cdot 10^{-2}$	$5.57 \cdot 10^{-2}$	$3.56 \cdot 10^{-2}$	4.00	$6.42 \cdot 10^{-2}$	$5.20 \cdot 10^{-2}$	$2.20 \cdot 10^{-2}$
1.30	$5.44 \cdot 10^{-2}$	$565 \cdot 10^{-2}$	$3.61 \cdot 10^{-2}$	4.50	$6.94 \cdot 10^{-2}$	$5.90 \cdot 10^{-2}$	$2.72 \cdot 10^{-2}$
1.35	$5.50 \cdot 10^{-2}$	$572 \cdot 10^{-2}$	$3.66 \cdot 10^{-2}$	5.00	$7.50 \cdot 10^{-2}$	$6.55 \cdot 10^{-2}$	$3.16 \cdot 10^{-2}$

Graph I.2 - Secondary structure transitions kinetics by ATR-IR spectra recording for a system of 40 mg/ml in the presence of 50 % ethanol (v/v) and 10 mM buffer (pH 7) at 22ºC. Native spectra also included for reference. Arrows indicate the evolution of every type of secondary structure in the first states of incubation time.

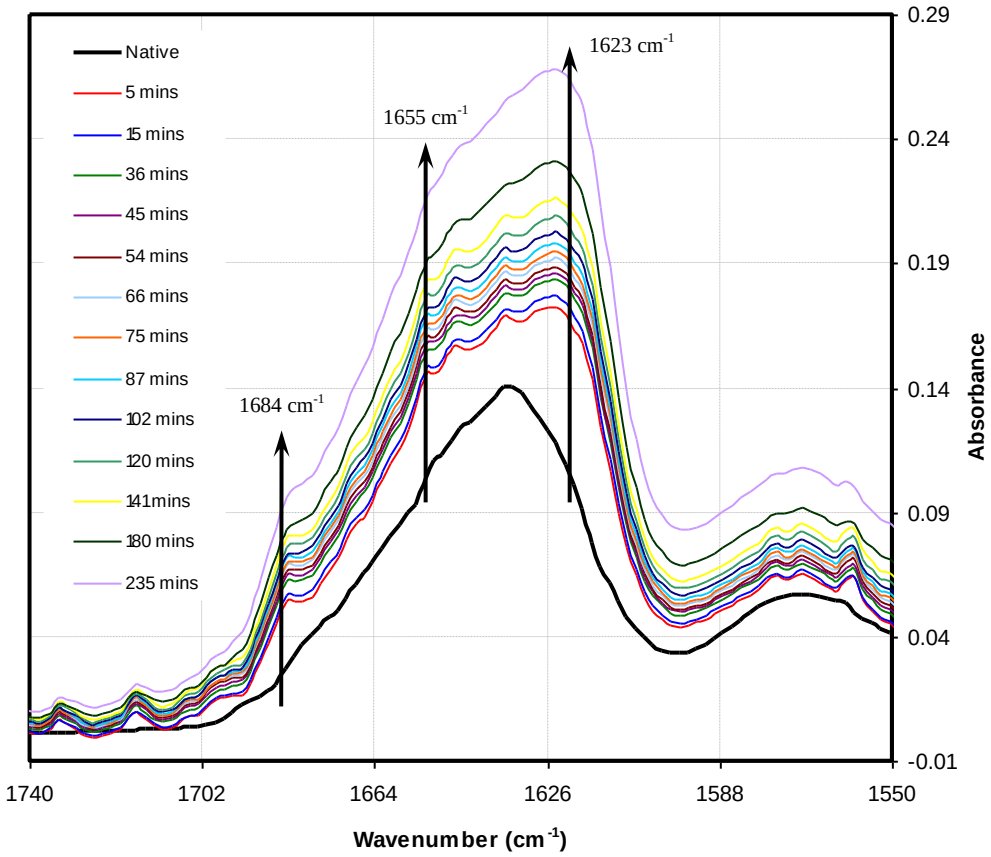
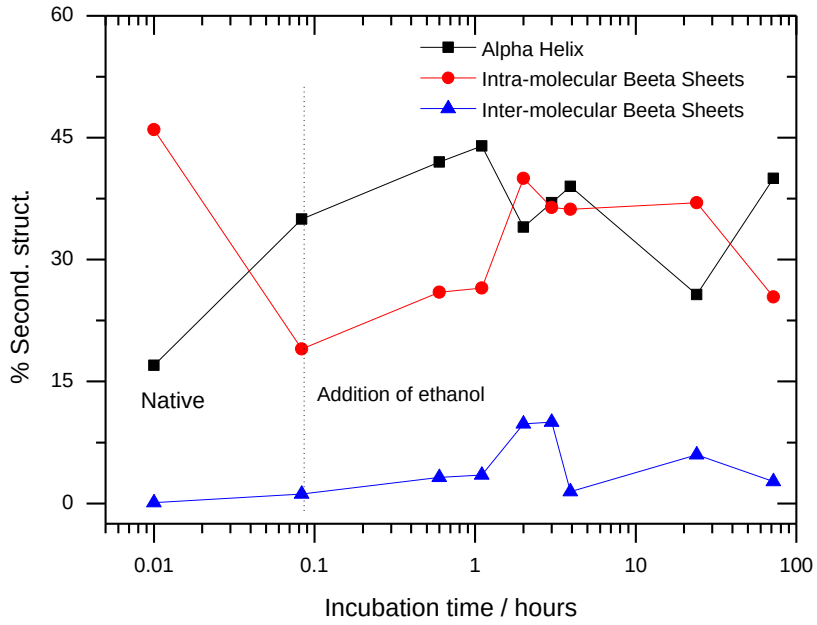


Table I.2 – ATR-IR absorbance data at wavenumbers associated to different secondary structures of the protein vs. the incubation time for the system with 40 mg/ml of protein concentration in the presence of 50% (v/v) ethanol and 10 mM buffer (pH 7) at 22°C.

