Backbone Dynamics of Bovine β-Lactoglobulin by ^{15}N NMR Spectroscopy

A thesis presented in partial fulfillment of the requirements for the degree of

Master of Science in Biochemistry

Institute of Fundamental Sciences and

Institute of Molecular Biosciences

Massey University

New Zealand.

Kristy Baker

2011

In Memory of Clinton John Reeve

Abstract

Bovine β-lactoglobulin (β-Lg) is a small 162 residue protein of unknown function from the whey component of milk, constituting ~50 % by dry mass. The protein is of great interest to the dairy industry due, in part, to its role in the fouling of dairy plants during heat treatment, and the significant operational costs this incurs. The structure of this protein is an eight stranded β-barrel with one long and two short flanking α helices. It is dimeric at neutral pH but dissociates at pH < 3.

In New Zealand herds there are three genetic variants, with variants A and B of bovine β-Lg predominating, while the C variant occurs at low levels in Jersey cows. However, despite the structural similarities of the three variants, milks containing one of A, B or C behaves differently when subjected to thermal processing. A greater understanding of factors that differentiate these protein variants is therefore important. In this study, ^{15}N nuclear magnetic (NMR) spectroscopy methods have been used to study the backbone dynamics of β-Lg A and B, at one temperature, and the hitherto unstudied C variant, at three temperatures. For follow-up functional studies a mutant protein, a covalently linked Ala34Cys dimer, was produced.

Acknowledgements

My first thanks go to my supervisor Dr. Patrick Edwards for providing me with a wealth of knowledge into the understanding of NMR spectroscopy based protein dynamics, assisting me with my figures, running the pulse sequences and taking the time to help me understand Linux and all the NMR programs. Many thanks to my other supervisors Professor Geoffrey Jameson for your encouragement, proof-reading my thesis, helping me keep things rolling in the critical end stages and for our interesting discussions and to Dr. Gill Norris for the resources and feedback into the molecular biology aspects of this project. I would like to express my gratitude and appreciation to Dr. Greg Sawyer, Trevor Loo and Dr. Alexander Goroncy for your sound advice, helpful insight and good cheer Also, thanks to the NMR group, former members and current; Hari, Jo, Martin, David and Nishit. I would especially like to thank my friends Jan, Ava and Carlene for keeping things fun and interesting and I would like to also thank other members of the structural biology group. I am grateful to my parents Grant and Lynette Baker, Craig's parents Malcolm and Fiona Lunn, my sister Rachel and brother Ted, for their love, support and frequent hot meals. Thanks again to Rachel for proof-reading my thesis. Also, I express my gratitude to Gribbles Veterinary Pathology for providing me with the means to finish and equipping me with new skills. Particular thanks to Dr. David Tisdall, Dr. Phil McKenna, Dr. Fraser Hill and Dr. Janice Thompson for encouraging me and taking an interest in my research. Thanks to Gaylene for the laughs and working with me. Thanks to Clint for being a good friend, wanting to know more about what I do and willing me to succeed. Sorry you can't be here to celebrate with me at the finish line. Thanks to Mark for all the coffees, listening to me grumble and not trying to tell me what to do to fix it. I couldn't have made it this far without your friendship. Thanks to the former Foundation for Research, Science and Technology (FRST) for providing funding, which enabled these studies, and to the Riddet Institute for providing me a scholarship. And finally I would like to thank Craig for constantly believing in me, making me laugh and sharing the bigger picture.

Glossary of Abbreviations

Å	Ångstrom (10^{-10} m)
Aa	Amino acid
AEC	Anion exchange chromatography
α-La	α-Lactalbumin
Amp	Ampicillin
Bis-tris	1,3-Bis(tris(hydroxymethyl)methylamino)propane
β-Lg	β-Lactoglobulin
BME	β-Mercaptoethanol
Bp	Base-pair
C	Carbon
°C	Degrees Celsius
CPMG	Carr-Purcell-Meiboom-Gill
Da	Dalton
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DsbC	Disulfide bond isomerase C
EDTA	Ethylene diamine tetra-acetic acid
EtBr	Ethidium bromide
EtOH	Ethanol
FID	Free induction decay
g	Gram
× *g*	Multiples of gravitational force
GER	Germany
H	Hydrogen
*Hin*dIII	DNA restriction endonuclease sourced from *Haemophilus influenza*
HMH	6-Hydroxy-6-methyl-3-heptanone
HSQC	Hetero-nuclear single quantum correlation
I	Italy
IEC	Ion exchange chromatography
IPTG	Isopropyl-β-D- thiogalactopyranoside
K	Kelvin
Kan	Kanamycin
Kb	Kilo bases
kDa	Kilo-Dalton
*Kpn*I	DNA restriction endonuclease sourced from *Klebsiella pneumonia*
LB	Luria Bertani media

m	Metre
mAU	Milli absorbance units
MCS	Multiple cloning site
MCS1	Multiple cloning site one
MCS2	Multiple cloning site two
µg	Micro gram
MHz	Mega hertz
mL	Milli litre
µL	Micro litre
mM	Milli molar (mmol L^{-1})
mol	Mole
ms	Millisecond
N	Nitrogen
*Nco*I	DNA restriction endonuclease sourced from *Gordonia rubripertincta*
*Nde*I	DNA restriction endonuclease sourced from *Neisseria denitrificans*
ng	Nanograms
nm	Nanometers
NMR	Nuclear Magnetic Resonance spectroscopy
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect Spectroscopy
ns	Nanoseconds
NZ	New Zealand
1D	One-dimensional
OD_{600}	Optical density (at a wavelength of 600 nanometres)
Pa	Pascal (= 10^{-5} bar, 145.05 $\times^{-6}$ psi)
PCR	Polymerase chain reaction
pH	Negative decadal logarithm of proton concentration
pKa	Acid dissociation constant, as negative decadal logarithm
ppm	Parts per million
ps	Picoseconds
R_1	Longitudinal (or spin-lattice) relaxation rate
R_2	Transverse (or spin-spin) relaxation rate
RBP	Retinol binding protein
RBS	Ribosome binding site
RCI	Random coil index
R_{ex}	Exchange induced relaxation rate
S^2	Squared order parameter
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
ss-NOE	Steady state-nuclear Overhauser effect
TAE	Tris-acetate-EDTA buffer
τ_e	Effective correlation time
Temp	Temperature
Tet	Tetracycline

τ_m	Molecular correlation time
TOCSY	Total correlation spectroscopy
2D	Two-dimensional
3D	Three-dimensional
USA	United States of America
UV	Ultraviolet light
V	Volts
v/v	Volume per volume
w/v	Weight per volume

Abbreviations of Nucleic Acids

One Letter Code	Base Represented
A	Adenine
T	Thymine
C	Cytosine
G	Guanine
U	Uracil

Abbreviations of Amino Acids

Amino Acid	3-Letter Code	1-letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic Acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Contents

List of Figures

LIST OF FIGURES

List of Tables

1

Introduction

1.1 Introduction

This chapter provides an introduction to the molecular biology of β-lactoglobulin, in particular, its structure and dynamics. The history of overcoming solubility issues upon recombinant expression, which is necessary for NMR spectroscopy investigations, is reviewed and a brief insight into NMR spectroscopy based protein dynamics is provided. The effects of heat treatment on purified variants A, B and C, and some of the factors that differentiate these structurally similar variants are discussed. Finally, the aims of this project are listed.

1.2 Milk

Since about 8000 BC, milk and dairy products have been an important nutritional component of the human diet in numerous regions around the world. Apart from human milk, which is of importance to its own neonate, other milks have been sourced by people from bovine, caprine and to a smaller degree ovine species. Milk is a complex mixture comprising proteins, lipids, carbohydrates, minerals, vitamins, colloidly dispersed salts, and water. The compositions and ratios of these milk constituents vary from species to species and also have been found to differ between breeds (Jenness, 1988). Some of the components that make up milk are synthesised in the mammary gland of the female, whereas others are transferred through the blood stream.

Milks analysed to date contain two protein groups, which have been distinguished as being acid precipitable or acid soluble at pH 4.6 and 20 °C. The former have been commonly grouped as the caseins and the latter as the whey proteins (Farrell *et al.*, 2004). The caseins along with the major whey proteins, β-lactoglobulin (β-Lg) and α-lactalbumin (α-La), are only expressed during lactation by specific secretory cells of the mammary gland and are targeted to the lumen to accumulate in milk (Larson, 1979).

The major whey proteins in milk include β-Lg, α-La, bovine serum albumin, lactoferrin and immunoglobulins (Farrell *et al.*, 2004). These proteins have been classified into their individual families based on homology with their primary amino acid sequence. Even though the biological function of β-Lg has been widely debated, other whey proteins have been identified to have immunological, bacteriostatic, enzymatic and/or other functional properties (Table 1.1).

Protein	**Composition in skim milk** (g/L)	**No. amino acids**	**Molecular weight** (kDa)	**Isoelectric point**	**Comments**
β-Lactoglobulin A	2-4	162	18.363	5.13	Unknown function
α-Lactalbumin	0.6-1.7	123	14.178	4.2-4.5	Subunit of lactose synthase
Bovine serum albumin	0.4	607	66.399	4.7-4.9	Non-specific carrier of hydrophobic molecules
Immunoglobulin G	0.3-0.6	>500	161 000*	5.5-6.8	Immunity function
Lactoferrin	0.02-0.1	689	76.110	8.81	Bacteriostatic role

Table 1.1 Whey proteins of milk and some of their properties (Farrell *et al.*, 2004, Edwards *et al.*, 2008).

* Molecular mass for G1, the major immunoglobulin. IgG2, IgM and IgA are present in much lower concentrations.

1.3 Bovine β-Lactoglobulin

Bovine β-Lg is the most studied milk protein as a consequence of its ready availability and its commercial significance. The β-Lg monomer is comprised of 162 amino acid residues and has a molecular weight of 18.3 kDa (Hambling *et al.*, 1992). It is a member of the lipocalin family of proteins, which are best known for the ability to bind small hydrophobic molecules within an internal hydrophobic cavity (Flower, 1996, Flower *et al.*, 2000). Intriguingly β-Lg's biological function is not known although there has been much speculation because of β-Lg's ability to bind many hydrophobic ligands, including palmitate (Wu *et al.*, 1999), retinoid species (Papiz *et al.*, 1986) and cholesterol (Kontopidis *et al.*, 2004) within its cavity, suggesting that it could be implicated in the transport of small molecules. Nonetheless, it may simply play a binding role.

β-Lg is expressed in the glandular epithelium of the mammary gland of most mammals including bovine, equine and porcine species, but to date has not been detected in the milk of humans, rodents or lagomorphs. It is the major whey protein in these milks, constituting about 10 % of total protein and approximately 50 % of whey protein, whereas in human milk, α-lactalbumin is the major whey protein component (Edwards *et al*., 2008).

To date, ten genetic variants of bovine β-Lg have been identified. In New Zealand bovine herds two β-Lg genetic variants, A and B, predominate (Farrell *et al*., 2004), while the C variant occurs at low levels in Jersey cows. These three variants differ by up to three amino acid substitutions and all share similar tertiary structure (Bewley *et al.*, 1997). β-Lg A and β-Lg B differ by both Asp64Gly and Val118Ala substitutions respectively, whereas β-Lg C differs from β-Lg B through a Gln59His substitution. The protein exists as a monomer, dimer or oligomer depending on the pH, temperature and ionic strength. It is dimeric above pH 3.5, including at neutral pH, but exists as a monomer at low pH and low salt concentration as a result of an intricate network of hydrophobic, electrostatic and hydrogen bond interactions (Sakurai *et al.*, 2001, Uhrínová *et al.*, 2000).

1.4 Molecular Structure of β-Lg

The molecular structure of bovine β-lactoglobulin has been determined by several groups using both X-ray crystallography (Brownlow *et al.*, 1997, Qin *et al.*, 1998a, Qin *et al.*, 1998b) and solution nuclear magnetic resonance (NMR) spectroscopy (Fogolari *et al.*, 1998, Kuwata *et al.*, 1999, Uhrínová *et al.*, 1998). These studies have demonstrated that bovine β-Lg shares the same general structure over a wide pH range, and in the presence and absence of ligands.

Bovine β-Lg is predominantly a β-sheet protein and hence has a considerable amount of β-sheet hydrogen bonding. The secondary structure consists of 15 % α-helix, 50 % β-sheet and 15–20 % reverse turn. The tertiary structure of β-Lg is composed of nine anti-parallel β-strands, of which eight contribute to a flattened calyx with a large internal hydrophobic pocket, a major three-turn α-helix, positioned on the outer surface of the calyx, and two minor flanking α-helices (Figure 1.1).

The flattened calyx, also known as the β-barrel, is conical and made up of two β-sheets. The β-B to β-D strands and the N-terminal half of the β-A strand form one sheet, and the β-E to β-H strands and the C-terminal half of the β-A strand form the other. The residue Leu22 is positioned at the midpoint of the β-A strand at the 90° bend. Its high content of rigid β-sheet structure means that β-Lg is stabilised by the considerable network of hydrogen bonds formed between strands, with the exception of strands β-B and β-D where all possible hydrogen bonds are not fulfilled (Figure 1.2).

The loops that connect the β-strands at the closed end of the barrel, B/C, D/E and F/G, are typically quite short, whereas those at the open end, A/B, C/D, E/F and G/H, are significantly longer and more flexible (Qin *et al*., 1998b, Uhrínová *et al*., 2000). A short 3_{10}-helix precedes β-A strand, and a second 3_{10}-helical turn lies in the long A/B loop, which forms part of the dimer interface together with strand β-I (Section 1.4.1). Short 3_{10}-helical turns are also found in the G/H loop and within the C-terminal region. The three-turn α-helix lies in the sequence between the β-H and β-I strands.

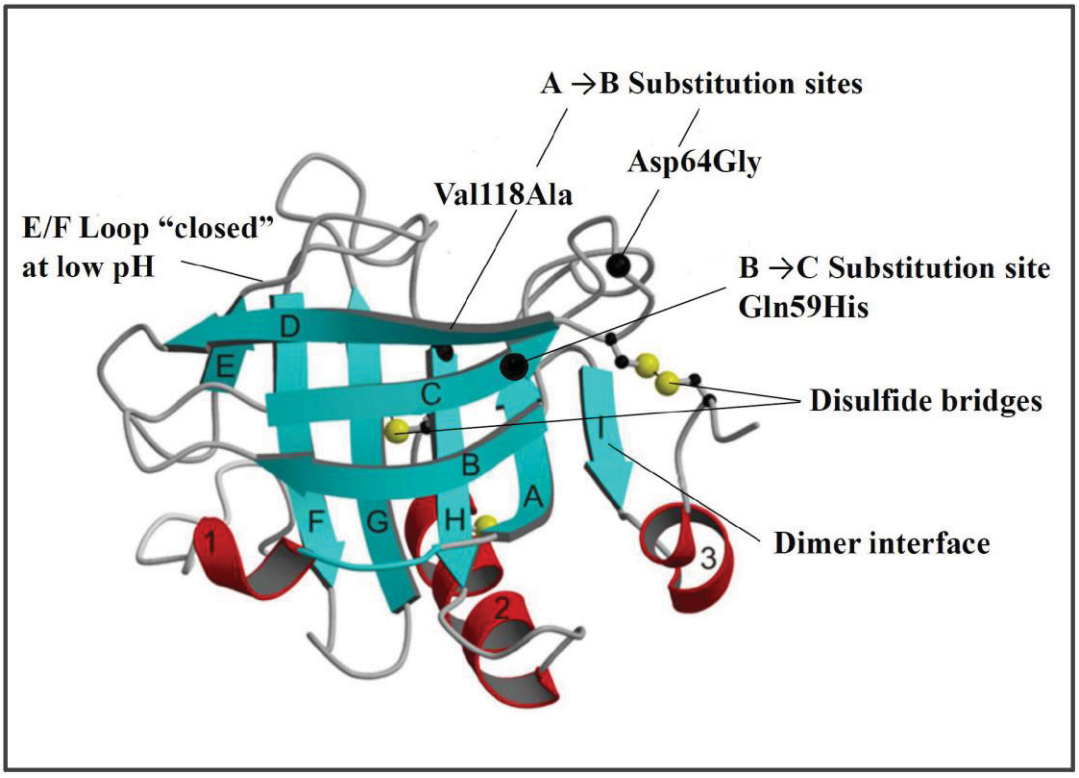


Figure 1.1 Structure of β-Lg.

Arrangement of β-sheets and helices of β-Lg A (from NMR data at pH 2.6) (Uhrínová *et al.*, 2000) generated with PyMOL (Delano, 2008). Teal represents the β-sheets (nine strands labelled A-I), whereas red represents the α-helices (three helices labelled 1-3). The positions of the cysteine residues are shown by the yellow circles and the variant A, B and C substitution sites with the larger black dots.

The large central cavity within the barrel is readily accessible to solvent at pH > 6. It serves as a principal binding site for a wide variety of hydrophobic molecules that is enabled by a stable cluster of 12 hydrophobic residues; Val15, Trp19, Tyr42, Leu46, Leu54, Phe82, Val92, Val94, Leu103, Phe105, Met107 and Leu122, which lend their side chains into the pocket. Trp19 (sited on the strand β-A just before the bend) and β-Lg's only other tryptophan residue, Trp61 (sited at the end of strand β-C), provide markers for investigating site-specific conformational changes. In the native structure Trp19 is hidden in the hydrophobic core, facing into the base of the cavity, whereas Trp61 is exposed to the solvent.

Of β-Lg's cysteine residues, four form two intra-molecular disulfide bridges (Figure 1.2 and Figure 1.3). One bridge links Cys66 (C/D loop) with Cys160 (close to the C-terminus), near the surface of the protein molecule, and the other links Cys106 (β-G strand) to Cys119 (β-H strand), in the interior of the molecule. The Cys106-Cys119 interaction forms a *cis*-disulfide bridge as opposed to a *trans*-disulfide bond as formed between the Cys66-Cys160 linkage. A free thiol (Cys121) is situated on strand β-H and

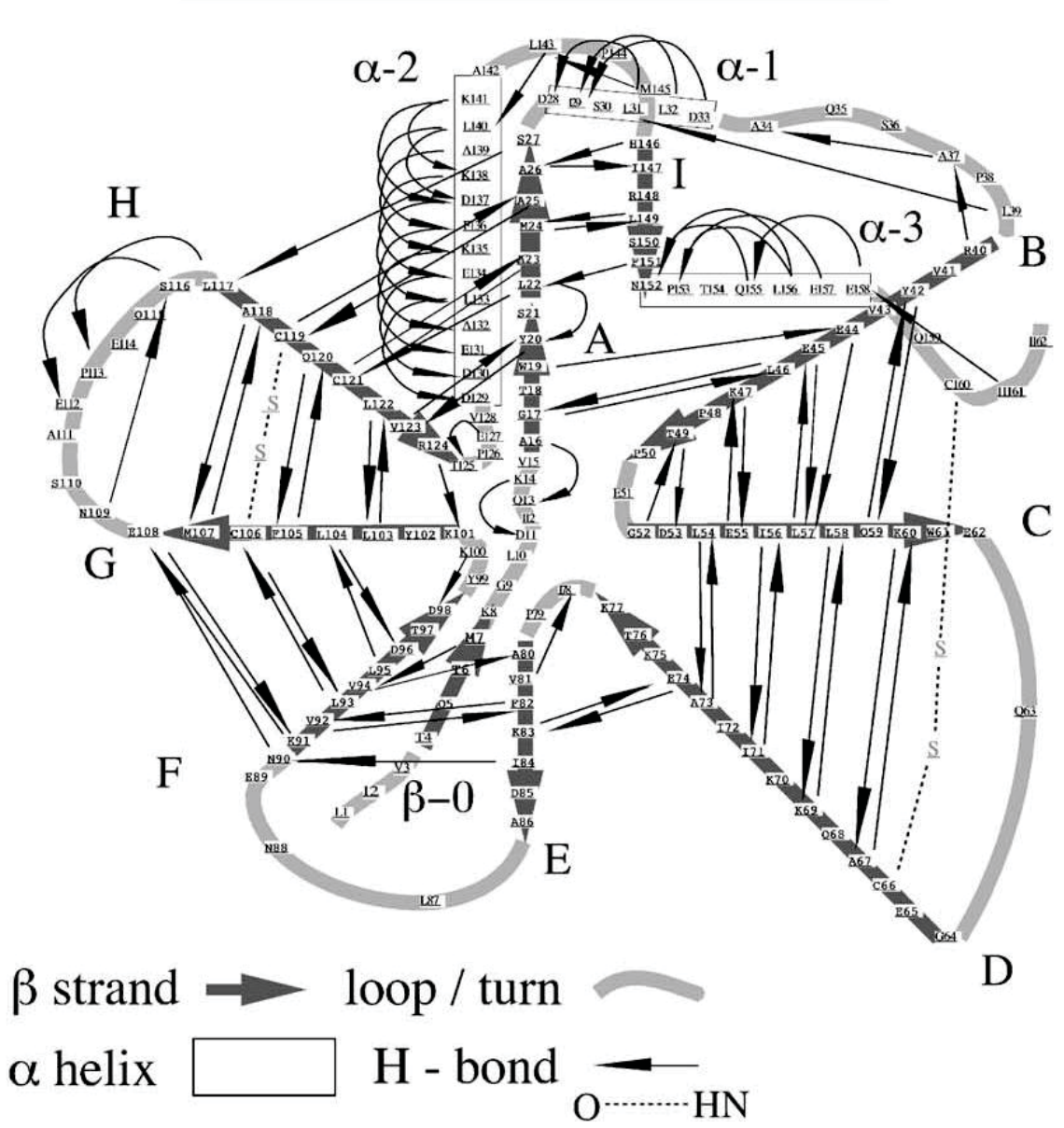


Figure 1.2 Topology of Bovine β-Lg.

Hydrogen-bonding pattern of β-Lg A (from X-ray data at pH 6.2 (Qin *et al*., 1998)). The nine β-strands are labelled alphabetically A–I and the α helical regions are numbered 1–3. The positions of the two disulfide bonds are shown by the dotted lines and the hydrogen bonds are from HN towards O. This figure was reproduced from Edwards *et al.* (2002).

is buried between the three-turn α-helix and the β-barrel.

1.4.1 Dimeric Interface

β-Lg exists predominantly as a homodimer at neutral pH, and at protein concentrations above 1 mg mL^{-1}, but dissociates primarily into monomeric species when the pH of the protein solution is lowered from 6.2 to 2.6, under conditions of low salt (Fugate & Song, 1980). Interestingly, much of the protein's native structure is retained (1.5.1), which is consistent with it being extremely acid stable at pHs as low as 2.4 (Sakurai & Goto, 2007).

Investigations into the monomer-dimer equilibrium (Uhrínová *et al.*, 2000, Sakurai & Goto, 2002, Joss & Ralston, 1996, Sakurai *et al.*, 2001) have provided information on the nature of the dimeric interface. At neutral pH, X-ray crystallography structural studies have shown that the dimer is stabilised by a variety of interactions formed between both the two anti-parallel β-I-strands and the two A/B-loops. Both hydrogen bonds and hydrophobic interactions play roles in stabilising the dimer at the β-I strands (Figure 1.3) and salt bridges formed by the positively charged Arg40 residues and the negatively charged Asp33 residues of the A/B loops stabilise the dimer as well as the subsequent hydrogen bonds formed between these loops. However, because of the small area of the interaction, the total energy stabilising the dimer is relatively small (Brownlow *et al*., 1997, Sakurai & Goto, 2002).

β-Lg's surface is positively charged at pH 3, but has negative and positive patches at its surface at pH 6-8 (Joss & Ralston, 1996, Qin *et al*., 1998a, Uhrínová *et al*., 2000). The favouring of monomeric species at low pH and low salt concentration is thought to stem from electrostatic repulsion between the protein monomers, which is modulated by the addition of neutral salt, shielding the positive charges of the protein and stabilising the dimer (Sakurai *et al*., 2001, Joss & Ralston, 1996). Uhrínová and coworkers speculated that a shift towards the monomeric population at pH 2.6 may also be the result of changes in the conformation of residues sited in the A/B loop, which is involved in forming part of the dimer interface, as the NMR structures at low pH showed there was up to 3.5 Å shift compared to the crystal structure (Uhrínová *et al*., 2000).

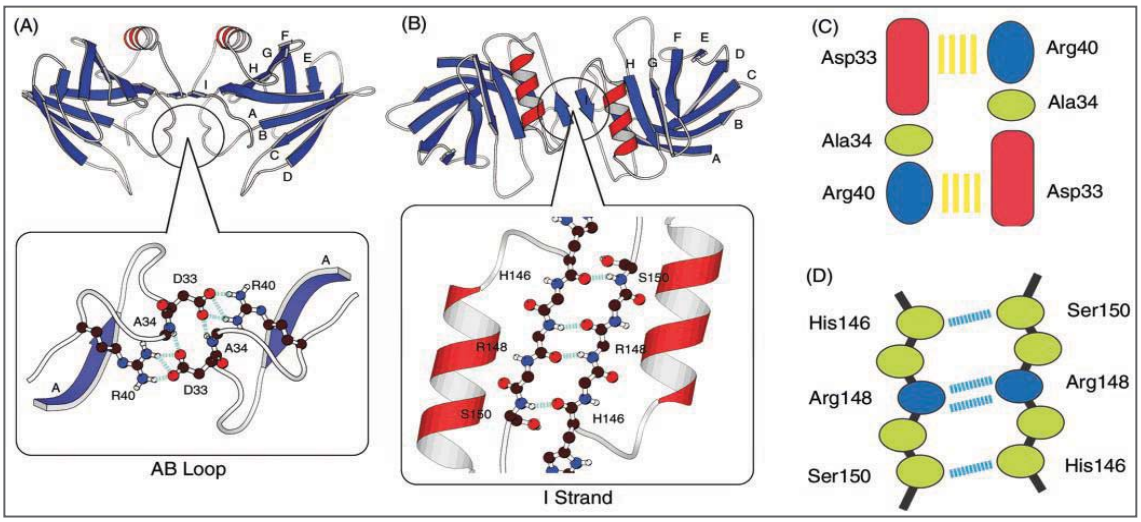


Figure 1.3 Structure of β-Lg's Dimeric Interface at pH 6.5, as Determined using X-Ray Crystallography.

Pictures created by Sakurai and Goto (2002). (A) *Side view* and (B) *top view* of the dimer interface with regards to β-I strands. *Close views* of (A) A/B loops and (B) β-I strands show intermolecular hydrogen bonds between side-chain and main-chain atoms. (C) Schematic illustration of the β-Lg dimer showing salt bridges between side chains of Arg40 and Asp33 at the A/B loops and (D) four hydrogen bonds formed between the β-I strand main chains. Sakurai *et al.* (2002) had created figures (A) and (B) with Molscript (Kraulis, 1991) and the Protein Data Bank code 1BEB (Brownlow *et al*., 1997).

1.4.2 The Tanford Transition

Although β-Lg exists in a native state over a wide pH range, it undergoes significant conformational transitions between pHs 6.3 and 8.2, which have been observed by a host of techniques including optical rotational dispersion (Tanford *et al.*, 1959), sedimentation titration (Taulier & Chalikian, 2001), X-ray crystallography (Qin *et al*., 1998a) and NMR spectroscopy (Sakurai & Goto, 2006, 2007). This series of pH-dependent conformational changes is known as the Tanford transition. It was first observed by a change in optical rotary dispersion at pH 7.0 (Tanford *et al.*, 1959) and is thought to be important as it could possibly be related to the function of the protein.

The Tanford transition involves the expansion in the volume of the protein molecule with the opening of a lid to the barrel (Loop E/F; residues 85-90), which is closed at below pH 6.2 and is opened at pH 7.1 and 8.2, as observed by X-ray structures of crystals at pH 6.2, 7.1 and 8.2 (Qin *et al*., 1998a). The opening of the lid is accompanied by a deprotonation of residue Glu89, which lies on the E/F loop and has an anomalous pKa of 7.5 (Tanford *et al*., 1959, Qin *et al*., 1998a). This residue is hidden in the closed form and is exposed in the open form. These conformational

changes have been thought to regulate the binding of ligands to β-Lg, as the lid could be involved in controlling entry into the central cavity.

A range of heteronuclear NMR spectroscopy studies tracked the conformational changes taking place throughout the Tanford transition and as a consequence a three-step mechanism was proposed shown, which is in Figure 1.4 (Sakurai & Goto, 2006, Sakurai & Goto, 2007, Sakurai *et al.*, 2009): 1) initially the carboxyl group of Glu89 is deprotonated, 2) next there is a fluctuation of hydrogen bonds among the backbone atoms of three residues: Ile84, Asn90 and Glu108, where Ile84 and Asn90 are positioned in the hinge region of the E/F loop and Glu108 hydrogen bonds to the backbone of this loop, and 3) unfolding of the β-D strand, E/F loop and G/H loop takes place.

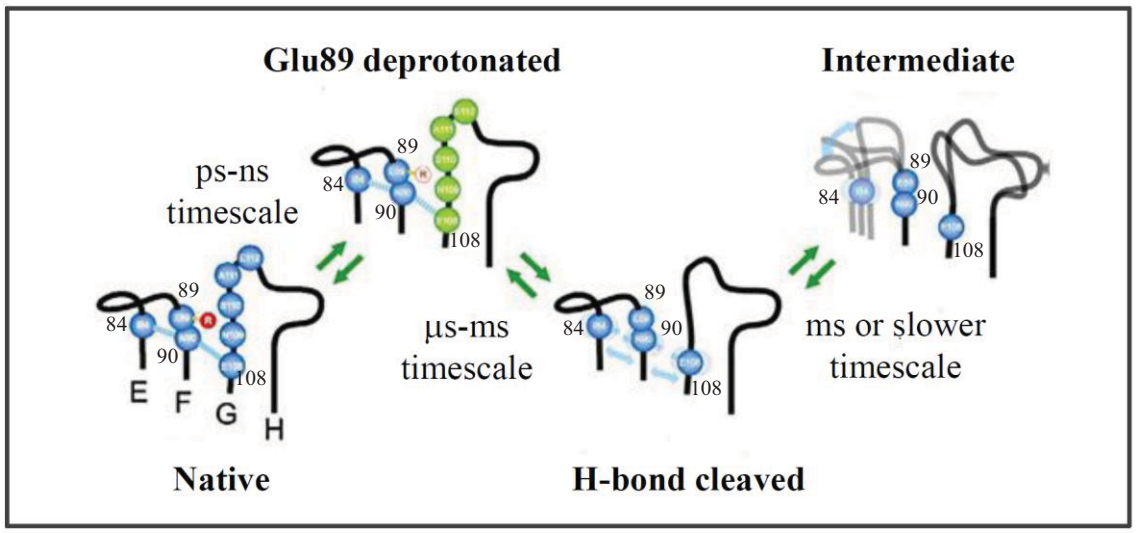


Figure 1.4 Molecular Process of the Closed-Open Tanford Transition of Bovine β-Lg.
In the first step Glu89 undergoes deprotonation of its carboxyl group. Second: the fluctuation or cleavage of hydrogen bonds between backbone atoms at Ile84, Asn90, and Glu108. Third: unfolding at the β-D strand and the E/F and G/H loops takes place (Sakurai *et al.*, 2009).