Incubation time (hours)	Intra molecular β-sheet 1623 cm^{-1}	α-helix 1655 cm^{-1}	Inter molecular β-sheet 1684 cm^{-1}	Incubation time (hours)	Intramolecular β-sheet 1623 cm^{-1}	α-helix 1655 cm^{-1}	Intermolecular β-sheet 1684 cm^{-1}
Native	$1.18 \cdot 10^{-1}$	$9.32 \cdot 10^{-2}$	$2.69 \cdot 10^{-2}$				
0 – Addition of ethanol	$1.66 \cdot 10^{-1}$	$1.32 \cdot 10^{-2}$	$5.16 \cdot 10^{-2}$	1.40	$1.99 \cdot 10^{-1}$	$1.57 \cdot 10^{-1}$	$7.18 \cdot 10^{-2}$
0.05	$1.73 \cdot 10^{-1}$	$1.34 \cdot 10^{-1}$	$5.46 \cdot 10^{-2}$	1.45	$1.99 \cdot 10^{-1}$	$1.57 \cdot 10^{-1}$	$7.21 \cdot 10^{-2}$
0.10	$1.74 \cdot 10^{-1}$	$1.35 \cdot 10^{-1}$	$5.52 \cdot 10^{-2}$	1.50	$2.00 \cdot 10^{-1}$	$1.58 \cdot 10^{-1}$	$7.23 \cdot 10^{-2}$
0.15	$1.76 \cdot 10^{-1}$	$1.36 \cdot 10^{-1}$	$5.62 \cdot 10^{-2}$	1.55	$2.01 \cdot 10^{-1}$	$1.59 \cdot 10^{-1}$	$7.25 \cdot 10^{-2}$
0.20	$1.77 \cdot 10^{-1}$	$1.37 \cdot 10^{-1}$	$5.69 \cdot 10^{-2}$	1.60	$2.02 \cdot 10^{-1}$	$1.59 \cdot 10^{-1}$	$7.29 \cdot 10^{-2}$
0.25	$1.78 \cdot 10^{-1}$	$1.38 \cdot 10^{-1}$	$5.77 \cdot 10^{-2}$	1.65	$2.03 \cdot 10^{-1}$	$1.60 \cdot 10^{-1}$	$7.31 \cdot 10^{-2}$
0.30	$1.79 \cdot 10^{-1}$	$1.39 \cdot 10^{-1}$	$5.84 \cdot 10^{-2}$	1.70	$2.04 \cdot 10^{-1}$	$1.61 \cdot 10^{-1}$	$7.36 \cdot 10^{-2}$
0.35	$1.80 \cdot 10^{-1}$	$1.39 \cdot 10^{-1}$	$5.91 \cdot 10^{-2}$	1.75	$2.05 \cdot 10^{-1}$	$1.61 \cdot 10^{-1}$	$7.41 \cdot 10^{-2}$
0.40	$1.81 \cdot 10^{-1}$	$1.41 \cdot 10^{-1}$	$5.99 \cdot 10^{-2}$	1.80	$2.06 \cdot 10^{-1}$	$1.62 \cdot 10^{-1}$	$7.47 \cdot 10^{-2}$
0.45	$1.82 \cdot 10^{-1}$	$1.41 \cdot 10^{-1}$	$6.09 \cdot 10^{-2}$	1.85	$2.07 \cdot 10^{-1}$	$1.63 \cdot 10^{-1}$	$7.53 \cdot 10^{-2}$
0.50	$1.83 \cdot 10^{-1}$	$1.43 \cdot 10^{-1}$	$6.16 \cdot 10^{-2}$	1.90	$2.08 \cdot 10^{-1}$	$1.64 \cdot 10^{-1}$	$7.59 \cdot 10^{-2}$
0.55	$1.83 \cdot 10^{-1}$	$1.43 \cdot 10^{-1}$	$6.21 \cdot 10^{-2}$	1.95	$2.10 \cdot 10^{-1}$	$1.65 \cdot 10^{-1}$	$7.64 \cdot 10^{-2}$
0.60	$1.85 \cdot 10^{-1}$	$1.44 \cdot 10^{-1}$	$6.29 \cdot 10^{-2}$	2.00	$2.10 \cdot 10^{-1}$	$1.66 \cdot 10^{-1}$	$7.7 \cdot 10^{-2}$
0.65	$1.86 \cdot 10^{-1}$	$1.45 \cdot 10^{-1}$	$6.38 \cdot 10^{-2}$	2.05	$2.11 \cdot 10^{-1}$	$1.66 \cdot 10^{-1}$	$7.75 \cdot 10^{-2}$
0.70	$1.87 \cdot 10^{-1}$	$1.46 \cdot 10^{-1}$	$6.44 \cdot 10^{-2}$	2.10	$2.12 \cdot 10^{-1}$	$1.67 \cdot 10^{-1}$	$7.78 \cdot 10^{-2}$
0.75	$1.88 \cdot 10^{-1}$	$1.47 \cdot 10^{-1}$	$6.53 \cdot 10^{-2}$	2.15	$2.13 \cdot 10^{-1}$	$1.68 \cdot 10^{-1}$	$7.84 \cdot 10^{-2}$
0.80	$1.88 \cdot 10^{-1}$	$1.48 \cdot 10^{-1}$	$6.60 \cdot 10^{-2}$	2.20	$2.14 \cdot 10^{-1}$	$1.69 \cdot 10^{-1}$	$7.89 \cdot 10^{-2}$
0.85	$1.89 \cdot 10^{-1}$	$1.49 \cdot 10^{-1}$	$6.67 \cdot 10^{-2}$	2.25	$2.15 \cdot 10^{-1}$	$1.70 \cdot 10^{-1}$	$7.95 \cdot 10^{-2}$
0.90	$1.90 \cdot 10^{-1}$	$1.50 \cdot 10^{-1}$	$6.72 \cdot 10^{-2}$	2.30	$2.17 \cdot 10^{-1}$	$1.71 \cdot 10^{-1}$	$8.00 \cdot 10^{-2}$
0.95	$1.91 \cdot 10^{-1}$	$1.50 \cdot 10^{-1}$	$6.77 \cdot 10^{-2}$	2.35	$2.18 \cdot 10^{-1}$	$1.71 \cdot 10^{-1}$	$8.04 \cdot 10^{-2}$
1.00	$1.92 \cdot 10^{-1}$	$1.51 \cdot 10^{-1}$	$6.79 \cdot 10^{-2}$	2.40	$2.19 \cdot 10^{-1}$	$1.72 \cdot 10^{-1}$	$8.05 \cdot 10^{-2}$
1.05	$1.92 \cdot 10^{-1}$	$1.52 \cdot 10^{-1}$	$6.85 \cdot 10^{-2}$	2.45	$2.20 \cdot 10^{-1}$	$1.72 \cdot 10^{-1}$	$8.06 \cdot 10^{-2}$
1.10	$1.93 \cdot 10^{-1}$	$1.52 \cdot 10^{-1}$	$6.91 \cdot 10^{-2}$	2.50	$2.31 \cdot 10^{-1}$	$1.80 \cdot 10^{-1}$	$8.33 \cdot 10^{-2}$
1.15	$1.94 \cdot 10^{-1}$	$1.53 \cdot 10^{-1}$	$6.95 \cdot 10^{-2}$	3.17	$2.48 \cdot 10^{-1}$	$1.90 \cdot 10^{-1}$	$8.87 \cdot 10^{-2}$
1.20	$1.95 \cdot 10^{-1}$	$1.54 \cdot 10^{-1}$	$6.97 \cdot 10^{-2}$	3.62	$2.60 \cdot 10^{-1}$	$1.98 \cdot 10^{-1}$	$9.28 \cdot 10^{-2}$
1.25	$1.96 \cdot 10^{-1}$	$1.55 \cdot 10^{-1}$	$7.04 \cdot 10^{-2}$	4.00	$2.68 \cdot 10^{-1}$	$2.04 \cdot 10^{-1}$	$9.58 \cdot 10^{-2}$
1.30	$1.97 \cdot 10^{-1}$	$1.56 \cdot 10^{-1}$	$7.11 \cdot 10^{-2}$	4.50	$3.09 \cdot 10^{-1}$	$2.29 \cdot 10^{-1}$	$1.09 \cdot 10^{-1}$
1.35	$1.98 \cdot 10^{-1}$	$1.56 \cdot 10^{-1}$	$7.14 \cdot 10^{-2}$	5.00	$3.56 \cdot 10^{-1}$	$2.60 \cdot 10^{-1}$	$1.27 \cdot 10^{-1}$

Data shown above were deconvoluted by Dr. Nadeem Javid using GRAMS 8 software (provided in the instrument used for measurements – see section 2.2.4) to obtain the percentage content of the different secondary structures: α-helix, intra- and inter-molecular β-sheets. Results shown in Graph I.3 indicates that ATR-IR measurements lead to very similar structural information qualitatively as CD (section 3.3.2) where α-helix increases just after adding ethanol and then decreases slowly until 10 days of incubation, while the β-sheet content decreases after adding ethanol and then starts increasing until 10 days of incubation at room temperature. There are differences between ATR-IR and CD in the time period for evolution of the different secondary structures which are because both techniques address different structures (CD, α-helix and ATR-IR, β-sheets) better than other, as mentioned in the results (section 3.3.2.2). Moreover, the protein concentrations are also different for both kinetics measurements (limitation of CD to use high concentration).

Graph I.5 – Percentages of the secondary structure of the protein from deconvolution of SAXS results (obtained by Dr. Nadeem Javid).