1.5 Solution Structures of Bovine β-Lg

Protein structures in solution, determined with high-field nuclear magnetic resonance spectroscopy (NMR), typically require monomeric proteins with molecular weights less than 40 kDa. To solve the structure of a protein molecule using NMR spectroscopy methods an isotopically labelled recombinant sample is typically required if the protein molecule is greater than 8 kDa. This is unlike X-ray crystallography methods where the isotopic label is not necessary and the protein can be obtained from native sources.

Until recently, the majority of β-Lg structural studies using multi-dimensional NMR techniques had been conducted at a pH between 2 and 3, in low-salt buffer (Fogolari *et al*., 1998, Uhrínová *et al*., 1998, Kuwata *et al*., 1999, Uhrínová *et al*., 2000). In this environment the protein is in its monomeric state. At neutral pH problems arise with NMR spectroscopy techniques, as the large size of the bovine β-Lg dimer causes the molecule to tumble slowly in solution, contributing to broadened peaks in the ^{1}H NMR spectrum (Figure 1.5). This peak broadening effect is further exacerbated by the dynamics of the monomer-dimer equilibrium. Therefore, the relatively small size of the β-Lg monomer at pH ~2 leads to better resolved spectra, making the protein more amenable to multidimensional NMR spectrometry, which is able to provide a high level of structural and dynamical information.

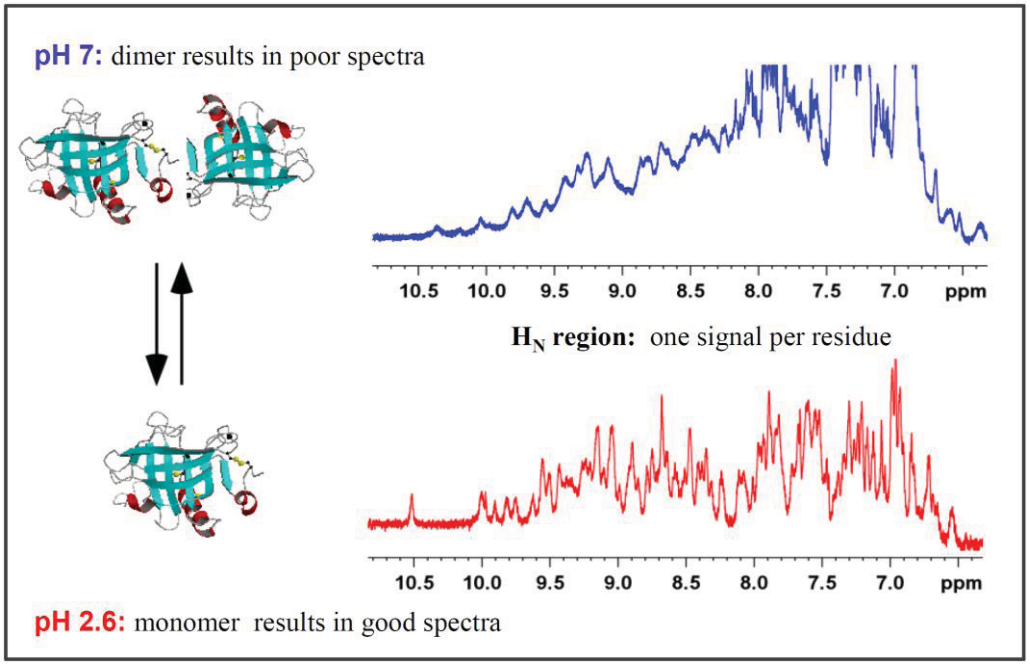


Figure 1.5 One-Dimensional NMR Spectra of β-Lg Sampled at Neutral pH and Low pH.
β-Lg exists predominantly as a dimer at neutral pH causative to broadened peaks in the spectrum. At pH 2.6 the β-Lg dimer has dissociated into monomeric species, contributing to well-resolved peaks that are amenable to NMR spectroscopy analysis.

1.5.1 Solution Structures of β-Lg at Low pH

Preliminary studies at pH 2.6, using homonuclear ^{1}H NMR spectroscopy experiments, showed that much of the secondary structure observed in the bovine β-Lg dimer at physiological pH was still present in the monomer (Ragona *et al.*, 1997). Also, to a great extent, the hydrophobic core appeared intact. However, UV-circular dichroism

(CD) measurements inferred there was some partial unfolding of the protein at low pH (Molinari *et al.*, 1996).

Three NMR spectroscopy structural studies have since been published independently of one another concerning β-Lg's structures at pH 2.0 to 2.6, supplying more information on its structure under these conditions. Fogolari *et al*. (1998) released the first comprehensive NMR spectroscopy structural characterisation using wild-type protein and ^{1}H NMR spectra, to produce a roughly resolved structure depicting the β-barrel, the C/D loop and the position of the main α-helix relative to the β-barrel. Following this, two concurrent studies, using recombinant ^{13}C- and ^{15}N-labelled β-Lg expressed in the methyltropic yeast *Pichia pastoris*, led to complete assignment of the protein's backbone and side chains atoms (Uhrínová *et al*., 1998, Kuwata *et al*., 1999), resulting in two independent, highly-resolved structures. Collectively, these studies have shown that the three-dimensional structure of the low-pH form of β-lactoglobulin is very similar to that of a subunit within the dimer at pH 6.2 (Qin *et al*., 1998a). However, some local differences arise with orientation, with respect to the β-barrel, the A/B loop, which is involved at the dimer interface (discussed in Section 1.4.1), the C/D loop, and the major three-turn α-helix, where its C-terminal end is tilted slightly more away from the N-terminal end of the β-H strand (Uhrínová *et al*., 2000). The hydrophobic cavity within the barrel is retained at low pH, with the stable cluster of 12 hydrophobic side chains still protruding into the cavity. The E/F loop, which moves to block the opening of the cavity over the pH range 7.2 - 6.2 during the Tanford transition, is in the 'closed' position at pH 2.6, and the side chain of Glu89 is buried, as expected (Uhrínová *et al*., 2000, Jameson *et al.*, 2002). At low pH, the protein becomes fully protonated, with a net charge of +21 proton charges (implications discussed in Section 1.4.1). These studies suggest that while the structure of β-Lg at low pH is not fully folded, it is very similar to the X-ray crystallography resolved structure of the dimer at pH 6.2. However, comparisons have not been made that note the discrepancies in structure at low pH, arising solely from the NMR spectroscopy and X-ray crystallography techniques alone, as an X-ray crystallography-derived structure has not been resolved for β-Lg at low pH.

1.5.2 Solution Studies of β-Lg at Neutral pH using an Ala34Cys Mutant Dimer

More recently, NMR spectroscopy measurements for the dynamics of the covalently linked Ala34Cys mutant dimer have given complementary information regarding the

structural changes associated with the Tanford transition (Sakurai & Goto, 2006, Sakurai *et al.*, 2009), established previously using X-ray crystallography (Qin *et al*, 1998a). The mutation was chosen so that Cys34 in the A/B loop, part of the dimer interface, would form a disulfide bridge with the corresponding Cys residue in the other monomer, producing a covalently linked dimer. The published (^{15}N,^{1}H) heteronuclear single quantum coherence (HSQC) spectrum for the mutant showed that the peaks were all of similar peak width; therefore it was suggested by the authors that peak broadening was due to the monomer-dimer equilibrium at neutral pH. When the spectrum was compared with that of wild-type β-Lg, under the same conditions, chemical shift differences for ^{15}N and ^{1}H, between the two sets, were less than 0.1 ppm, except for Cys34 and those near the mutation site. This illustrated that there was no great significant structural digression in the mutant when compared to native protein under neutral conditions, except for the site close to the mutation. UV-CD spectroscopy measurements verified that both proteins had similar tertiary and secondary structure, and tryptophan fluorescence monitoring Trp19 and Trp61 residues (Section 1.4) showed that the environment around the tryptophan residues in the mutant and the native β-Lg structures were similar at the physiological pH of cows' milk.

1.6 Expressing Isotopically Labelled β-Lg for NMR Spectroscopy

Full structural and dynamical characterisations of bovine β-Lg have been hampered by the length of time it has taken to develop a system that expresses and purifies adequate yields of correctly folded isotopically labelled protein. Examination of the structure-function relationships of a protein, using multidimensional NMR spectroscopy, requires an efficient expression system, which can incorporate the necessary isotope(s). When the protein is labelled with ^{13}C and/or ^{15}N labels, NMR spectroscopy experiments can be conducted that transfer magnetisation between specific groups of nuclei. This facilitates the assignment of NMR peaks to the appropriate amino acid residue of the ^{15}N,^{1}H-HSQC spectrum (Section 2.4.2), as well as gaining site-specific dynamical insights using specific NMR spectroscopy pulse sequences. Relaxation studies of β-Lg alone necessitate an 18 mg mL^{-1} (~1 mM) protein solution and, depending on the tube used, volumes required vary between 250 μL and 500 μL. If the expression system is not capable of producing adequate yields of correctly folded soluble protein, studies are extremely costly especially when using ^{13}C-D-glucose as the carbon source.

1.6.1 Recombinant Expression in Yeast

Up until recently the highest yields of recombinant β-Lg had been obtained from the lower eukaryote *Pichia pastoris* (Kim *et al.*, 1997, Denton *et al.*, 1998). The two NMR spectroscopy resolved structures of β-Lg A had been resolved with protein expressed in this system. However, the over-expression of labelled protein with *P. pastoris* has its disadvantages as it produces a recombinant protein with an additional mix of N-terminal extraneous residues in the purified protein. The protein also co-purifies with significant quantities of carbohydrate, which can otherwise interfere with quantitative NMR spectroscopy analyses and may cause β-Lg to behave in an unpredictable manner. Expression prior to this in both *Saccharomyces cerevisiae* (Totsuka *et al.*, 1990) and *Kluyveromyces lactis* eukaryotic yeast systems (Rocha *et al.*, 1996) produced insufficient amounts of recombinant protein for biophysical investigations.

1.6.2 Heterologous Expression in *Escherichia coli*

Expression of recombinant protein in a prokaryotic host has many benefits including rapid cell growth, easy culturing techniques and the availability of a myriad of established commercial kits and plasmids, which assist in protein expression. Ariyaratne and co-workers (2002) created the first expression system with an *Escherichia coli* host that produced adequate amounts of β-Lg necessary for 3D structural studies (15 mg of purified β-Lg per litre of culture). Expression prior to this in *E. coli* resulted in the formation of misfolded protein manifesting as inclusion bodies (Batt *et al.*, 1990). Ariyaratne's group engineered a synthetic leaderless BLG A gene, by removing the secretory signal sequence, to enable the targeting of recombinant protein to accumulate in the cytoplasm, rather than the periplasm. Restriction sites were also introduced into the gene sequence to create 'cassettes' for future site-directed mutagenesis investigations, and translation codons were changed based on high usage in *E. coli* (Nakamura *et al.*, 1995), decreasing the incidence of protein misfolding (Dillon & Rosen, 1990). It is important to note that these nucleotide base changes did not alter the translated primary sequence of β-Lg. Once the protein was purified, Edwards and coworkers (2004) were able to get preliminary insights into the dynamical nature of the A and B variants using NMR spectroscopy methods.

Although this expression procedure provided hope for further dynamical studies, these experiments were not reproducible, especially when producing variant B, as all new stocks of enterokinase purchased not only cleaved the fusion tag that aided purification,

but also β-Lg into many fragments (Professor Geoffrey B. Jameson; personal communication). However, after a period of frustration, this same group tackled issues of expression in *E. coli* by developing a new strategy to produce high-yields of 'native-like' β-Lg (Ponniah *et al.*, 2010). As studies had shown that correct disulfide bond formation was fundamental in attaining the wild-type conformation (Jayat *et al.*, 2004, Hattori *et al.*, 2005, Invernizzi *et al.*, 2008), solubility obstacles were removed by co-expressing the signal sequence deficient-disulfide bond isomerase C (DsbC) with leaderless β-Lg (Ariyaratne *et al*., 2002) in the cytoplasm of the double-mutant *E. coli* Origami (DE3) strain. The *E. coli* mutant was chosen as both the thioredoxin (*trxB*) and glutathione oxido-reductase (*gor*) genes had been inactivated (Bessette *et al.*, 1999) to ensure that the cytoplasmic region was oxidising, promoting disulfide-bond formation, as in a 'typical' *E. coli* host this area does not provide favourable redox conditions for these events to occur. In conjunction with DsbC isomerase expression, reproducible amounts of correctly folded β-Lg A and B variants, adequate for NMR spectroscopy analysis, were obtained producing 8.3 mg of protein per litre of culture post-purification. Studies prior to this showed that in oxidising environments, such as the Origami cytoplasm or the periplasmic space, the proof-reading protein DsbC is able to effectively catalyse disulfide bond formation and reshuffle non-native disulfides to their wild-type arrangement in proteins with multiple disulfide bonds (Bessette *et al*., 1999, Levy *et al.*, 2001, Zapun *et al.*, 1995).

Proteins were co-expressed with the aid of the pETDuet-1 dual-expression plasmid from Novagen. The two target genes encoding leaderless DsbC isomerase and bovine β-Lg variant A/B proteins were ligated into MCS1 and MCS2, respectively (Ponniah *et al*., 2010). The construct incorporated T7 bacteriophage transcription and translation signals that provide high-level protein expression (Figure 1.6).

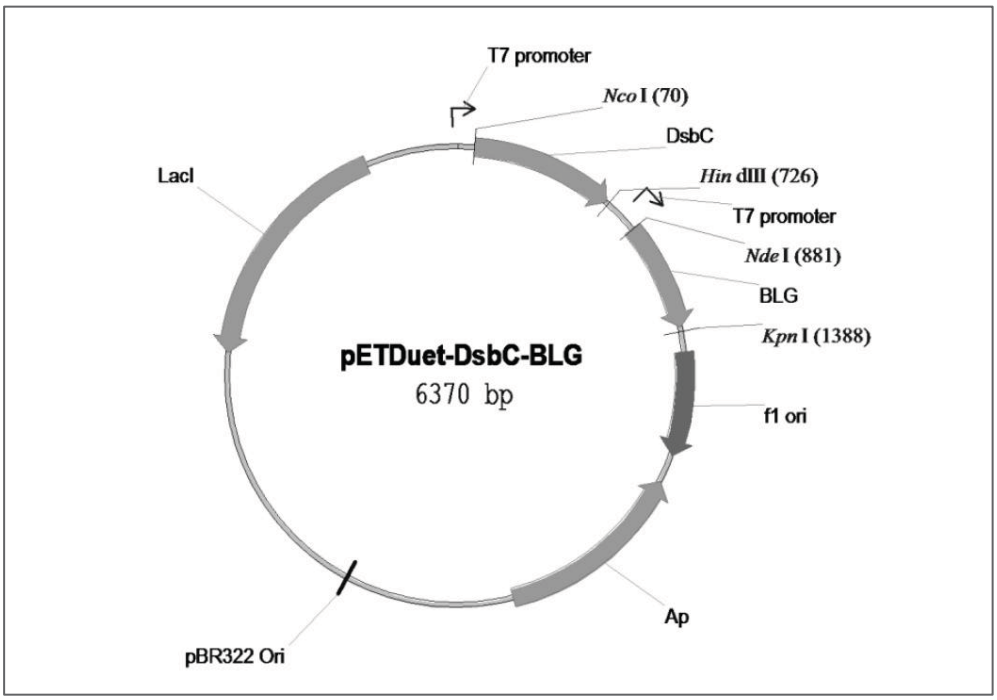


Figure 1.6 Expression Constructs Developed by Ponniah *et al*. (2010).
The pETDuet-1 plasmid was designed for the co-expression of two target genes. In these studies the target genes were leaderless DsbC (inserted into MCS1) and leaderless BLG (inserted into MCS2) genes. Both MCS' are preceded by a T7 promoter/lac operator and a ribosome binding site. The vector carries the pBR322-derived ColE1 replicon, the LacI gene and the ampicillin resistance gene (Ap).