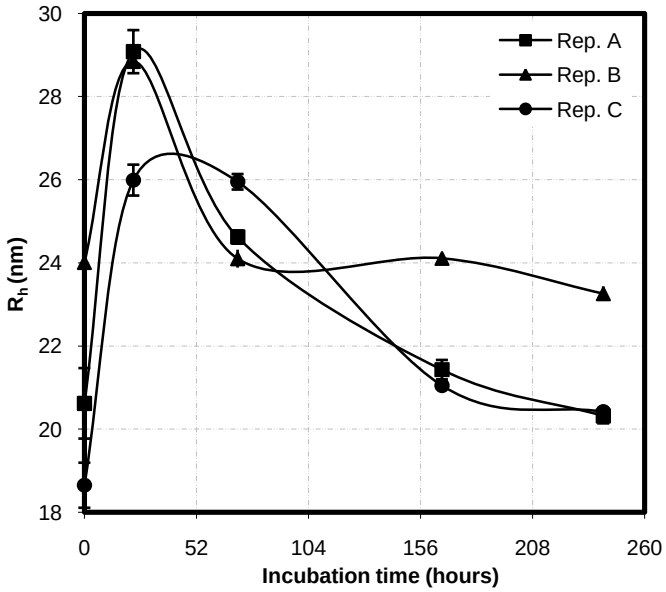
APPENDIX J

Table J.1 summarises the apparent hydrodynamic radii calculated for every repetition of experiments (R_{hi}) for systems with different concentrations of protein (1, 10 and 40 mg/ml) in the presence of 50% (v/v) ethanol and 10 mM buffer (pH 7) measured over 10 days of incubation at 22ºC. Different repetitions are named with an alphabet letter followed by the protein concentration of the system; hence "system A1" indicates repetition "A" of the system with "1 mg/ml of β-Lg." Data in Table J.1 are plotted vs. the incubation time in Graphs J.1 to J.3.

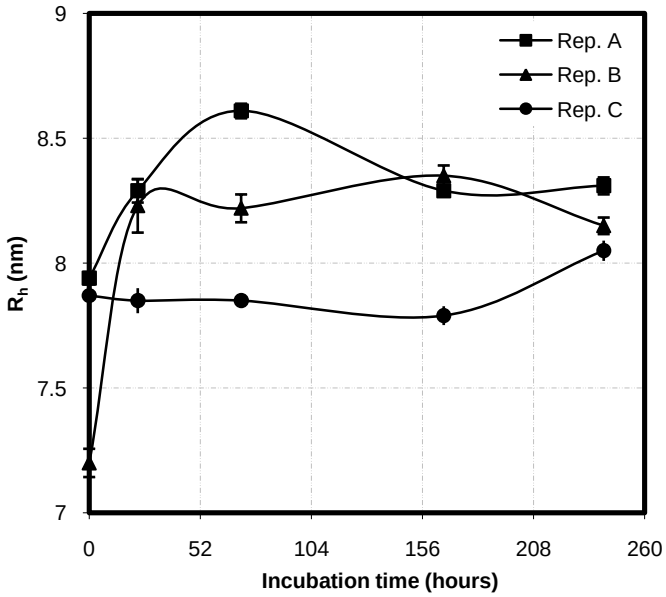
Table J.1 - Apparent hydrodynamic radius (R_{hi}, nm) calculated for individual experiments over 10 days of incubation at 22ºC in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7).

1 mg/ml			
Rh (nm)	**A1**	**B1**	**C1**
Day 0	$21 \pm 4.2 \cdot 10^{-1}$	$24 \pm 3.8 \cdot 10^{-1}$	$19 \pm 2.7 \cdot 10^{-1}$
Day 1	$29 \pm 2.6 \cdot 10^{-1}$	$29 \pm 1.0 \cdot 10^{-1}$	$26 \pm 1.9 \cdot 10^{-1}$
Day 3	$25 \pm 7.9 \cdot 10^{-2}$	$24 \pm 1.1 \cdot 10^{-1}$	$26 \pm 9.3 \cdot 10^{-2}$
Day 7	$21 \pm 1.2 \cdot 10^{-1}$	$24 \pm 6.2 \cdot 10^{-2}$	$21 \pm 3.8 \cdot 10^{-2}$
Day 10	$20 \pm 9.1 \cdot 10^{-2}$	$23 \pm 8.8 \cdot 10^{-2}$	$20 \pm 4.4 \cdot 10^{-2}$
10 mg/ml			
Rh (nm)	**A10**	**B10**	**C10**
Day 0	$7.9 \pm 1.2 \cdot 10^{-2}$	$7.2 \pm 2.8 \cdot 10^{-2}$	$7.9 \pm 1.0 \cdot 10^{-2}$
Day 1	$8.3 \pm 2.3 \cdot 10^{-2}$	$8.2 \pm 5.4 \cdot 10^{-2}$	$7.9 \pm 2.5 \cdot 10^{-2}$
Day 3	$8.6 \pm 1.5 \cdot 10^{-2}$	$8.2 \pm 2.8 \cdot 10^{-2}$	$7.9 \pm 8.4 \cdot 10^{-3}$
Day 7	$8.3 \pm 9.7 \cdot 10^{-3}$	$8.4 \pm 2.1 \cdot 10^{-2}$	$7.8 \pm 1.9 \cdot 10^{-2}$
Day 10	$8.3 \pm 1.7 \cdot 10^{-2}$	$8.2 \pm 1.7 \cdot 10^{-2}$	$8.1 \pm 2.0 \cdot 10^{-2}$
40 mg/ml			
Rh (nm)	**A40**	**B40**	**C40**
Day 0	$4.6 \pm 1.9 \cdot 10^{-2}$	$5.4 \pm 6.7 \cdot 10^{-3}$	$4.8 \pm 3.9 \cdot 10^{-3}$
Day 1	$5.7 \pm 1.2 \cdot 10^{-2}$	$5.5 \pm 1.3 \cdot 10^{-2}$	$5.6 \pm 1.3 \cdot 10^{-2}$
Day 3	$6.4 \pm 1.6 \cdot 10^{-2}$	$6.5 \pm 2.0 \cdot 10^{-2}$	$6.5 \pm 1.3 \cdot 10^{-1}$
Day 7	$6.8 \pm 2.0 \cdot 10^{-2}$	$7.1 \pm 1.6 \cdot 10^{-2}$	$6.7 \pm 1.0 \cdot 10^{-1}$
Day 10	$6.5 \pm 2.3 \cdot 10^{-2}$	$6.9 \pm 1.3 \cdot 10^{-2}$	$6.9 \pm 3.8 \cdot 10^{-2}$

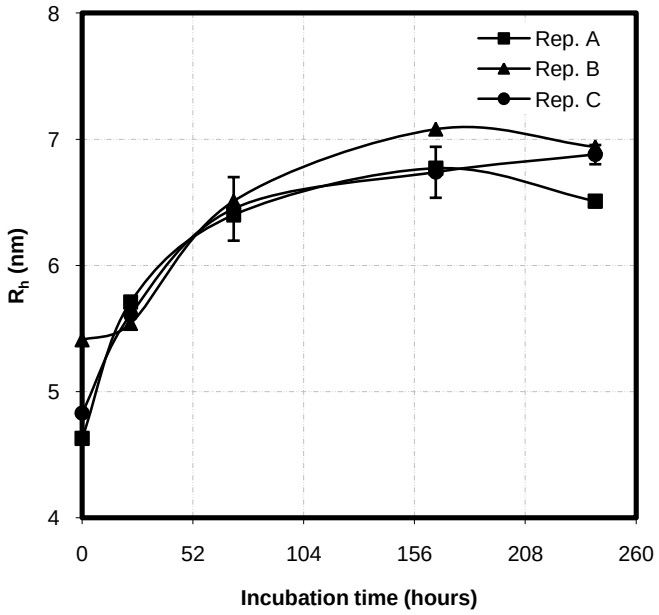
Graph J.1 – Apparent hydrodynamic radius (R_{hi}, nm) vs. incubation time for 1 mg/ml of β-Lg in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7), measured just after adding ethanol (intercept in the vertical axis) and after 1, 3, 7 and 10 days of incubation at 22ºC.