1.6.3 Purification of β-Lg Variants

Ponniah and co-workers (2010) purified recombinant β-Lg from *E. coli* using an established protocol exploiting the extremely acid stable nature of β-Lg, a distinctive feature not generally shared by most other proteins (Mailliart & Ribadeau-Dumas, 1988). β-Lg was predominantly found in the soluble fraction of the pH 2.6 and 7 % NaCl (w/v) partially purified cell lysate, while the precipitate was found to contain most of the bacterial protein contaminants. The recombinant protein was recovered by salting out the centrifuged supernatant with 30 % NaCl (w/v), and the precipitate containing sufficiently pure β-Lg was solubilised in neutral buffer. Ponniah and co-workers suggested that the salting out of proteins at low pH acted as an efficient method of ridding β-Lg of species that were not correctly folded in this system, as 'native-like' β-Lg is very acid stable.

1.7 Exploring Protein Dynamics Using High-Field NMR Spectroscopy

Nuclear magnetic resonance spectroscopy (NMR) is a powerful tool that has allowed the characterisation of protein dynamics in solution, illuminating the mechanisms by which these molecules function and also behave in response to different treatments such as heat. It is uniquely suited to study many dynamical processes as site-specific information can be acquired for a diversity of movements across a spectrum of time-scales (Figure 1.7). NMR spectroscopy parameters are sensitive to the types of motions exhibited by a typical protein, capturing movements taking picoseconds, such as the rapid libration of a backbone N-H vector, to events that take seconds or much longer such as protein unfolding (for a review see Palmer *et al*., 1996).

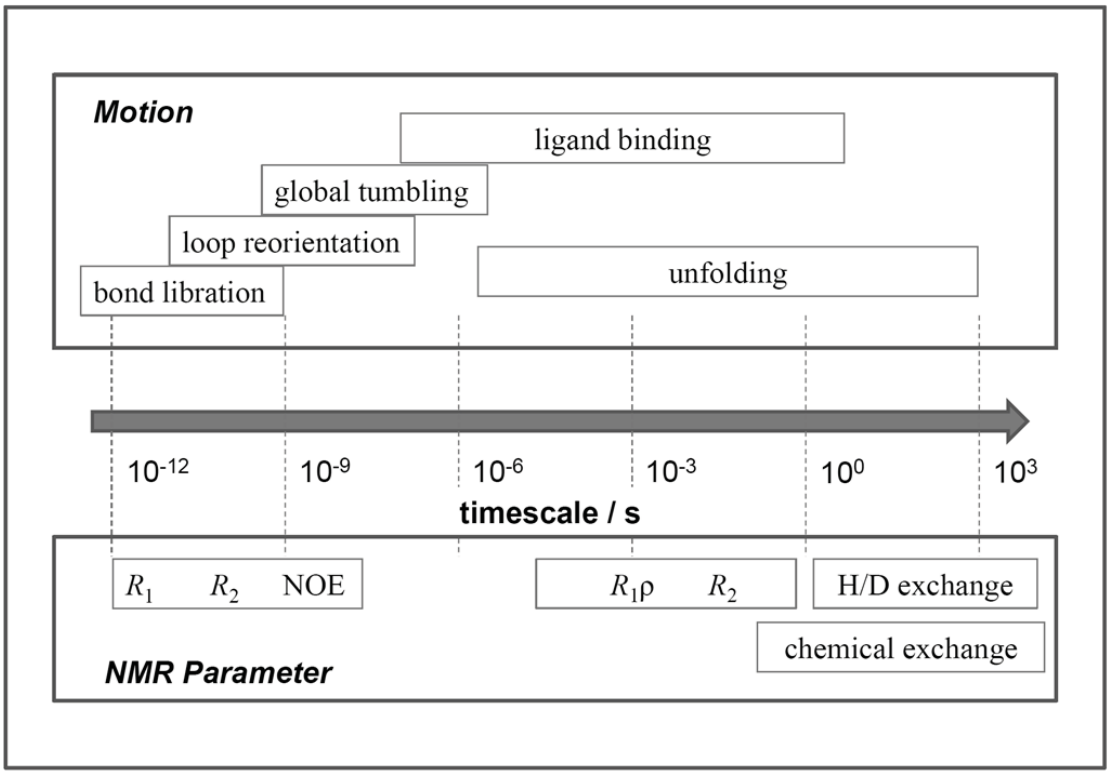


Figure 1.7 Time-scales of Protein Dynamics Measurable by NMR Spectroscopy.
NMR spectroscopy parameters are sensitive to the types of motion exhibited by the protein backbone. Backbone dynamics of the β-Lg variants have been probed *via* ^{15}N R_1, R_2 and NOE experiments (Uhrínová *et al*., 2000).

1.7.1 ^{15}N Relaxation Experiments and Model-Free Analysis

Heteronuclear ^{15}N relaxation experiments have been used extensively to characterise the backbone dynamics and motional properties of protein molecules in aqueous solution.

Detailed information of internal motions can be acquired from the calculation of backbone amide ^{15}N relaxation rates, granted these internal motions are of the order or more rapid than the overall molecular tumbling time (τ_m). Modified versions of the ^{15}N,^{1}H-HSQC give a series of 2D spectra from which R_1 (^{15}N longitudinal relaxation rate), R_2 (^{15}N tranverse relaxation rate) and ^{1}H-^{15}N ss-NOEs (steady-state heteronuclear Overhauser effect) are obtained for the backbone N-H vector of each assigned residue. This raw data can then be interpreted into the context of motional processes relating to the protein backbone by means of the 'model-free' protocol. The model-free protocol employs the Lipari and Szabo model-free formalism (Lipari & Szabo, 1982a, Lipari & Szabo, 1982b), with the extension of a two time-scale model. The mechanism of ^{15}N relaxation is relatively simple compared to ^{13}C or ^{1}H relaxation, as there is no homonuclear *J* coupling to be concerned with. A comprehensive explanation of these experiments and the model-free protocol is provided in Materials and Methods (Section 2.4.1).

Using the model-free protocol, the motion of the backbone N-H bond is described by three parameters. The order parameter (S^2) describes the amplitude of motion inside a cone (Figure 1.8). Typically, the lower the S^2 value the more flexible the N-H bond is deemed. The correlation time (τ_e) defines the time-scale of the internal dynamics, describing the picosecond to nanosecond time-scale of the motion, whereas the conformational exchange constant (R_{ex}) accounts for slower conformational exchange movements, pointing towards motions on micro- to millisecond time-scales, hence generally only affecting a few residues.

1.7.2 Examination of Backbone Dynamics with Other Methods

Alternative methods have been developed for estimating backbone dynamics of proteins based on NMR and X-ray crystallography derived tertiary structures, or alternatively from NMR backbone assigned chemical shifts. Explanations of these methods are found in Sections 2.4.8 and 2.4.9.

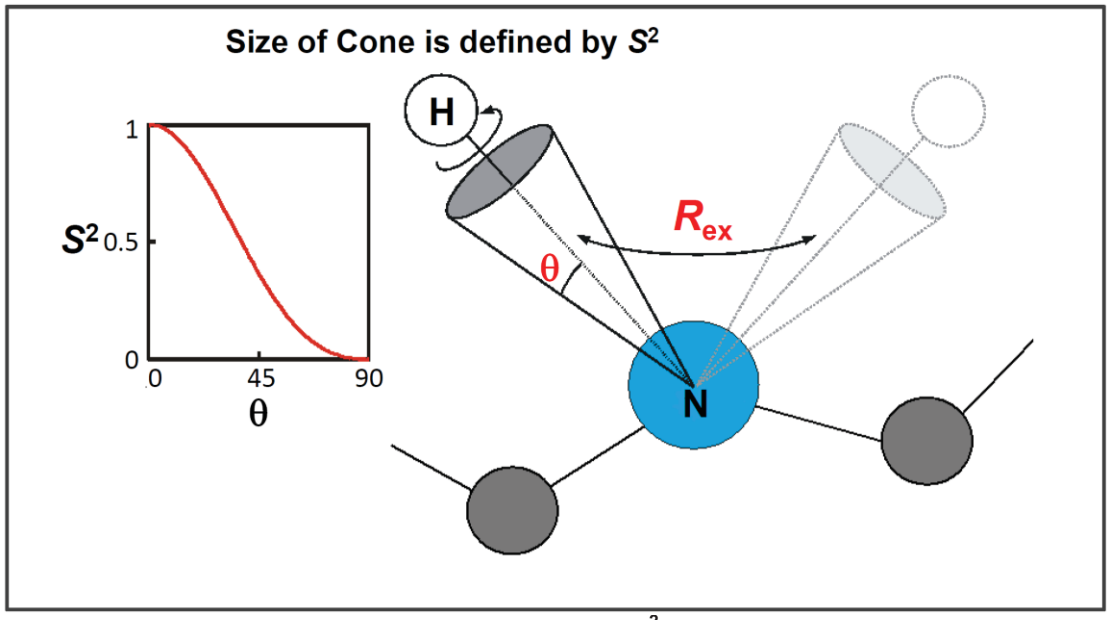


Figure 1.8 Description of the Order Parameter, S^2, and the Conformational Exchange Parameter R_{ex}.

The amplitude of motion is defined by the S^2 order parameter, whereas R_{ex} describes motions of N-H bonds undergoing conformational exchange. Typically, higher values of S^2 point towards higher restraint of the N-H bond, with 1.0 meaning that the bond is totally rigid. As values become lower the amplitude of motion becomes greater. Values of R_{ex} describe the conformational exchange of the backbone amide on a ms - μs time-scale. Figure reproduced from Rule & Hitchens (2006).

1.7.3 Backbone Dynamics of β-Lg

The backbone dynamics of β-Lg A, at low pH and at 37 ˚C, have been interpreted, primarily with ^{15}N relaxation data, as experimental data were not of satisfactory quality to permit their processing with the more sophisticated model-free method (Kuwata *et al*., 1999, Uhrínová *et al*., 2000). Experimental values showed that 'on the whole' the β-Lg's backbone is not particularly mobile. However, relaxation parameters corresponding to the N-terminus (residues 1-10) point towards a high degree of flexibility in this region, except for Met7 that is seen to form an H-bond with Val94 in the Z-lattice crystal structure (Qin *et al*., 1998b). Relaxation values for backbone amides of residues sitting in the β-strands, except in strands β-B, β-D, and β-I, whose hydrogen bonding potentials are not fully achieved, are rigid in nature, as the rather uniform relaxation values are somewhat high. Thirteen residues, with high NOE values greater than one SD (Standard Deviation) above the mean were found predominantly in the β-stands and the major three-turn α-helix, with most having side-chains lending themselves into the hydrophobic cavity (Uhrínová *et al*., 2000). Uhrínová *et al*. noted that some residues in the A/B loop, with high relaxation values, are markedly rigid in

behaviour, because of the occurrence of the short α-helix (residues 29-32) in this section. Five residues sitting in the C/D, E/F and G/H loops, positioned at the top of the barrel, displayed relaxation parameters that are indicative of conformational exchange on a micro- to millisecond time-scale. The authors speculated that these residues are positioned in regions that could facilitate the ingress of ligands into the hydrophobic core.

1.8 Studies of β-Lg Variants

1.8.1 Polymorphic β-Lg Variants

The occurrence of genetic polymorphism among milk proteins was first reported by Aschaffenberg and Drewry (1955), who observed that when milk samples from individual cows were subjected to electrophoresis on filter paper, samples produced either one or another or a mixture of two electrophoretically distinct bands, which were subsequently named β-Lg A and β-Lg B. Since then ten genetic polymorphic variants are known (A, B, C, D, E, F, G, H, W, X, Y, Z and Dr). Genetic polymorphism has been a consequence of a mutation to the genetic sequence, giving rise to an amino acid substitution along the translated polypeptide chain, which has since been able to be maintained by bovine populations.

1.8.2 Structural Differences of β-Lg Variants A, B and C

The effect of the substitutions on the secondary and tertiary structure of the β-Lg A, B and C protein variants appears to be subtle. A comparison of trigonal crystal forms between the A and B variants showed that the B variant has a less well packed core than the A variant, caused by the Val118Ala substitution (A→B) on the β-H strand as Val118 introduces a larger side-chain into the "core" region of the barrel (Qin *et al.*, 1999). The second substitution (Asp64Gly), sitting between β-strands C and D, causes a changed charged allocation in the C/D loop. Orthorhombic crystals structures of the three A, B and C variants (Bewley *et al.*, 1997) at neutral pH, showed that the Gln59His substitution (B→C) on strand β-C appears to have its main effect in altering stabilising interactions on the surface by switching a hydrogen bond formed with Glu44, positioned on strand β-B, with a salt bridge. Differences in structure observed amongst β-Lg A, B and C are not due to a difference in the location of the free thiol group as this is conserved within the variants.

1.8.3 β-Lg from Other Species

The most recent non-bovine β-Lg structure published is of rangiferine (reindeer) β-Lg, which shares similar monomeric tertiary structure and dimeric quaternary structure to bovine β-Lg, which is unsurprising as its primary sequence is 95 % identical and its peptide chain is also 162 residues in length (Oksanen *et al.*, 2006). Equine β-Lg shares 58 % primary sequence identity with bovine β-Lg but is monomeric over a wide pH range, whereas porcine (pig) β-Lg shares 63 % amino acid identity with bovine β-Lg and is dimeric below pH 5 and monomeric at pH 5 and above. As a consequence both can be analysed by NMR spectroscopy at neutral pH (Kobayashi *et al.*, 2000, Ugolini *et al.*, 2001). A complete structure of equine β-Lg has not been released; however, it has been studied by NMR spectroscopy to understand the process of protein folding (Kobayashi *et al*., 2000). The crystal structure of porcine β-Lg at pH 3.2 reveals that it dimerises by exchanging the N-terminal region domains forming a rather different quaternary structure in comparison to bovine β-Lg (Hoedemaeker *et al.*, 2002). The core β-barrel structure is superimposable with differences mainly found in the flexible loop regions.

1.9 Effects of Heat on β-Lg

1.9.1 Effects on Bovine β-Lg During Heat Treatment of Milk

In the dairy industry, heat treatment of milk and its products is an unavoidable operation, as it is essential for both food safety and technological purposes. Of technological importance is bovine β-Lg's role in the industrial thermal processing of milk and the characteristics imparted to milk products in regards to heat-induced changes affecting milk stability. Its behaviour during heating has attracted significant attention as it has been linked to fouling of heat exchangers caused by the deposition of material during pasteurisation and Ultra-High-Temperature (UHT) processing, the spoilage of thermally treated milk products generating off flavours and the production of sediments or gels in stored milks. The denaturing of β-Lg on the other hand is also advantageous as the formation of gels can be induced, imparting functional characteristics in foods (Belloque & Smith, 1998).

Of commercial significance is that milks that contain one of either β-Lg A, B or C variants behave differently when subjected to thermal processing, though these proteins differ only by up to three residues in sequence and all share similar tertiary structure.