Graph J.2 - Apparent hydrodynamic radius (R_{hi}, nm) vs. incubation time for 10 mg/ml of β-Lg in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH7), measured just after adding ethanol (intercept in the vertical axis) and after 1, 3, 7 and 10 days of incubation at 22ºC.



Graph J.3 - Apparent hydrodynamic radius (R_{hi}, nm) vs. incubation time for 40 mg/ml of β-Lg in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH7), measured just after adding ethanol (intercept in the vertical axis) and after 1, 3, 7 and 10 days of incubation at 22ºC.



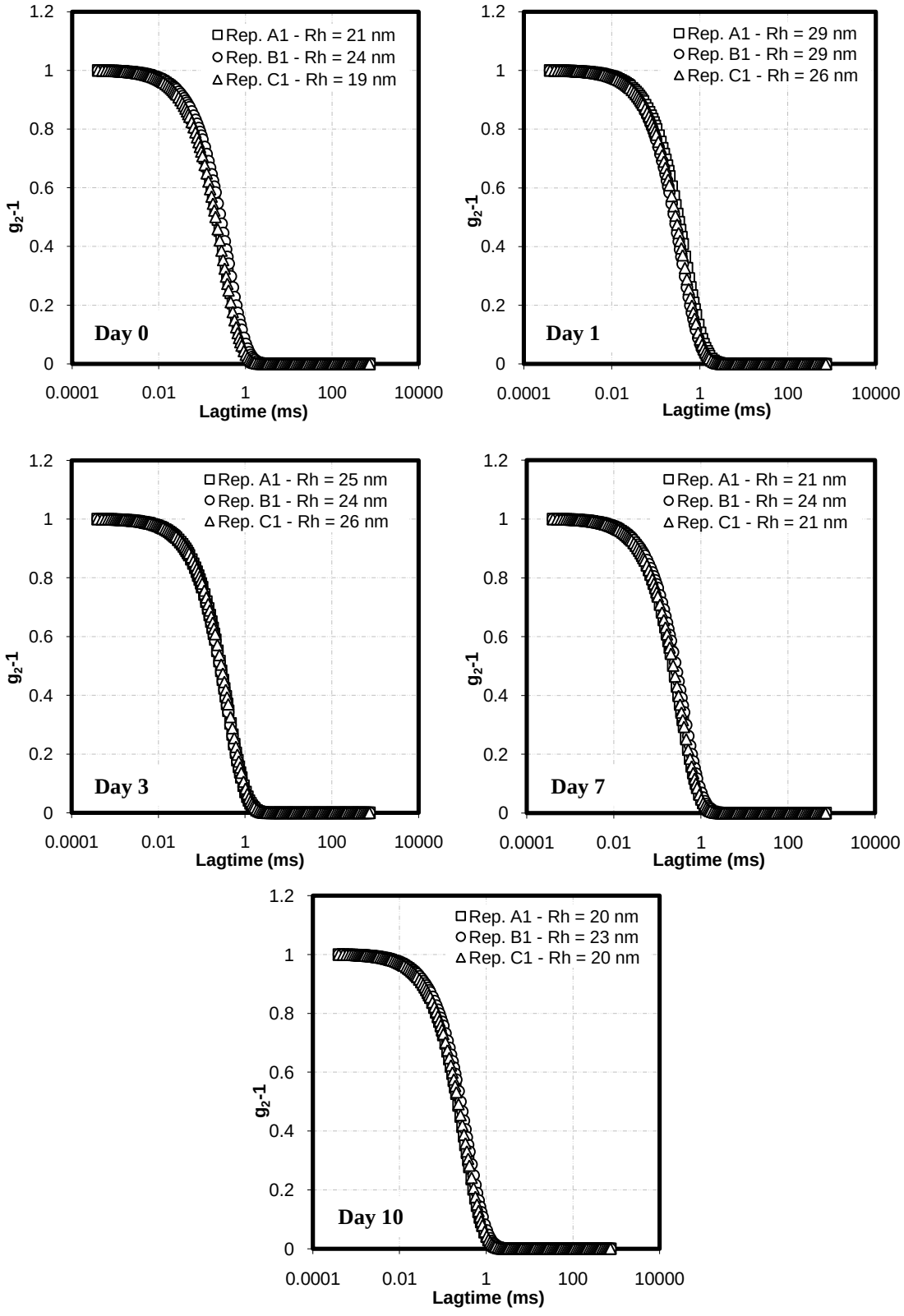
APPENDIX K

Graphs from K.1 to K.3 represent the normalised autocorrelation functions obtained for every repetition of experiments (A, B, C) for every protein concentration (1, 10 and 40 mg/ml) for each incubation time at room temperature.

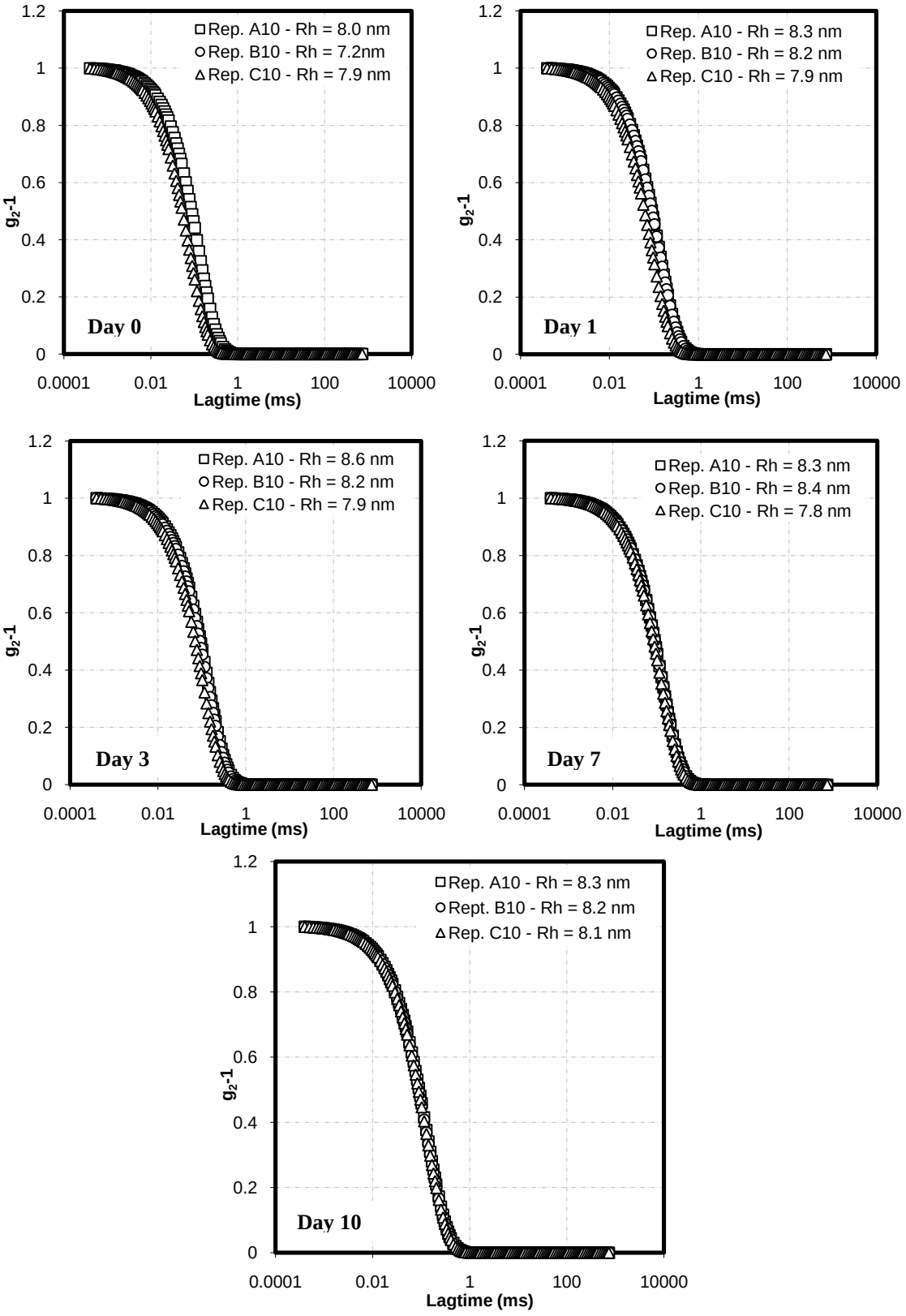
For a better understanding of these graphs, each legend shows the repetition together with the concentration of protein in the corresponding sample. Hence, "Rep. A1" indicates the repetition A done for a sample containing "1 mg/ml" of β-Lg. For every system, the concentration of buffer was 10 mM (pH 7) and ethanol concentration was 50 % (v/v). Samples were incubated at room temperature (22°C) during the specified time: from 0 (addition of ethanol) to 10 days.