Stability of the three variants is in the order of B < A < C, where B is least stable (Manderson *et al.*, 1998, Manderson *et al.*, 1999a, Manderson *et al.*, 1999b).

1.9.2 The Effects of Heat Treatment to Purified β-Lg Variants A, B and C

Between the temperatures of 30 ˚C and 55 ˚C dimeric β-Lg dissociates into monomers (Hambling *et al.*, 1992), and as temperature increases, β-Lg unfolds and aggregates through a series of parallel and consecutive steps, some of which involve disulfide bond interchange, and some of which involve hydrophobic-driven association reactions (Manderson *et al.*, 1999b).

Manderson and co-workers (1998) showed that heat-treated solutions of purified β-Lg A contain higher populations of aggregated and stable unfolded monomeric protein species than β-Lg B or C samples, which had been treated in an identical manner, analysed *via* 1D and 2D polyacrylamide gel electrophoresis (PAGE) (Manderson *et al.*, 1998). Also, it was shown that some species are held together by a combination of non-covalent interactions and disulfide bonding, as shown by alkaline-PAGE and sodium dodecyl sulfate (SDS)-PAGE, providing evidence against the idea that β-Lg aggregates *via* a simple polymerase reaction. This group speculated that the differences in the proportions of β-Lg species that form disulfide bonded aggregates are due to the polymorphisms of the variants. These investigations showed that β-Lg A favours the formation of hydrophobically driven associations and the formation of non-native monomers as intermediates in the aggregation pathway.

Following PAGE analyses, the scope of their investigations were broadened and extended by probing dilute solutions of purified β-Lg A, B and C that were heated for 10 minutes at temperatures between 40 ˚C and 95 ˚C. Analyses were performed with CD spectroscopy and with tryptophan and ANS (1,8-anilinonaphthalene sulfonate) fluorescence (Manderson *et al.*, 1999a, Manderson *et al.*, 1999b), which showed that during the initial stages of heat treatment the chiral environment for Trp19 is lost, a thiol group becomes available, an ANS binding site is formed and changes occur to the far-UV-CD spectrum. Results from these studies coupled with those from PAGE analyses showed that these changes corresponded with loss of native-like protein observed by alkaline PAGE, and not the loss of monomeric protein as assessed by SDS-PAGE (i.e. protein that had not aggregated by disulfide bonds). The following ongoing aggregation of the protein that occurs, as monitored by light scattering, is possibly linked to changes in tryptophan fluorescence and at the highest temperatures a change in ANS fluorescence.

The account of differences in milk in response to heat treatment by each of the three variants is complex. However, it is thought to stem from structural differences of the variants that are due to the effects of the ion-pair involving His59 (β-Lg C), a destabilising hollow caused by the Val118Ala (A→B) substitution and a altered charge distribution within the C/D loop arising from the the Asp64Gly (A→B) substitution (Manderson *et al.*, 1998, Manderson *et al.*, 1999a, Manderson *et al.*, 1999b).

Manderson and co-workers speculated that structural differences in β-Lg-C could possibly arise from a difference in arrangement of the Cys66-Cys160 bond, as the α and β carbons of residue 59 are in a different configuration in β-Lg-C (Gln59His), as observed with near-UV-CD measurements of disulfide bonds at 46 ˚C, whereas the smaller differences in spectra between the A and B variants are caused from the Val118Ala substitution rather that the Asp64Gly. These researchers attributed differences in the denaturation curves of β-Lg A, B and C at pH 6.7 and pH 7.4 to structural differences caused by the polymorphisms. A factor that the authors did not consider in their investigations is the differences in the dynamics of the variants.

1.9.3 Preliminary NMR Spectroscopy Studies looking at Site-Specific Changes in β-Lg upon Increases in Temperature

Although the β-Lg variants have slightly different thermal stabilities, the different conformational changes that take place within each variant are likely to be subtle and more readily detectable by tracing the changes of as many as possible of the individual residues in the variants using techniques such as NMR spectroscopy. As a lead up to such a comparative study, and follow-on to a study published by Belloque and Smith (1998) who tracked hydrogen/deuterium (H/D) exchange of 22 residues at three temperatures (45 ˚C, 55 ˚C and 75 ˚C), Edwards *et al.* (2002) tracked thermal resistance in many residues in native β-Lg A at pH 3.0, by means of H/D exchange using a 400 MHz NMR spectrometer. The exchange behaviour of the backbone amide protons, examined between 37 ˚C and 80 ˚C, showed that the unfolding of the β-barrel above 75 ˚C was extensive, whereas regional variations in structure were observed at lower temperatures. Residues positioned in loops and the terminal regions of the protein molecule showed swift H/D exchange at 37 ˚C, while H/D exchange in the α and β structural regions were observed in the following order with increasing temperatures: D-E strand (55-60 ˚C); C-D strand and α-helix (60-65 ˚C); A-B, A-I and E-F strands (65-70 ˚C); and A-H, B-C and F-G strands (75-80 ˚C) (Edwards *et al*., 2002). At 80 ˚C

the only identifiable H_N signal was observed for Phe105 positioned in the G-H pair of disulfide linked β-strands. D/H exchange and SDS-PAGE experiments showed that many of the effects of heating β-Lg to 80 ˚C were reversible**.**

To date NMR spectroscopy relaxation data published with respect to comparisons of the three variants or dynamical comparisons at any temperatures of the β-Lg A, B or C variants has not been released. Such studies would be fundamental in illuminating the mechanisms of the site-specific conformational changes, which take place within the protein/s.

1.9.4 Other Factors Affecting Heat Treatment of β-Lg

Although other milk components, protein concentration, pH, temperature, ionic strength and buffer type all appear to contribute to the stability of β-Lg, making the account more complex, what is known is that studies of solubility, conformation, thiol reactivity, gelation and thermal aggregation all show differences between the genetic variants. For a comprehensive review of these investigations refer to Hill *et al.* (1997).

1.10 The Lipocalin Family

β-Lg's eight-stranded β-barrel is a major structural motif found in a family of proteins called the lipocalins, which share small segments of sequence homology (Brownlow *et al*., 1997). The lipocalins are a diverse, widely distributed family of extracellular proteins that are best known for binding small hydrophobic or amphiphilic molecules (Flower, 1996, Pervaiz & Brew, 1985).

The three-dimensional structure of the lipocalins points towards a very close structural relationship, even though the sequence conservation may appear to be only slightly significant. The shared fold and structural alignments of several of the family members are depicted in Figure 1.9 (A) and (B). The β-strands, labelled B-H, are connected by a series of +1 hairpins, whereas β-strands A and B are connected by a large Ω linkage that incorporates a 3_{10} like helix. An additional 3_{10}-like helix is also positioned preceding the β-A strand. The major α-helix present between β-strands H and I is conserved in all of the lipocalins; however, its length and location relative to the axis of the flattened barrel varies amongst its members. Pair-wise sequence comparisons amongst the lipocalins frequently drop below 20 %; however, each share either two or three characteristic conserved sequence motifs.

It is the differences of the internal cavity and the external loop structure which affords a variety of different binding modes each capable of receiving ligands of different size, shape and chemical nature. β-Lg has the ability to bind retinol (Papiz *et al*., 1986), long-chained fatty acids found in milk (Wu *et al.*, 1999) and a number of other small hydrophobic molecules *in vitro* (Kontopidis *et al*., 2004). Other members of the lipocalins include β-Lg proteins from non-bovine species, the pheromone-binding proteins, MUP-I and α-$_{2u}$-globulin, present as the major proteins in mouse and rat urine respectively, bovine odorant-binding protein (OBP) which may be implicated in binding and transporting odorants, bilin-binding protein (BBP) from the cabbage butterfly, *Pieris brassica,* which binds biliverdin, and the retinol-binding protein (RBP), the main protein for transporting retinol in serum.

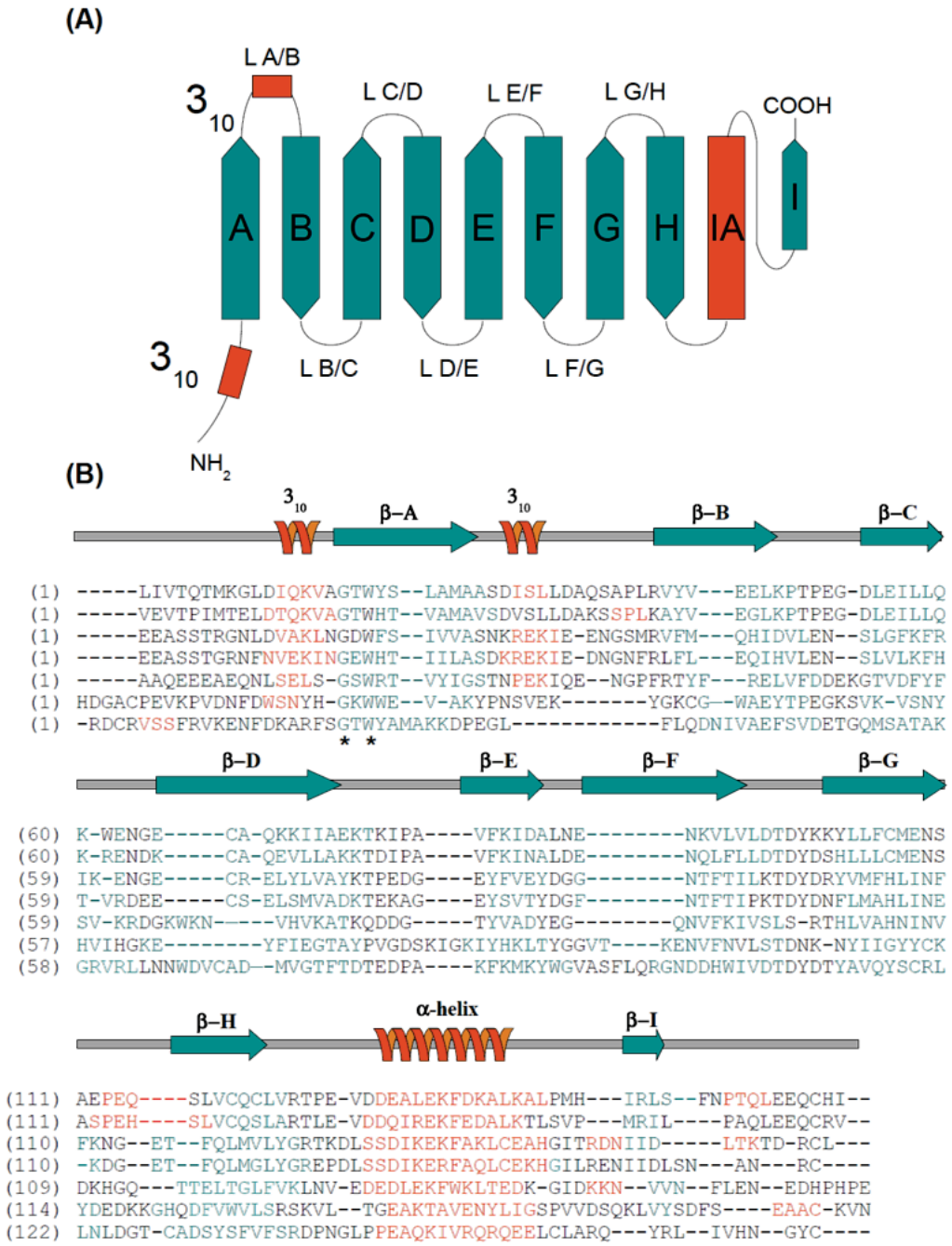


Figure 1.9 Schematic Diagram and Structural Alignments amongst Several Lipocalin Family Members.

Schematic diagram of the lipocalin fold. Figure modified from Flower (2000). Teal depicts β-strands whereas orange represents α-helical regions. (B) Structural alignment based on topological equivalence of the lipocalins in which PDB coordinates have been deposited. BLG, bovine β-Lg (Brownlow *et al.*, 1997); PLG, porcine β-Lg (Hoedemaeker *et al.*, 2002); A2uG, α_{2u}-globulin from rat (Chaudhuri *et al.*, 1999); MUP, major urinary protein from mouse (Böcskei *et al.*, 1992); OBP, odorant-binding protein (Tegoni *et al.*, 1996); BBP, bilin-binding protein (Huber *et al.*, 1987); and RBP, retinol-binding protein (Cowan *et al.*, 1990). Figure adapted from Brownlow (1997). The two absolutely most conserved residues in the GXW motif are marked by asterisks.

1.11 Function of β-Lg

Although the biological role of β-Lg is not yet known, functional studies have mainly focused on its putative role in binding nutrients in the cow's milk, and transporting them intact to the feeding neonate's intestine (Papiz *et al.*, 1986, Said *et al.*, 1989). β-Lg's extremely acid and proteolytic resistant nature, coupled with its ability of binding a broad spectrum of small hydrophobic molecules, such as retinol and palmitic acid within its barrel (Wu *et al*., 1999), have supported this idea.

The popular hypothesis that β-Lg acts as a retinol transporter is suspect, although *in vitro* studies have shown its association with retinol to form soluble complexes (Futterman & Heller, 1972), and its protein structure and configuration of intron and exon genetic sequences in relation to the protein's 3D conformation is very similar to that of the lipocalin RBP (Ali & Clark, 1988). Structural studies have shown the tip of retinol's isoprene tail protruding from the β-barrel (Kontopidis *et al*., 2002), being exposed to the environment, unprotected, and *in vitro* investigations have shown the elocation of retinol from its complex with β-Lg with a retinol degradation product (Hemley *et al.*, 1979), questioning this idea. A study displaying the specific enhanced intestinal uptake of retinol bound to β-Lg in suckling rats (Said *et al*., 1989), is debatable in identifying its function as β-Lg is absent in rat milk. Retinol in milk is primarily associated with fat globules, and the binding of retinol to β-Lg in milk has not yet been observed. An alternative argument is that retinol is simply transported from mother to neonate through the fat phase of the milk (Kontopidis *et al*., 2004).

It has been proposed that β-Lg functions to remove free fatty acids in milk, produced by pregastric lipases, to assist in the digestion of milk (Pérez *et al.*, 1992). This has been supported by evidence that β-Lg has been isolated with free fatty acids, primarily palmitic and oleic acids, from fresh milk (Pérez *et al.*, 1989) and competition experiments with retinol and palmitic acid demonstrated only palmitate binding to β-Lg (Kontopidis *et al*., 2002). However, this hypothesis is questionable as β-Lgs from a number of animals, including equine and porcine species, cannot bind to lipids (Pérez *et al.*, 1993), Also NMR spectroscopy studies showed that β-Lg's affinity for palmitate is lost at low pH, questioning the notion of that it is a transporter protein, as it would lose its ligand in the stomach (Ragona *et al.*, 2000).

Binding studies may only reflect β-Lg's affinity for a vast range of small hydrophobic molecules. Another suggestion is that it serves as a supply of nutritional protein for the calf, with perhaps some other function in the mother (Sawyer & Kontopidis, 2000). However, because β-Lg is extremely resistant to acid proteolysis and has even been isolated intact from the breast milk of humans, who have drunk bovine milk, its role as a dietary source of protein seems unlikely.

1.12 Aims of this Investigation

Overall Goals: To develop the relationships between structure and dynamics of bovine β-Lg by means of heterologously expressed forms of natural variants and selected mutants using high-field NMR spectroscopy.

Hypothesis A: *That β-Lg's observed motional behaviour is related to its structure.*

Aims:

i) To develop a suitable system for the expression of β-Lg C.

ii) To assign the $^{15}N,^{1}H$-HSQC cross-peaks of β-Lg C.

iii) To investigate the ^{1}H-^{15}N backbone dynamics of the isotopically labelled bovine β-Lg C variant at 305 K using ^{15}N relaxation data and the model-free protocol.

iv) To compare β-Lg A respective ^{15}N model-free order parameters with those derived with two other methods.

v) To develop a system for the expression of covalently linked dimeric β-Lg protein (A34C).