The theoretical autocorrelation function obtained from the average hydrodynamic radius (R_h) shown in Table 3.7 is represented as a thick black line. However, this line is hided by the raw data in most cases due to the good fit using the cumulant method. For details about the fitting of the autocorrelation functions obtained in DLS measurements please see APPENDIX B.

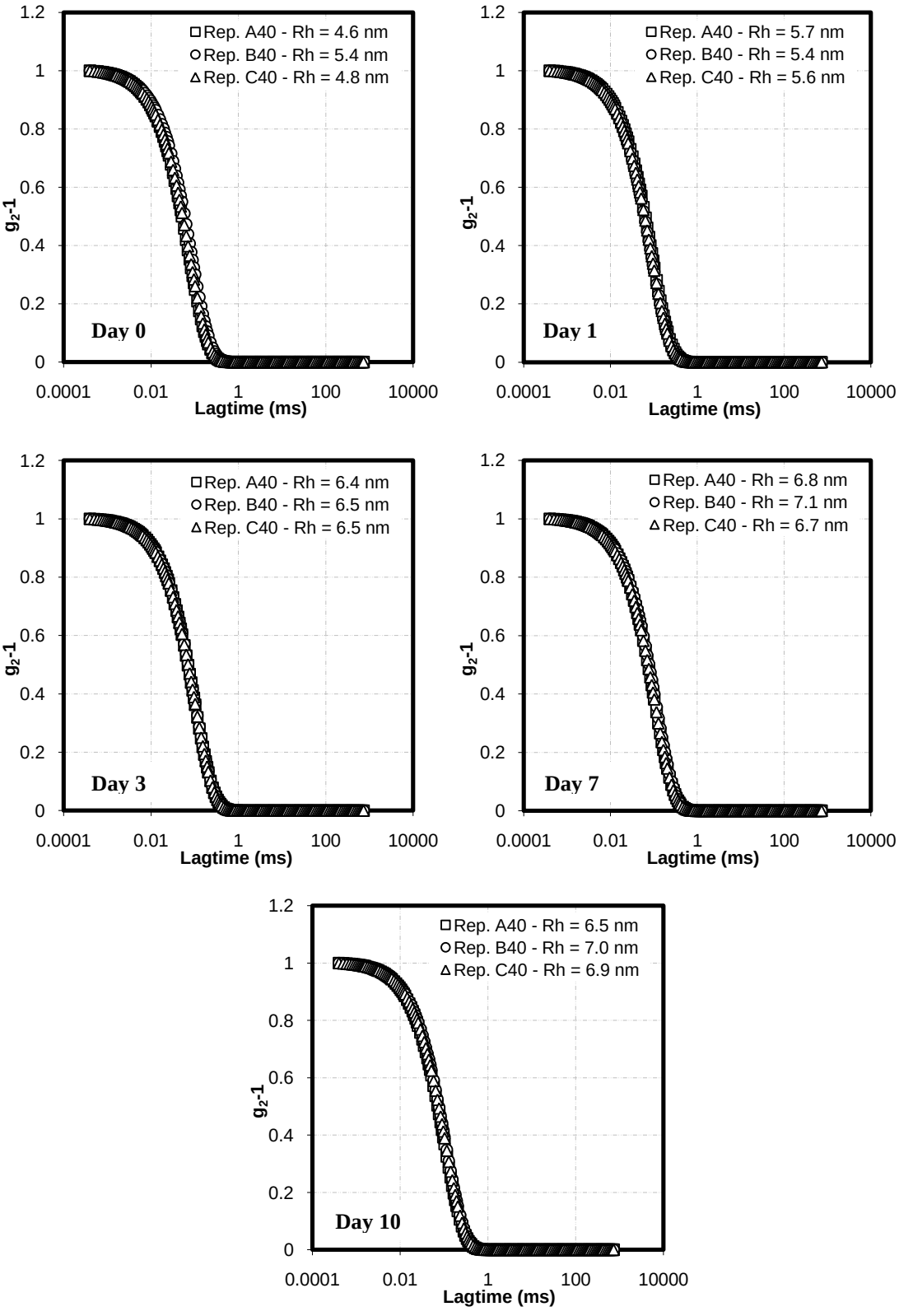
Graphs K.1 – Normalised autocorrelation functions obtained for 1 mg/ml system incubated over 10 days at 22ºC for every repetition from day 0 to day 10.



Graphs K.2 – Normalised autocorrelation functions obtained for 10 mg/ml system incubated at 22ºC over 10 days for every repetition from day 0 to day 10.



Graphs K.3 – Normalised autocorrelation functions obtained for 40 mg/ml system incubated over 10 days for every repetition from day 0 to day 10.

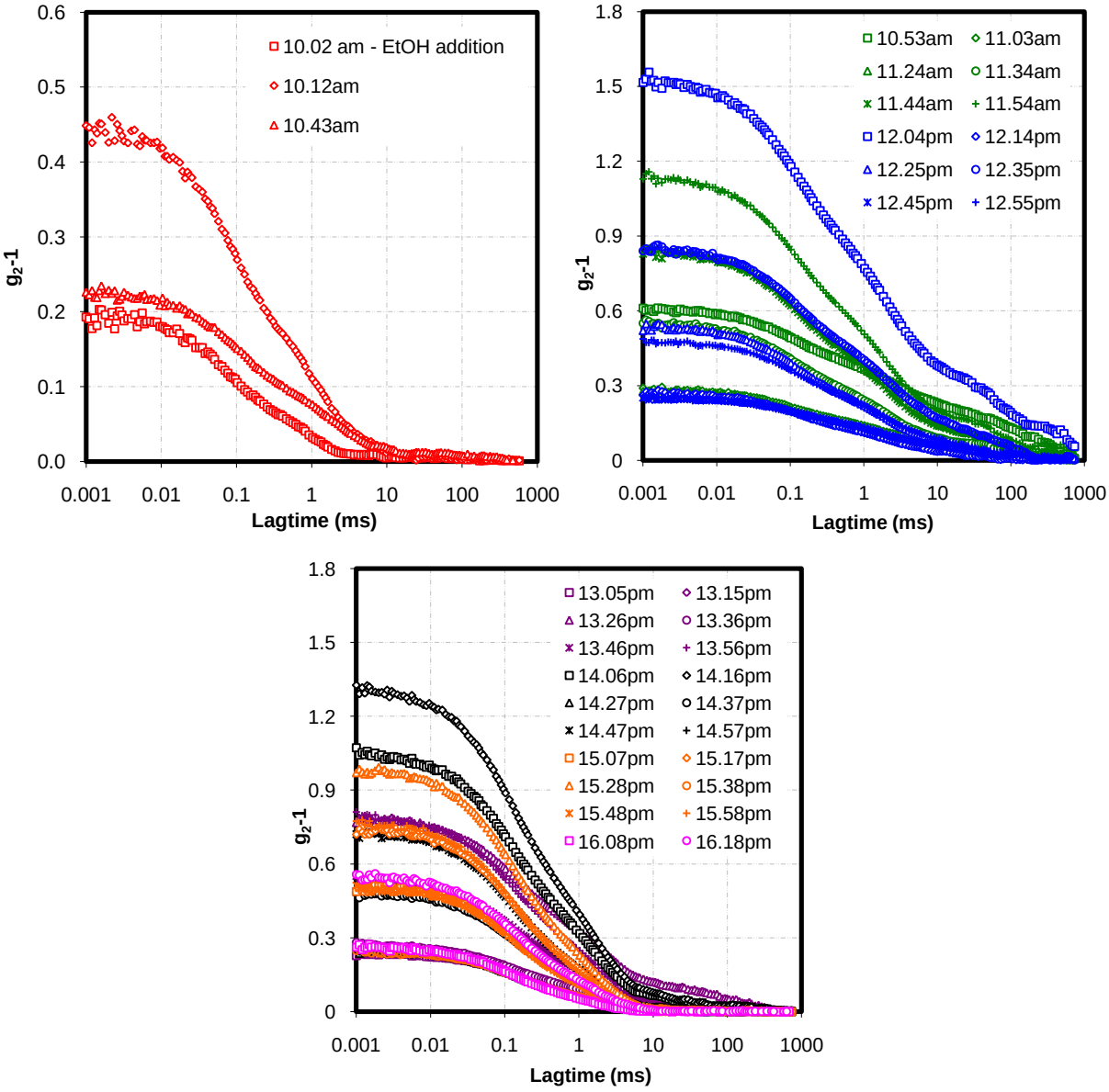


APPENDIX L

Graphs L.1 show the autocorrelation functions for the system with 1 mg/ml measured in short periods of time – from addition of ethanol and up to 6 hours of incubation at room temperature every 10 minutes – in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7). Spectra recorded are presented in three different plots for clarity, according with the hydrodynamic radius vs. incubation time shown in Graph L.2. Average values of the hydrodynamic radius (R_h) obtained from the 10-days systems for 1 mg/ml (Table 3.7) were also included in Graph L.2 to better understand the different behaviour of this sample. Table L.1 shows the values of the hydrodynamic radius (R_{hi}) obtained from the fitting of the autocorrelation functions shown in Graph L.1, together with the error of the fit. The tails observed in the autocorrelation functions in Graph L.1 are not entirely understood, since they were not observed in any other system studied before when 50 % (v/v) of ethanol was present, including 1 mg/ml during 10 days of incubation (APPENDIX K). This sample behaves clearly differently from the rest of samples studies; therefore it was not included in the general discussion. Some problems occurring with the sample during preparation and/or measurements could explain these suspicious results.

Graph L.3 shows the autocorrelation functions obtained from the system of 4 mg/ml in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7) measured in short periods of time at 22°C. Table L.3 summarises the hydrodynamic radius (R_{hi}) calculated from the fit of the autocorrelation functions in Graph L.3, together with the error of the fit.

Graph L.1 - Autocorrelation functions obtained for 1 mg/ml of β-Lg in the presence of 50% (v/v) ethanol and 10 mM buffer (pH 7) measured in short periods of time up to 6 hours at 22°C. Spectra recorded at different times of incubation are plotted in different graphs: A) from addition of ethanol up to 40 minutes; B) from 40 minutes up to 170 minutes; C) from 170 minutes up to 360 minutes.



Graph L.2 - Growth kinetics of aggregation for 1 mg/ml of protein in the presence of 50 % (v/v) ethanol and 10 mM buffer (pH 7) at 22ºC. The empty square intercepting the vertical axis corresponds to the aqueous solution of the protein (native state). Full squares correspond with the hydrodynamic radius (R_{hi}) calculated from short time DLS measurements. Empty triangles correspond with average hydrodynamic radius (R_h) obtained in the long-term DLS measurements, done for the same system (Table 3.7).

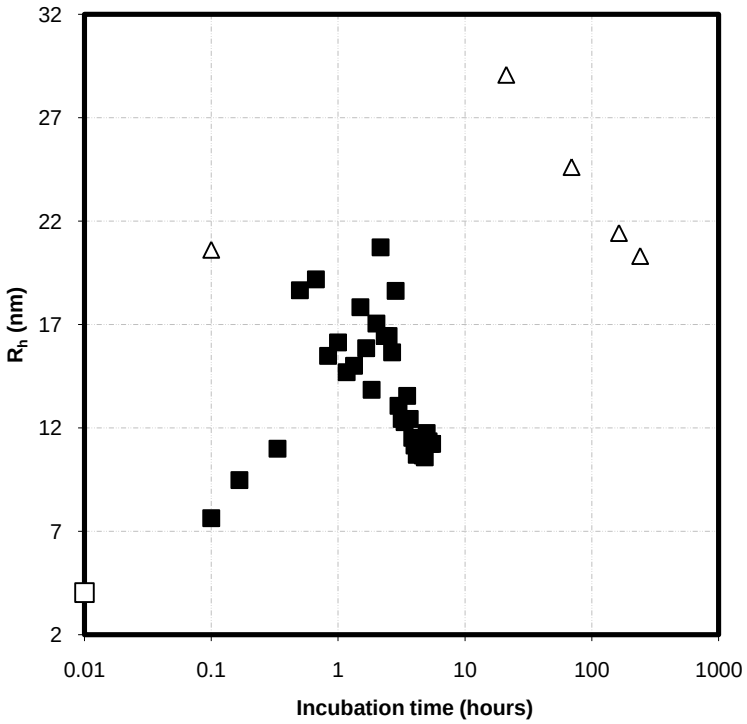


Table L.1 - Hydrodynamic radius calculated from the fitting of the autocorrelation functions presented in Graph L.1 – 1 mg/ml of protein in the presence of 50% (v/v) ethanol and 10 mM buffer (pH 7), measured in short periods of time up to 6 hours at 22°C.

Incubation time (hours)	**R_{hi} (nm)**	**Incubation time (hours)**	**R_{hi} (nm)**
0 –Addition of ethanol	$7.6 \pm 9.9 \cdot 10^{-2}$	3.33	$19 \pm 9.4 \cdot 10^{-2}$
0.17	$9.5 \pm 8.0 \cdot 10^{-2}$	3.50	$13 \pm 1.1 \cdot 10^{-1}$
0.67	$11 \pm 3.2 \cdot 10^{-1}$	3.67	$12 \pm 8.9 \cdot 10^{-2}$
0.83	$19 \pm 3.2 \cdot 10^{-1}$	3.83	$12 \pm 1.2 \cdot 10^{-1}$
1.00	$19 \pm 2.1 \cdot 10^{-1}$	4.00	$14 \pm 1.8 \cdot 10^{-1}$
1.33	$15 \pm 1.1 \cdot 10^{-1}$	4.17	$12 \pm 8.1 \cdot 10^{-2}$
1.50	$16 \pm 1.4 \cdot 10^{-1}$	4.33	$12 \pm 1.6 \cdot 10^{-1}$
1.67	$15 \pm 1.9 \cdot 10^{-1}$	4.50	$11 \pm 1.1 \cdot 10^{-1}$
1.83	$15 \pm 1.4 \cdot 10^{-1}$	4.67	$11 \pm 8.4 \cdot 10^{-2}$
2.00	$18 \pm 1.5 \cdot 10^{-1}$	4.83	$11 \pm 7.7 \cdot 10^{-2}$
2.17	$16 \pm 2.2 \cdot 10^{-1}$	5.00	$11 \pm 9.6 \cdot 10^{-2}$
2.33	$14 \pm 1.6 \cdot 10^{-1}$	5.17	$11 \pm 1.2 \cdot 10^{-1}$
2.50	$17 \pm 1.3 \cdot 10^{-1}$	5.33	$11 \pm 9.9 \cdot 10^{-2}$
2.67	$21 \pm 2.0 \cdot 10^{-1}$	5.50	$12 \pm 1.2 \cdot 10^{-1}$
2.83	$16 \pm 1.3 \cdot 10^{-1}$	5.67	$11 \pm 1.2 \cdot 10^{-1}$
3.00	$16 \pm 1.5 \cdot 10^{-1}$	5.83	$11 \pm 5.6 \cdot 10^{-2}$
3.17	$16 \pm 7.7 \cdot 10^{-2}$	6.00	$11 \pm 1.0 \cdot 10^{-1}$

Graph L.3 - Autocorrelation functions obtained for 4 mg/ml of β-Lg in the presence of 50% (v/v) ethanol and 10 mM buffer (pH 7) measured in short periods of time up to 6 hours at 22°C.