Hypothesis B: C*hanges in temperatures can be correlated with the differences in protein flexibility behaviour observed for bovine β-Lg-C.*

Aims:

i) To assign $^{15}N,^{1}H$-HSQC cross-peaks of β-Lg C at 313 K and 320K.

ii) To investigate and compare the model-free ^{1}H-^{15}N backbone dynamics of the isotopically labelled bovine β-Lg C variant at 305 K, 313 K and 320 K.

Hypothesis C: *Changes in molecular dynamics caused by both external and internal amino-acid substitutions can be correlated with the differences in protein flexibility behaviour observed for these natural variants of bovine β-Lg.*

Aims:

i) To assign the ^{15}N,^{1}H-HSQC cross-peaks of β-Lg A, β-Lg B at 305 K.

ii) To investigate the ^{1}H-^{15}N backbone dynamics of β-Lg variants A, B and C at 305 K.

iii) To investigate and compare the model-free ^{1}H-^{15}N backbone dynamics of isotopically labelled recombinant β-Lg A, β-Lg B and β-Lg C variants at 305 K.

2

Materials and Methods

2.1 General Materials and Methods

Appendix A.1 lists the sources of common chemicals and buffers used throughout this thesis. Abbreviations and/or formulae for these chemicals are also included in this appendix.

2.1.1 Purified Water

Purified water was sourced from a Sybron/Barnstead NANOpure II filtration system (Maryland, US). Two ion exchangers and two organic filter cartridges were incorporated in this system. All water used in this thesis was obtained from this source, unless otherwise stated. Purified water is simply referred to as H_2O throughout this study.

2.1.2 General Buffers and Solutions Used in this Study

General buffers and solutions used in this study are included in Table 2.1.

Solution Name	Method
RF1 solution	100 mM RbCl, 50 mM $MnCl_2$, 30 mM potassium acetate, 10 mM $CaCl_2$, pH adjusted to 5.8 with concentrated acetic acid
RF2 solution	10 mM RbCl, 10 mM MOPS, 75 mM $CaCl_2 \times 6H_2O$, 15 % v/v glycerin, adjusted to pH 5.8 with concentrated NaOH
6 × DNA loading dye	0.2 % (w/v) bromophenol blue in H_2O, and 50 % (v/v) glycerol
Tris-acetate-EDTA buffer (TAE)	40 mM Tris-HCl, 20 mM acetic acid, 1 mM EDTA; pH 8.0
Lysis buffer (20 mM Bis-Tris-HCl, pH 6.5)	20 mM Bis-Tris, adjusted to pH 6.5 with concentrated HCl; filtered with a 2.2 μm filter membrane (Millipore; MA, USA)
High salt buffer	20 mM Bis-Tris, 1 M NaCl, adjusted to pH 6.5 with concentrated HCl; filtered with a 2.2 μm filter membrane
Dialysis buffer	50 mM disodium hydrogen orthophosphate, adjusted to pH 7.5 with sodium dihydrogen orthophosphate
0.1 M HCl	10 % 1 M HCl (v/v) added to 90 % H_2O (v/v)
KH_2PO_4 buffer (pH 6.5)	50 mM potassium dihydrogen orthophosphate, adjusted to pH 6.5 with dipotassium hydrogen orthophosphate; filtered with a 2.2 μm filter membrane
KH_2PO_4 buffer (pH 2.6)	50 mM orthophosphoric acid, adjusted to pH 2.6 with 50 mM KOH
Sodium azide	3 % (w/v) sodium azide in H_2O

Table 2.1 List of general solutions and buffers used in this study.

2.1.3 Media

Sterilisation

Solutions were filter sterilised with sterile syringes (Terumo Corporation; Tokyo, Japan) and sterile 0.22 μm syringe filters (Millipore; MA, USA).

Solutions and media were autoclaved at 121 °C and 2×10^5 Pa for 20 minutes.

Luria-Bertani (LB) Media

LB media 25.0 g

Made up to 1 L with H_2O and was autoclaved.

Solid LB Media was prepared by adding 1.5 % (w/v) agar to LB broth prior to autoclaving.

Minimal Media Components

Table 2.2 comprises a list of components that constitutes minimal media used in Section 2.3.3. ^{13}C-D-glucose and ^{15}N-ammonium sulfate, required for ^{13}C- and ^{15}N-labelling of β-Lg protein variants, were purchased from Cambridge Isotope Laboratories (Andover, USA).

Solution Name	Method
5 × M9 salts	64 g $Na_2HPO_4.7H_2O$, 15 g KH_2PO_4, 2.5 g NaCl; dissolved in purified water to a volume of 1 L; autoclaved
1 × M9 salts	1 part 5 × M9 salts to 4 parts purified water; autoclaved
1M $CaCl_2$ solution	1 M $CaCl_2$ dissolved in purified water; autoclaved
1M $MgSO_4$ solution	1 M $MgSO_4$ dissolved in purified water; autoclaved
100 × Trace elements solution	(per 100 mL solution) 0.08 g $FeCl_3.6H_2O$, 0.06 g $MnCl_2.4H_2O$, 0.04 g $CoCl_2.6H_2O$, 0.04 g $ZnSO_4.7H_2O$, 0.03 g $CuCl_2.2H_2O$ and 0.002 g H_3BO_3
Thiamine solution	(per 12 mL solution) 0.12 g thiamine dissolved in purified H_2O to a volume of 12 mL; filter sterilised
10 % (w/v) Ammonium sulfate solution	(per 20 mL solution) 2 g of $(NH_4)_2SO_4$ dissolved in H_2O to a volume of 20 mL; filter sterilised
10 % (w/v) ^{15}N Ammonium sulfate solution	(per 20 mL solution) 2 g of ^{15}N labelled $(NH_4)_2SO_4$ dissolved in H_2O to a volume of 20 mL; filter sterilised
20 % (w/v) D-glucose solution	(per 40 mL solution) 8 g of D-glucose dissolved in H_2O to a volume of 40 mL; filter sterilised
20 % (w/v) ^{13}C-D-glucose solution	(per 40 mL solution) 8 g of ^{13}C-labelled D-glucose dissolved in H_2O to a volume of 40 mL; filter sterilised

Table 2.2 Chemicals and solutions that constitute minimal media used in this study.

Antibiotic Stock Solutions

The antibiotics were prepared as stock solutions (Table 2.3), filter sterilised (Section 2.1.3 and then stored in one mL aliquots at -20 °C. Antibiotics were added to cooled (~50 °C) sterile media or to starved cultures to final concentrations.

Antibiotic	**Stock Solution (mg mL^{-1})**	**Final Concentration (μg mL^{-1}) *E. coli***	**Source**
Ampicillin (Na-salt)	100 in H_2O	100	USB Corporation (Cleveland, US)
Tetracycline (-hydrochloride)	12.5 in ethanol (70 %)	12.5	Sigma (Auckland, NZ)
Kanamycin (-sulfate)	15 in H_2O	15	Sigma (Auckland, NZ)

Table 2.3 Antibiotics, stock solutions, final concentration and sources.

2.1.4 Glycerol Stocks

1 mL glycerol stocks were generated by mixing 800 μL of culture with 200 μL of 70 % (w/v) sterile glycerol, snap frozen in liquid nitrogen and then stored at -80 °C.

2.1.5 Measurement of Optical Density (OD) of Cultures

The cell density of a culture was determined by measuring the optical density (OD) at a wavelength of 600 nm with a Smart Spec^{TM} Plus Spectrophotometer (BioRAD, Milan, Italy). Sterile culture medium was used as a reference.

2.2 Methods for Deoxyribonucleic Acid (DNA) Work

2.2.1 Site-Directed Mutagenesis Strategy

Sited-directed mutagenesis is a technique widely used by molecular biologists to study the structure and function of proteins. It is a controlled process that introduces rationalised, site-specific mutations in DNA sequences, by creating single or multiple base substitutions, insertions or deletions. Mutations in the DNA sequence result in a mutation in the mRNA sequence, leading to a mutated protein. Proteins are mutated for a variety of reasons including for comparative structural and/or functional analyses, enhancing protein solubility, increasing purification efficiency and/or to introduce desired functional traits within the studied protein or its gene. In this study site-specific DNA base-pair substitution/s were required to generate expression vectors necessary for the production of isotopically labelled bovine β-Lg variant C and both variant A and B covalently bonded dimers (Ala34Cys).

PCR Site-Directed Mutagenesis of the BLG Gene Using a Rational Approach

The plasmids were mutated at defined sites in the BLG gene insert and amplified with a PCR-based protocol developed by Stratagene, using the QuikChange site-directed mutagenesis kit (Stratagene, CA, USA). This method requires a supercoiled double stranded DNA plasmid with an insert of interest, and two complementary oligonucleotide primers with the desired base substitution(s), which anneal to opposite sides of the vector construct. The proof-reading *PfuTurbo* DNA polymerase extends each primer during thermo-cycling to produce a mutated plasmid with staggered nicks. After PCR, DNA is treated with the restriction endonuclease, *Dpn*I, which will cut only fully or hemi-methylated 5'-G^{m6} ATC-3' sequences in duplex DNA. The digestion of the methylated parental DNA with *Dpn*I, results in selection of the PCR synthesised DNA (Nelson & McClelland, 1992). DNA extracted from the majority of *E. coli* strains is dam methylated, which happens shortly after synthesis by an enzyme adding a methyl group to the sequence 5'-GATC-3', therefore making it predisposed to digestion with *Dpn*I (Geier & Modrich, 1979).

The PCR-generated unmethylated plasmid, incorporating the introduced base change(s), is resistant to *Dpn*I digestion and used to transform *E. coli* competent cells, after which the staggered nicks within the plasmid are repaired by the *E. coli* host. The ratio of mutation-harbouring recombinant clones, using this method, achieves levels above 95 %

for 2.9 kb templates and more than 83 % for templates of 8 kb length (Papaworth *et al.*, 1992).

2.2.2 Bacterial Strains Used in this Thesis

Escherichia coli bacterial strains used in these studies are included in Table 2.4. The cells were used for cloning and expression of recombinant β-Lg.

Bacterial Strain	Genotype	Source	Comments
Top10 (*E. coli)*	F$^-$ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 Δ*ara*D139 (*ara-leu*)7697 *gal*U *gal*K *rps*L *end*A1 *nup*G	Invitrogen	Routine cloning, sub-cloning
Origami B (*E. coli*)	F$^-$ *ompT hsd*S_B(r_B^- m_B^-) *gal dcm lacY1 ahpC* (DE3) *gor522*:: Tn*10 trxB* (Kan^R, Tet^R)	Novagen	Expression host Improved disulfide bond formation in *E. coli* cytoplasm

Table 2.4 List of *E. coli* strains used in this thesis.

2.2.3 Template Plasmid Constructs Used

Template pETDuet-DsbC-BLG (variants A and B) plasmid constructs (Table 2.5) were provided by Dr. Komala Ponniah, Massey University. The expression vectors had been designed to co-express recombinant bovine β-Lg (A or B) and DsbC isomerase proteins in *E. coli* cells.

Vector	Size (bp)	Selection and other features	Source
pETDuet-DsbC-BLG A	6370	Amp^r; two MCS preceded by PT7 and RBS; pBR322 origin; *lacI*; leaderless DsbC inserted into site 1 (NcoI/HindIII); leaderless BLG A* (Ariyaratne *et al.*, 2002) inserted into site 2 (NdeI/KpnI)	Novagen
pETDuet-DsbC-BLG B	6370	Amp^r; two MCS preceded by PT7 and RBS; pBR322 origin; *lacI*; leaderless DsbC inserted into site 1 (NcoI/HindIII); leaderless BLG B (Ponniah *et al.*, 2010) inserted into site 2 (NdeI/KpnI)	Novagen

Table 2.5 List of expression vectors used as template DNA for PCR site-directed mutagenesis.

* The synthetic BLG A DNA sequence, designed by Ariyaratne *et al.* (2002), and its respective translated primary sequence are listed in Appendix A.4.

2.2.4 DNA Concentration

The DNA concentrations of plasmid samples were measured using a Nanodrop ND-1000 spectrophotometer.

2.2.5 Methods for Plasmid Purification

E. coli cells carrying plasmids were cultivated overnight in 5 mL LB medium with the appropriate antibiotics. Between 2 to 4 mL of overnight cell growth was harvested for 60 seconds at 14 000 × *g*. Plasmids were isolated and purified using the "High Pure Plasmid Isolation Kit" (Roche, Mannheim, GER) according to the manufacturer's instructions.

2.2.6 DNA Agarose Gel Electrophoresis Methods

DNA was separated on the basis of size by agarose gel electrophoresis. Routinely 8 µL of PCR reaction was mixed with 2 µL of 6 × loading dye (Table 2.1) and loaded onto a standard 1 % agarose Tris-acetate-EDTA gel (Table 2.1). The gels were routinely run at 80 V until the dye front had migrated to the other end of the gel. The gels were stained with an ethidium bromide solution (0.5 µg mL^{-1}) for 20 minutes and destained in H_2O for 5 minutes, before visualisation using an Ultraviolet trans illuminator (Bio-Rad). A 1 Kb plus DNA ladder (Appendix A.4) was used to estimate DNA lengths.

2.2.7 Transformation Methods

Preparation of Competent E. Coli Cells (Hanahan, 1983)

E. coli Top10/Origami B (DE3) cells were grown aerobically at 37 °C to an OD_{600} of 0.5-0.8. Prior to cultivating Origami B cells, kanamycin and tetracycline were added to LB growth media, to final concentrations (Table 2.3). This was necessary to maintain strain selection. Both *E. coli* strains were centrifuged (2 700 × *g* for 10 minutes at 4 °C), and then resuspended in ice-cold RF1 buffer (Table 2.1). Cells were placed on ice for a period of 30 minutes and centrifuged at 2 700 × *g* for 10 minutes at 4 °C, after which the supernatant was decanted. The cell sediment was resuspended in ice-cold RF2 buffer (Table 2.1), and incubated on ice for 10-15 minutes. Cells were distributed in 100 µL aliquots, snap frozen in liquid nitrogen and stored at -80 °C.

Transformation and Selection of E. Coli Cells

Competent *E. coli* Top10/Origami B (DE3) cells were transformed with pETDuet-DsbC-BLG plasmid using the heat shock method (Hanahan, 1983) as follows: 100 µL of Top10 or Origami B cells were incubated with 80 ng of plasmid constructs and placed on ice for 30 minutes, after which the reaction mixture was heat shocked at 42 °C for 90 seconds. Cells were returned to ice for two minutes before being transferred to a bench top shaker at 37 °C for 60 minutes. Top10 and plasmid transformation reactions were plated onto LB-ampicillin agarose plates. Origami B and plasmid transformation reactions were

spread on LB agarose plates containing ampicillin, kanamycin and tetracycline at the appropriate concentrations (Table 2.3).

2.2.8 Site-Directed Mutagenesis using the Polymerase Chain Reaction (PCR)

Site-Directed Mutagenesis Methods

Each PCR reaction (Table 2.6) consisted of 5 μL of 10 × reaction buffer, 2 μL (10 ng) of the appropriate plasmid (double-stranded DNA template), 1 μL (10 pmol) dNTP mix (deoxyribonucleic acids) purchased from Roche, and 1.25 μL (125 ng) of each respective forward and reverse primer (complementary mutagenic oligonucleotides), purchased from Integrated DNA Technologies (USA). Sterile H_2O was added to a final volume of 50 μL. 1 μL (2.5 units) of *PfuTurbo* DNA polymerase (thermostable proof-reading DNA polymerase) was added to each PCR reaction.