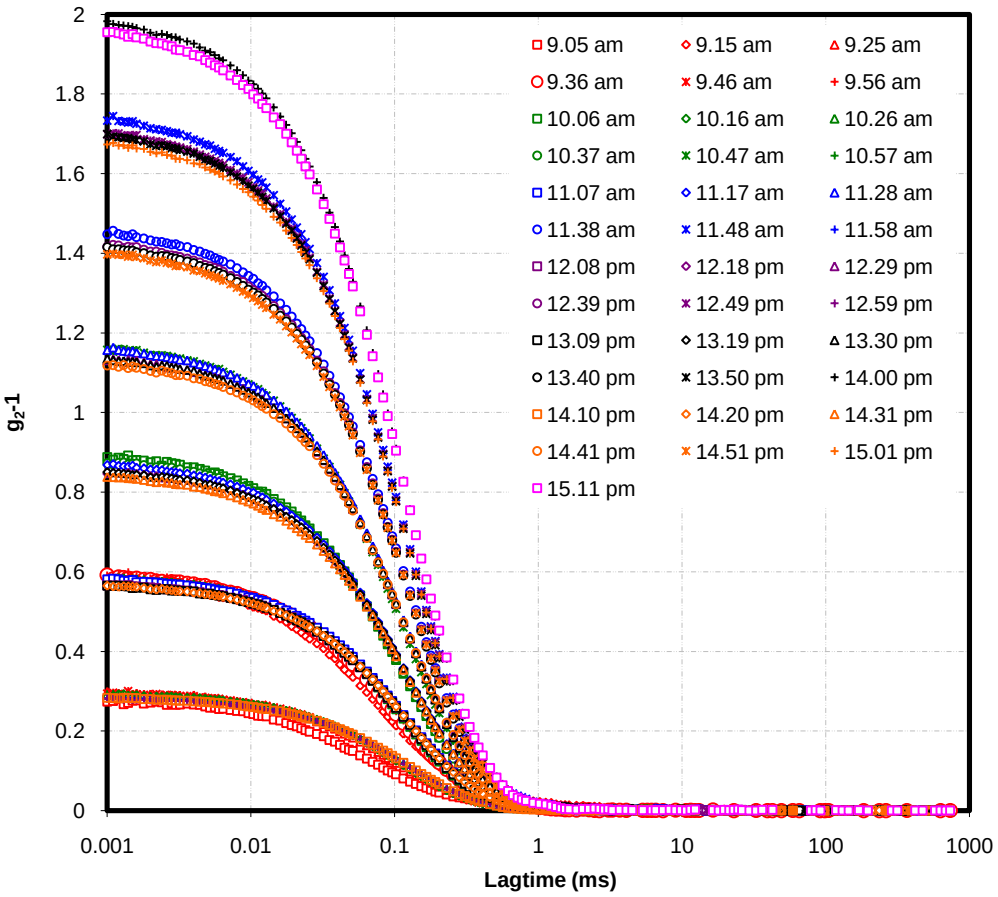


Table L.3 - Hydrodynamic radius calculated from the fit of the autocorrelation functions presented in Graph L.1 – 4 mg/ml of protein concentration in the presence of 50% (v/v) ethanol and 10 mM buffer (pH 7) measured in short periods of time up to 6 hours at 22°C.

Incubation time (hours)	**R_{hi} (nm)**	**Incubation time (hours)**	**R_{hi} (nm)**
0 –Addition of ethanol	$4.5 \pm 3.0 \cdot 10^{-2}$		
0.17	$5.2 \pm 4.0 \cdot 10^{-2}$	3.17	$7.2 \pm 2.7 \cdot 10^{-1}$
0.33	$6.1 \pm 3.7 \cdot 10^{-1}$	3.33	$7.2 \pm 2.4 \cdot 10^{-2}$
0.50	$6.2 \pm 2.1 \cdot 10^{-1}$	3.50	$7.4 \pm 1.5 \cdot 10^{-1}$
0.67	$6.8 \pm 1.8 \cdot 10^{-1}$	3.67	$7.2 \pm 2.0 \cdot 10^{-1}$
0.83	$6.5 \pm 5.2 \cdot 10^{-1}$	3.83	$7.5 \pm 1.8 \cdot 10^{-2}$
1.00	$6.6 \pm 3.8 \cdot 10^{-1}$	4.00	$7.3 \pm 2.5 \cdot 10^{-1}$
1.17	$7.1 \pm 5.6 \cdot 10^{-1}$	4.17	$7.1 \pm 2.3 \cdot 10^{-1}$
1.33	$7.1 \pm 2.2 \cdot 10^{-1}$	4.33	$7.7 \pm 8.3 \cdot 10^{-3}$
1.50	$7.0 \pm 1.9 \cdot 10^{-1}$	4.50	$7.6 \pm 1.1 \cdot 10^{-2}$
1.67	$7.1 \pm 1.1 \cdot 10^{-1}$	4.67	$7.5 \pm 1.3 \cdot 10^{-2}$
1.83	$7.3 \pm 1.9 \cdot 10^{-1}$	4.83	$7.4 \pm 1.5 \cdot 10^{-1}$
2.00	$7.0 \pm 3.0 \cdot 10^{-1}$	5.00	$7.4 \pm 2.3 \cdot 10^{-2}$
2.17	$6.7 \pm 3.1 \cdot 10^{-1}$	5.17	$7.8 \pm 1.4 \cdot 10^{-1}$
2.33	$7.4 \pm 1.0 \cdot 10^{-1}$	5.33	$7.1 \pm 2.6 \cdot 10^{-1}$
2.50	$7.3 \pm 1.5 \cdot 10^{-1}$	5.50	$6.5 \pm 5.2 \cdot 10^{-2}$
2.67	$7.1 \pm 2.0 \cdot 10^{-2}$	5.67	$7.1 \pm 2.2 \cdot 10^{-1}$
2.83	$6.8 \pm 5.2 \cdot 10^{-2}$	5.83	$7.6 \pm 3.2 \cdot 10^{-1}$
3.00	$7.4 \pm 1.6 \cdot 10^{-2}$	6.00	$7.5 \pm 1.7 \cdot 10^{-2}$

APPENDIX M

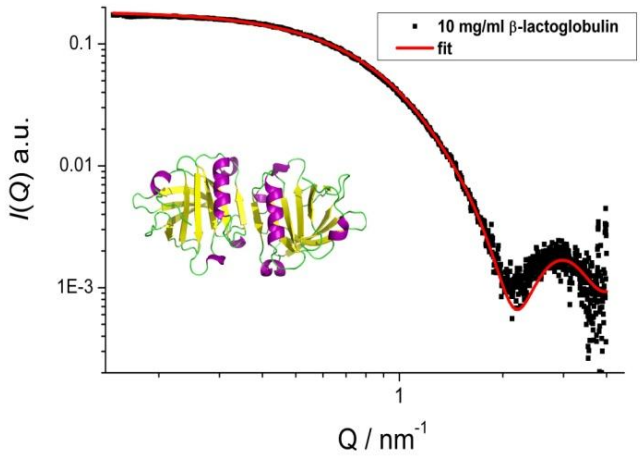
Synchrotron small angle X-ray scattering (SAXS) measurements were carried out at beamline SWING at SOLEIL synchrotron source in Gif-Sur-Yvette (France) by Dr. Nadeem Javid (University of Strathclyde) and Dr. Karsten Vogtt (Institute for Soft Matter and Functional Materials, Helmholtz Centre, Berlin). These measurements were done as collaboration for this project. Results are published in Soft Matter (Vogtt et al., 2011).