PCR Reaction No.	Plasmid DNA Template Constructs	Site-directed β-Lg insert Mutation (Translation)	Forward Primer Sequence 5' →3'	Reverse Primer Sequence 5' →3'
1	pETDuet-DsbC-BLG B	β-Lg-BΔGln59His (β-Lg C)	CTC GAG ATT CTG CTG CAC AAA TGG GAA AAC GGT GA	TC ACC GTT TTC CCA TTT GTG CAG CAG AAT CTC GAG
2	pETDuet-DsbC-BLG A	β-Lg AΔAla34Cys	GAT ATT AGT CTG CTG GAT TGT CAG TCG GCG CCG CTG CGT	ACG CAG CGG CGC CGA CTG ACA ATC CAG CAG ACT AAT ATC
3	pETDuet-DsbC-BLG B	β-Lg BΔAla34Cys	GAT ATT AGT CTG CTG GAT TGT CAG TCG GCG CCG CTG CGT	ACG CAG CGG CGC CGA CTG ACA ATC CAG CAG ACT AAT ATC

Table 2.6 List of PCR template DNA constructs and complementing primers.

These primers were used for PCR site-directed mutagenesis and amplification of plasmid constructs. Site-directed substitutions were targeted towards the BLG gene insert, by using primers engineered with base-pair mismatches (noted by an underline in the primer sequences), to produce the bovine β-Lg C protein variant (reaction one) and β-Lg AΔAla34Cys and β-Lg BΔAla34Cys (reactions two and three, respectively) mutant proteins upon subsequent translation.

PCR components were mixed in a sterile PCR tube. The reaction underwent repeated cycles of denaturation, annealing and elongation in a PCR cycler T gradient (Biometra). The cycling parameters used for each PCR reaction are listed in Table 2.7.

	Temperature	Time	Process
	95 °C	30 seconds	Denaturation
	95 °C	30 seconds	Denaturation
16 Cycles	55 °C	2 minutes	Annealing
	68 °C	6 minutes	Extension
	10 °C	∞	Cooling

Table 2.7 PCR conditions used for site-directed mutagenesis and amplification of expression vectors.

8 μL of PCR products were removed from the reaction mix and used to confirm PCR amplification by means of agarose gel electrophoresis (Section 2.2.6).

Digestion of Parental Plasmid Template with DpnI Enzyme

The remaining PCR reaction mixtures were placed in a 37 °C environment for one hour, after 1 μL of *Dpn*I enzyme had been added to each tube and mixed. Following digestion of parental DNA template, mutated plasmid constructs were transformed (Section 2.2.7) into *E. coli* Top10 strains (Table 2.4).

DNA Sequencing

Prior to sequencing the BLG gene insert, a single colony (a positive transformant) was selected aseptically. These cells were cultivated, and their harbouring plasmids propagated and isolated as described in Section 2.2.5. Sequencing of the BLG gene insert within the plasmid vectors was then performed by the Alan Wilson Centre sequencing service using Big Dye Terminator V3.1.

After the rationalised mutation was confirmed by sequencing, plasmids were transformed into *E. coli* Origami B expression cells (Section 2.2.7). A positive transformant was selected by isolating a colony aseptically and transferring it to 5 mL of LB broth containing antibiotics for cell selection (kanamycin and tetracylin) and an antibiotic for plasmid selection (ampicillin) to appropriate concentrations. The cells grew aerobically overnight at 30 ˚C and were then stored as glycerol stocks (Section 2.1.4).

2.3 Protein Biochemical Methods

2.3.1 Determination of Protein Concentration

The concentration of protein samples was determined by the Bradford assay (Bradford, 1976) or UV absorption (Scopes, 1974).

In instances when protein concentration was determined using the UV absorption method, it was measured at 280 nm on a Smart Spec™ Plus Spectrophotometer (BioRAD). The absorbance was converted into mg mL^{-1} using the Beer-Lambert law which states that:

$$Protein\ Concentration = \frac{Absorbance_{280} \times Molecular\ Weight}{Molar\ Extinction\ Coefficient_{280} \times Path\ Length} \quad (1)$$

where the protein concentration is calculated as mg mL^{-1}, the molar extinction coefficient for β-Lg is 17600 cm^{-1} M^{-1}, the optical path length is 1 cm and the molecular weight is equal to 18.4 kDa.

2.3.2 Polyacrylamide Gel Electrophoresis Methods (PAGE)

Protein molecules were separated by means of SDS-PAGE using a discontinuous buffer system with the methods described by Laemmli (1970). Reduced and non-reduced SDS-PAGE were performed using the the BioRAD Mini-PROTEAN 3 Cell and BioRAD MiniSub Cell GT apparatus (BioRAD Laboratories, California, USA).

SDS-PAGE

Stock Solutions

Acrylamide/Bis (40 % w/v) in H_2O), PAGE resolving gel buffer (1.5 M Tris-HCl; pH 8.8), PAGE stacking gel buffer (0.5 M Tris-HCl; pH 6.8), ammonium persulfate solution (APS) (10 % w/v in H_2O) and 10 % SDS solution (10 % w/v in H_2O) were made.

Five × SDS-PAGE Buffer

5 g SDS, 15 g Tris, 72 g glycine were made up to 1 L with H_2O.

Two × Reduced-SDS Sample Buffer

10 mL 1.5 M Tris (pH 6.8), 6 mL 20 % SDS, 30 mL glycerol, 15 mL β-mercaptoethanol, and bromophenol blue (1.8 mg) were made up to 100 mL with H_2O.

Two × Non-Reducing-SDS Sample Buffer

25 mL 1.5 M Tris (pH 6.8), 6 mL 20 % SDS, 30 mL glycerol and bromophenol blue (1.8 mg) were made up to 100 mL with H_2O.

The discontinuous system was comprised of 4 % acrylamide stacking and 15 % acrylamide separating gels containing 0.1 % SDS. 5 μL protein samples were mixed with 5 μL sample buffer (reduced or non-reduced) and boiled for two minutes. Once samples were loaded into the stacking gel, electrophoresis was carried out at 200 volts, until the dye front reached the end of the gel (approximately 45 minutes). Gels were stained with 0.1 % Coomassie Brilliant Blue R-250 and destained with a solution comprising 25 % (v/v) methanol , 10 % acetic acid (v/v) and 65 % H_2O (v/v).

2.3.3 Heterologous Expression of β-Lg Variants

Initial Innoculum

Origami B (DE3) [pETDuet-DsbC-BLG] cells were aseptically transferred from a glycerol stock into a 100 mL LB broth, containing appropriate concentrations of ampicillin, tetracycline and kanamycin antibiotics (Table 2.3), by means of a sterile plastic stick. The cells were left to grow aerobically at 37 °C overnight. The cells were then harvested by centrifugation (2 700 × *g* for 15 minutes). LB medium was removed and cells were washed to remove residual traces of LB broth by resuspending in 100 mL sterile 1 × M9 salts (Table 2.2) and centrifuging (2 700 × *g* for 15 minutes) to pellet cells. Supernatant was removed and cells were resuspended in 4 mL of sterile 1 × M9 salts.

Induction and Labelling of β-Lg and DsbC Proteins in Minimal Media

Cells were grown aerobically at 37 °C in sterile minimal media, which had been supplemented with sterile 2 mM $MgSO_4$, 0.2 mM $CaCl_2$, 0.3 mg mL^{-1} thiamine and 1 % of a 100 × trace elements solution (Table 2.2). After a 1.5 hour starvation period, the media was supplemented with sterile 1 % ammonium sulfate, 4 % D-glucose and ampicillin. When over-expressing ^{15}N labelled β-Lg protein, ^{15}N enriched ammonium sulfate was used as the sole nitrogen source. When over-expressing ^{15}N-^{13}C labelled β-Lg protein both ^{15}N-enriched ammonium sulfate was used as the sole nitrogen source and ^{13}C-D-glucose was used as the sole carbon source (Table 2.2). When an OD_{600} of 0.5 was reached, β-Lg and DsbC protein expression were induced by the addition of IPTG (0.5 mM final concentration). As a control, a culture was grown as above, using unlabelled ammonium sulfate and D-glucose, and without the addition of IPTG. The cells

were harvested by centrifugation (2 700 × *g* for 15 minutes at room temperature). Cells were transferred to a 25 °C environment and continued to grow aerobically for a period of twenty hours.

Cell Harvesting and Cell Lysis

E. coli cells from induction medium were harvested by centrifugation (4 500 × *g* for 20 minutes at 4 °C). The cells were resuspended in approximately 12 mL of lysis buffer (20 mM Bis-Tris-HCl, pH 6.5) and homogenised by two passes through the French Press (AMINCO, Silver Spring, MD, US) or by sonication. Cell lysate was centrifuged (30 000 × *g* for 20 minutes at 4 °C) to separate insoluble matter from soluble matter. Samples of the supernatant and pellet fractions were collected for solubility analysis.

2.3.4 Purification of Recombinant β-Lg

Anion Exchange Chromatography

Anion exchange chromatography of the cell lysate involved two key steps: first the binding of negatively charged protein to the oppositely charged ionic resin and second the elution of protein from the charges of the resin. Displacement of proteins from the resin was controlled by altering the ionic strength of the elution buffer from a 0 M to 1 M NaCl solution, whereby the proteins with the weaker ionic interactions elute first and the molecules that have stronger ionic interactions elute further along as salt concentration is increased.

Cell supernatant was filtered through a 0.8 μm syringe filter to remove any residual debris, before being subjected to purification by means of anion exchange chromatography with an Äkta explorer 10 S instrument (GE HealthCare; USA). The clarified lysate was applied to a HR 30/10 column (GE Healthcare; USA) packed in-house with a SOURCE 15Q resin (10 mm × 30 cm; bed volume 24 mL), which had been equilibrated with low salt buffer at 4 °C. Protein was eluted stepwise with a gradient of 0, 5, 10, 20, 25 and 100 % high salt buffer (five column volumes each) at a flow rate of 1.5 mL min^{-1}. The low salt buffer contained 20 mM Bis-Tris-HCl (pH 6.5); the high salt buffer contained 20 mM Bis-Tris-HCl (pH 6.5) and 1 M NaCl.

30 mL anion-exchanged purified fractions containing β-Lg were combined and then dialysed against 50 volumes of 20 mM $NaH_2PO_4.H_2O$ pH 7.5 buffer for 20 hours to remove salt. Following dialysis, protein samples were concentrated to 5 mg mL^{-1} by

means of ultra-filtration, utilising Vivaspin concentrators (Vivascience AG, Hannover, GER) with a molecular weight cut off of 10 kDa.

Acid and Sodium Chloride Precipitation

β-Lg was purified from remaining proteins by acid and sodium chloride precipitation using a modification of the method described by Mailliart and Ribadeau-Dumas (1988). To fractionate recombinant β-Lg protein from remaining protein contaminants, the pH of the dialysed protein sample was adjusted to 2.6 with the drop-wise addition of 0.1 M HCl. The protein sample was agitated for 15 minutes at room temperature and then centrifuged (30 000 × *g* at 4 °C for 20 minutes) to remove insoluble protein. Sodium chloride was added to the supernatant to a final concentration of 7 % (w/v) then stirred gently for a further 20 minutes. The precipitate was removed by centrifugation (30 000 × *g* for 20 minutes at 4 °C) and NaCl was added to the supernatant to a final concentration of 30 % (w/v). The mixture was left to stir gently at room temperature for an additional 20 minutes. The protein solution was centrifuged for one hour (30 000 × *g* at 4 °C). Supernatant was discarded and an appropriate amount of KH_2PO_4 (pH 6.5) buffer was added to solubilise the protein precipitate. The pellet was left to equilibrate with the buffer on ice for one hour in a 4 °C environment.

Size-Exclusion Chromatography

Size-exclusion chromatography of recombinant β-Lg variants was performed using an Äkta Explorer 10 S FPLC and a Superdex 75 10/300 size-exclusion column (GE Healthcare; USA) (10 mm × 30 cm; bed volume 24 mL). Before gel filtration the protein sample was centrifuged at 13 000 × *g* (4 °C for 30 minutes) and filtered to remove any particulate matter. It was isocratically eluted using filtered 50 mM KH_2PO_4 buffer (pH 6.5). The flow rate was maintained at 0.5 mL min^{-1} and 1.2 mL fractions were collected for the duration of the analysis. The presence of protein was monitored by observing the change in absorbance at 280 nm. The molecular masses of eluting proteins were estimated using SDS-PAGE (Section 2.3.2).

2.4 Nuclear Magnetic Resonance (NMR) Methods

2.4.1 Theory of 'Model-Free' Analysis of ^{15}N Backbone Dynamics of Proteins

The following theory of the model-free analysis of ^{15}N backbone dynamics of proteins is based on theory published by Rule and Hitchens (2006).

The Experimental Parameters: ^{15}N R_1, R_2 and ss-NOE

To quantify the dynamics of the protein backbone using ^{15}N NMR spectroscopy, three experimental parameters are usually measured. These are the **^{15}N R_1 (longitudinal relaxation rate), R_2 (tranverse relaxation rate)** and **ss-NOE (steady state-NOE)**. At equilibrium the ^{15}N magnetisation lies along the *z* axis of the spectrometer's magnet. By applying a RF pulse at the absorption frequency of ^{15}N, the ^{15}N magnetisation is tipped into the *x*-*y* plane. R_1 is the rate at which ^{15}N magnetisation realigns along the *z* axis, whereas R_2 is the rate at which magnetisation is loses coherance from the *x*-*y* plane. Both these rates are measured using HSQC pulse sequences that have been modified by holding the N magnetisation along the -*z* axis (for R_1), or in the *x*-*y* plane (for R_2), for varying amounts of time. For macromolecules, such as proteins, the value of R_2 is greater than R_1.

Unlike the first two parameters, the ss-NOE is a dimensionless number rather than a rate. However, its size does rely on a dynamical process. The ss-NOE is the relative change in intensity of the ^{15}N signal measured with and without irradiation of the attached H (amide proton). Irradiation of the NHs destroys the H *z* magnetisation. However, owing to the spatial closeness of the attached nitrogen the N magnetisation is affected to some degree as well. If the NHs are irradiated for a long enough period (typically ~3 seconds) no further changes in the N *z* magnetisation takes place. This is known as the steady state.

^{15}N relaxation is assumed to be due to magnetic fields stemming from fluctuations of the NH dipole-dipole interaction and the ^{15}N chemical shift anisotropy. The chemical shift is a measure of electron density, which is a 3D typically asymmetric property. In solution, because of molecular tumbling, the chemical shift is seen as a single averaged value. However, because of its 3D nature this characteristic still affects the relaxation parameters.