The software provided at the beamline was used for data reduction. Different protein concentrations of β-Lg were used in 20 mM phosphate buffer only and with 50 % (v/v) ethanol in the presence of 10 mM phosphate buffer at pH 7. The 10 mg/ml and 40 mg/ml (1 and 4 wt %) concentrations samples with and without ethanol in buffer were measured for freshly prepared and 7 days old samples. The background subtracted intensities of 10 mg/ml concentrations, with and without ethanol, were analyzed by Guinier's analysis to estimate the sizes (Balakrishnan, Javid, Weingaertner and Winter, 2008). The pair distribution functions (*P(r)*) were calculated by using the programme GNOM (Semenyuk and Svergun, 1991). The ab initio low-resolution X-ray solution structures were determined by using the programme GASBOR (Svergun, Petoukhov and Koch, 2001) which reconstructs the protein structure by chain like ensemble of dummy residues. The programme DAMAVER (Volkov and Svergun, 2003) was used to average the 10 calculated structures to build the protein structure.

The static light scattering measurements were carried by using 3 DDLS spectrophotometer (LS instruments, Fribourg, Switzerland) using vertically polarized He-Ne laser light (25 mW with wavelength of 632.8 nm) with an avalanche photodiode detector at various angles. The background intensities were subtracted from the intensities of the protein solutions in buffer with and without ethanol.

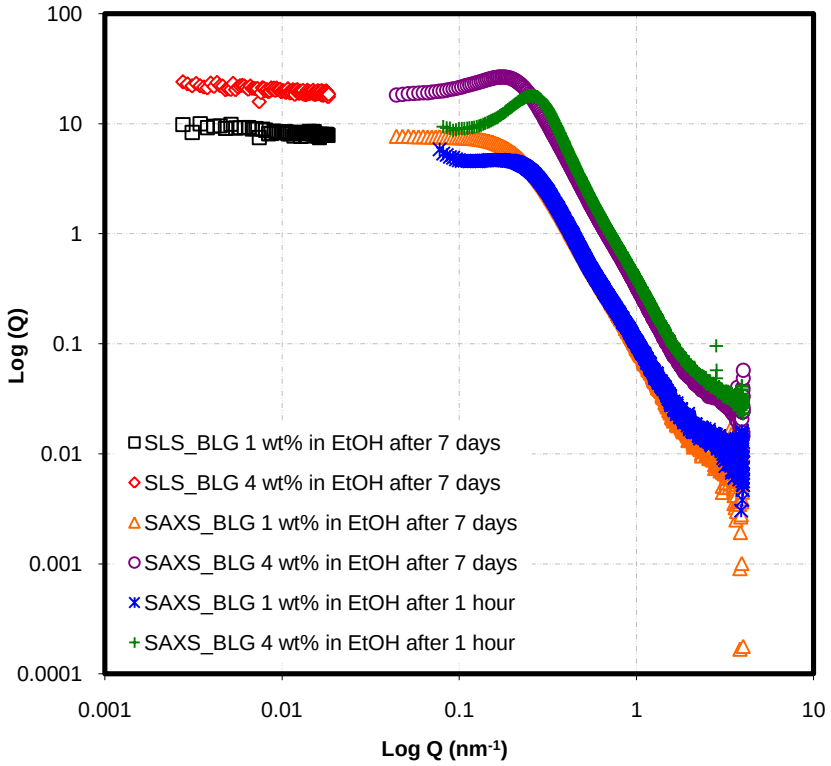
Firstly, β-Lg in its native state was measured and SAXS results were fitted based on the pdb structure 1BEB. This model represents well the β-Lg dimer since the crystal structure of the protein and its structure in solution do not differ significantly from each other. This model also suggests that the β-Lg dimer can be represented as two attached spheres. This result is shown in Graph M.1.

Graph M.1 - SAXS intensities of a β-Lg solution (10 mg/ml in 20 mM buffer at ambient conditions) and fit performed with the program CRYSOL based on the pdb-structure 1BEB. The good agreement of the data with the fit shows that the solution structure of dimeric β-Lg is well represented by its crystal structure (Vogtt et al., 2011).



The background subtracted intensities of various β-Lg concentrations (freshly prepared and 7 days old) in 50 % (v/v) ethanol measured by SAXS and SLS are shown in Graph M.2.

Graph M.2 - Small angle X-ray Scattering (SAXS) and static light scattering (SLS) intensity patterns for various protein concentrations with 50 % (v/v) ethanol and 10 mM phosphate buffer at pH 7.0. SLS: squares, β-Lg 1 wt % (10 mg/ml) with ethanol after 7 days; diamonds, β-Lg 4 wt % (40 mg/ml) with ethanol after 7 days. SAXS: triangles, β-Lg 1 wt % (10 mg/ml) with ethanol after 7 days; circles, β-Lg 4 wt % (10 mg/ml) with ethanol after 7 days; stars, β-Lg 1 wt % (10 mg/ml) with ethanol after 1 hour; crosses, β-Lg 4 wt % (40 mg/ml) with ethanol after 1 hour.



The increasing trend of scattering intensity at low *Q*-value for freshly prepared samples indicates that the protein molecules aggregates into large clusters for both of the concentrations. The correlation peak in the case of the 4 wt % concentration suggests that the protein aggregates are subjected to structural ordering due to the balance of intermolecular attractive and repulsive interactions. The protein aggregates have characteristic inter-particle distance of 22.6 nm calculated from the peak position in the structure factor (Javid et al., 2007). The aggregates undergo

different structural changes with time, and form stable size aggregates indicated by the plateau of the scattering intensities at low *Q*-values (7 days old samples). The plateaus of small angle light scattering (SAXS) patterns also confirm the existence of equilibrium clusters and no large scale structures (network/gel) are found. The radius of gyration for the cluster in case of 1 wt % (10 mg/ml) β-Lg in the presence of 50% (v/v) ethanol and 10 mM buffer after 7 days was found to be 6 nm by Guinier's analysis corresponding to geometric radius of 7.8 nm and diameter of value 15.6 nm by assuming the spherical shape of the oligomer. The DLS method yielded hydrodynamic radius of 8.1 ± 0.3 nm (Table 3.7) which is in accordance with the SAXS data. The β-Lg solution in buffer (20 mM) yielded the radius of gyration of the protein molecule as 2.1 nm with geometric radius of 2.7 nm (DLS showed hydrodynamic radius of 3.3 ± 0.2 nm suggesting the dimeric β-Lg structure (Table 3.3)). The oligomers are more likely to be hexameric in nature and they arrange themselves having inter-particle distance of 27.8 nm for 4 wt % (40 mg/ml) of protein concentration. The pair distribution functions and the corresponding ab initio molecular envelop shapes reconstructed from GASBOR are shown in Graph M.3. The maxima of the peaks in pair distribution functions indicate the approximate radius of the particles, which match well with the results from Guinier's analysis.

Graph M.3 - Pair distribution functions for dimeric (squares) and hexameric (circles) oligomeric structures along with the ab initio molecular envelop shapes reconstructed from GASBOR.