The three measured parameters are given by:

$$R_1(\omega_H, \omega_N) = \left(\frac{d^2}{4}\right)\left(J(\omega_H - \omega_N) + 3J(\omega_N)\right) + 6J(\omega_H + \omega_N) + \frac{1}{3}{\omega_N}^2\Delta\sigma^2 J(\omega_N) \quad (2)$$

$$R_2(\omega_H, \omega_N) = \left(\frac{d^2}{4}\right)\left[\left(J(\omega_H - \omega_N) + 3J(\omega_N)\right) + 6J(\omega_H + \omega_N) + 6J(\omega_H) + 6J(0)\right] + \left(\frac{1}{18}\right){\omega_N}^2\Delta\sigma^2\left(3J(\omega_N) + 4J(0)\right) + R_{ex} \quad (3)$$

$$NOE(\omega_H, \omega_N) = 1 + \left(\frac{\gamma_H}{\gamma_N}\right)\frac{d^2\left(6J(\omega_H+\omega_N) - J(\omega_H-\omega_N)\right)}{4R_1(\omega_H,\omega_N)} \quad (4)$$

where $d = (\mu_0 h\gamma_H\gamma_N)/(r^3 8\pi^2)$. μ_0 is the permeability of free space ($4\pi \times 10^{-7}$ Js^2C^{-2} m^{-1}), h is Planck's constant (6.626×10^{-34} Js), γ_H (26.75×10^7 $radT^{-1}$ s^{-1}) and γ_N (-2.712×10^7 $radT^{-1}$ s^{-1}) are gyromagnetic ratios of hydrogen and nitrogen-15 nuclei respectively, r is the proton-nitrogen bond length (1.02×10^{-10} m), ω_H and ω_N are the Larmor frequencies for proton and nitrogen, respectively, and $\Delta\sigma$ is the chemical shift anisotropy of the N-H bond (-172 ppm (Kroenke *et al.*, 1999)).

The Spectral Density Function: the Link between the Molecular Dynamics and the Experimental Parameters

Though equations 2-4 are rather large, in essence they illustrate that each relaxation parameter relies on the amplitude of a function $J(\omega)$ at a range of ω (frequencies). ***J(ω)* is the spectral density function (SDF)**. Each backbone ^{15}N will possess its own SDF, dictated by the type of movement it is experiencing (i.e. its dynamics). As a consequence, the SDF establishes the connection between the parameters that quantify the dynamical behaviour, which is what is desired, and the experimental parameters, which are able to be measured.

Although measurements of the amplitude of $J(\omega)$ at many frequencies are preferable, assessment of equations 2-4 shows that for relaxation parameters derived at **only one field strength** there is a limit of five frequencies in which amplitudes can be sampled: 0, ω_N, $(\omega_H-\omega_N)$, ω_H and $(\omega_H + \omega_N)$ i.e. 0 %, ~10 %, ~90 %, 100 % and ~110 % of the spectrometer's ^1H operating frequency. A distinguishing feature of R_2 is that it also

relies on the value of $J(\omega)$ at zero frequency, *J(0),* as well. As a result, on a 700 MHz spectrometer, the function $J(\omega)$ is able to be sampled for each ^{15}NH at 0, 71, 629, 700 and 771 MHz.

To ensure a more comprehensive explanation of R_2 we have to allow for the likelihood that the ^{15}N nucleus may be switching between two distinct chemical shift environments. This is provided for by the additional term R_{ex}. This term is typically zero for relatively rigid NH bonds, but can be significant for ^{15}N positioned in flexible loops.

The next step is to provide a logical formulation for the SDF. If the protein molecule is assumed to be tumbling isotropically with no internal motion, then the SDF can be determined by the parameter $\boldsymbol{\tau_m}$, the **molecular correlation time.** This is approximately the time for the molecule to rotate through 1 radian. This simple SDF is given by:

$$J_m(\omega, \tau_m) = \frac{2}{5}\left(\frac{\tau_m}{1+\omega^2{\tau_m}^2}\right) \quad (5)$$

The SDF for two cases is shown below (Figure 2.1). The red trace is for a typical protein with a τ_m equal to 9 ns, and the blue trace is for a much smaller, faster rotating molecule with τ_m of 9 ps.

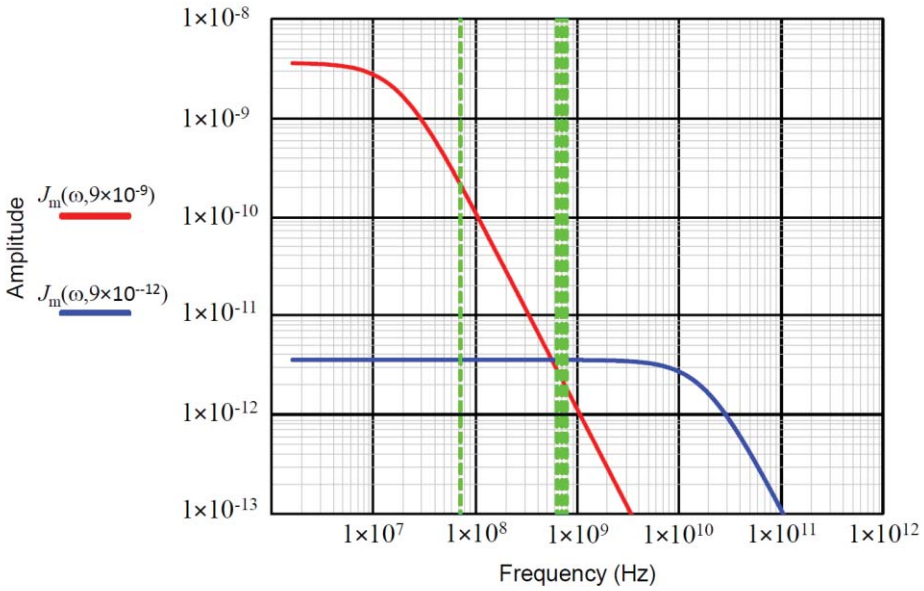


Figure 2.1 SDFs for Rigid Spheres.

The green vertical lines correspond to the frequencies that relaxation parameters depend on using a 700 MHz spectrometer. R_2 is also influenced by the value of the functions at 0 Hz.

The protein molecule (red) has a large amplitude at J(0), but is decreasing at other frequencies, to which the relaxation parameters are sensitive. The molecule that is tumbling faster (blue) has a flatter, lower SDF, with not as much potential to promote relaxation at a particular frequency, but can influence relaxation over a wider frequency range. In a globular protein molecule the backbone ^{15}N can experience both global molecular tumbling and faster internal motions. Therefore, the SDF of the backbone ^{15}N in a more realistic representation, with global tumbling and internal dynamical movements, looks like a combination of both the SDFs shown in Figure 2.1.

Theoretical Description of the SDF

In order to describe the more complete SDF the **order parameter** for the NH bond, $\boldsymbol{S^2}$ is introduced:

$$S^2(\alpha) = \left[cos(\alpha)\frac{(1+COS(\alpha))}{2}\right]^2 \quad (6)$$

where α is the half angle of the cone which the N-H bond is expected to librate. E.g. an NH bond librating in a cone with $\alpha = 30^\circ$ gives $S^2 = 0.65$.

S^2 decreases as the NH bond motion becomes more flexible i.e. as the cone encompassing the movement gets bigger. This is illustrated in the following plot (Figure 2.2).

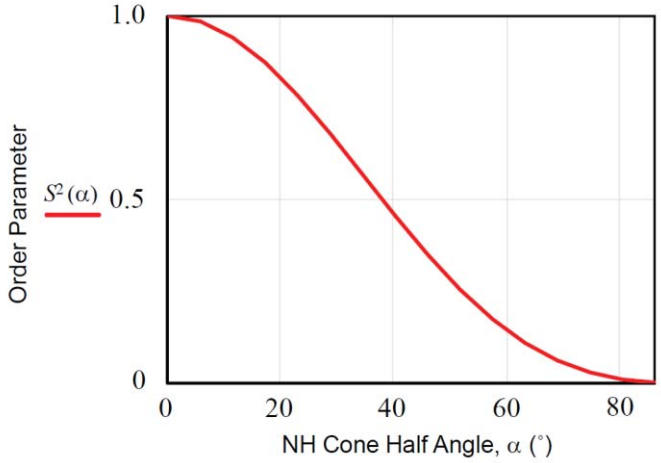


Figure 2.2 Plot Tracking Changes in S^2 as α is Increased.

Libration of a backbone amide within a cone becomes more rigid as the NH bond orientation becomes more constrained. For example at a cone half angle of 30 degrees the order parameter is approximately 0.65, at 15 degrees S^2 is about 0.9 and when fully restrained S^2 is equal to 1.

τ_e defines the correlation time of the bond librating through one radian. The overall correlation time of the bond τ depends on both τ_e and the global correlation time τ_m. Assuming the two motions are independent τ_e is described by:

$$\tau(\tau_m, \tau_e) = \left[\left(\frac{1}{\tau_m}\right) + \left(\frac{1}{\tau_e}\right)\right]^{-1} \quad (7)$$

which is the overall correlation time for the NH bond.

Lipari and Szabo combined these parameters to derive the following SDF for a backbone NH bond, assuming that the internal dynamics and global motions of protein are independent from one another and providing that the molecule is rotating isotropically (i.e. tumbling like a sphere) (Lipari & Szabo, 1982a, Lipari & Szabo, 1982b):

$$J(\tau_m, \tau_e, \omega, S^2) = \left[S^2 \frac{\tau_m}{(1+\omega^2 {\tau_m}^2)} + (1 - S^2) \frac{\tau(\tau_m, \tau_e)}{(1+\omega^2 \tau(\tau_m, \tau_e)^2)}\right] \frac{2}{5} \quad (8)$$

Equation 8 is called 'model-free' because it does not prescribe the mechanism of the NH bond movement, just the amplitude of motion within the cone (S^2) and the rate of the movement (τ_e). Despite the fact that the Lipari Szabo formalism is only an approximation of the *actual* SDF of a particular NH, which could be a consequence of more complex movements, in many cases it is adequate to account for the experimental data.

The Lipari-Szabo spectral density for different combinations of S^2 and τ_e for a molecule with a τ_m typical of a medium sized protein (9 ns) is plotted in Figure 2.3.

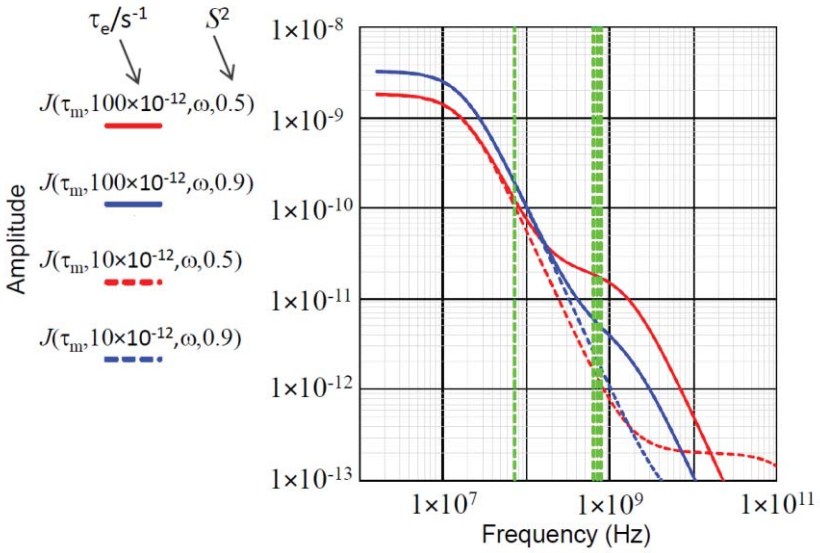


Figure 2.3 Lipari-Szabo SDF.

In this log-log scaled plot the frequencies at which the SDF is sampled are shown by the green vertical lines. R_2 is also sensitive to motions at zero frequency.

Basics of ^{15}N Model-Free Fitting

For a more stringent evaluation of internal motions of each backbone amide all three relaxation parameters need to be considered in concert. A slightly modified version of the SDF is used as the basis of model-free fitting of ^{15}N relaxation data to make these assessments. The basics of the method are described below.

Estimating the Global Correlation Time τ_m

The first thing that is required to estimate the internal dynamics of each backbone amide is an estimate for τ_m. This can be derived from R_2 and R_1 measurement from ^{15}N with relatively rigid NH bonds using the following equation:

$$\tau_m estimate = \left(\frac{1}{\omega_N}\right)\sqrt{\left(6\frac{R_2}{R_1}\right) - 17} \quad (9)$$

Therefore, every residue with respective R_1, R_2 and NOE data, and which is not excluded using the following criteria, is used to calculate a value for τ_m.

1. Residue data are excluded from the τ_m estimate if the respective R_1/R_2 ratios are not within one standard deviation from the mean. This is necessary to minimise the likelihood of using residues with R_{ex} contributions (increased R_2) or long τ_e values (increased R_1).
2. To prevent the selection of residues with low S^2 and relatively long τ_e values,

residue data are excluded from the τ_m estimate if NOE values are less than 0.7 at a field strength of 700 MHz.

Every residue for which there is appropriate relaxation data can be used to calculate τ_m. The average of these is used for the initial estimate of τ_m in the model-free fitting.

Fitting R_1, R_2 and the NOE Using the Model-Free SDF

Up till this point the simple model for the SDF expressed has four parameters, τ_m, S^2, τ_e and R_{ex}, to fit the experimental R_1, R_2 and NOE data. Hence the values for these parameters could be guessed, and then the expected values for R_1, R_2 and the NOE could be calculated and compared with the experimental values, changing the dynamic parameters until a close match with the experimentally determined parameters is obtained.

The following shows what would be expected for some typical values of τ_e and S^2 (it is assumed that τ_m is known), using the theory in Section 2.4.1:

(a) Long τ_e, high S^2

$\tau_m = 9 \times 10^{-9}$ s $\quad \tau_e = 400 \times 10^{-12}$ s $\quad S^2 = 0.85$

$R_1(\tau_m, \tau_e, S^2) = 1.244$ s^{-1}

$R_2(\tau_m, \tau_e, S^2, 0) = 16.702$ s^{-1}

NOE$(\tau_m, \tau_e, S^2) = 0.678$

(b) Short τ_e, high S^2, NOE increased, R_1 and R_2 relatively unaffected compared to (a)

$\tau_m = 9 \times 10^{-9}$ s $\quad \tau_e = 10 \times 10^{-12}$ s $\quad S^2 = 0.85$

$R_1(\tau_m, \tau_e, S^2) = 1.072$ s^{-1}

$R_2(\tau_m, \tau_e, S^2, 0) = 16.479$ s^{-1}

NOE$(\tau_m, \tau_e, S^2) = 0.878$

(c) Short τ_e, low S^2, R_1 and R_2 reduced, NOE increased compared to (a)

$\tau_m = 9 \times 10^{-9}$ s $\quad \tau_e = 10 \times 10^{-12}$ s $\quad S^2 = 0.65$

$R_1(\tau_m, \tau_e, S^2) = 0.834$ s^{-1}

$R_2(\tau_m, \tau_e, S^2, 0) = 12.617$ s^{-1}

NOE$(\tau_m, \tau_e, S^2) = 0.808$

The best estimates for R_1 and R_2 can be used to compute a new estimate for τ_m. If this doesn't agree with the initial estimate then a round of fitting can be performed and the process repeated until a self consistent solution is achieved.

Extended Model-Free Formalism

For some backbone amides, motions are more complex and hence the description of the internal motion in terms of a single order parameter and the associated correlation time does not render a sufficiently good match between the calculated and experimental R_1, R_2 and NOE values. The extended model-free spectral density function, J_{EMF}, resolves this to some extent by assuming that the NH bond can librate in a cone inside another cone, i.e. two motions with variable amplitudes and time-scales. Therefore, S^2 is broken down into S^2_s and S^2_f where $S^2 = S^2_f \times S^2_s$.

$$J_{EMF}(\tau_{m.}\tau_e, \omega, {S^2}_f, {S^2}_s) = \frac{2}{5}{S^2}_f \left[{S^2}_s \frac{\tau_m}{(1+\omega^2{\tau_m}^2)} + (1 - {S^2}_s)\frac{\tau(\tau_m\tau_e)}{(1+\omega^2\tau(\tau_m\tau_e)^2)}\right] \qquad (10)$$

S^2_s corresponds to S^2 as defined by the simple model-free formalism. S^2_f describes the amplitude of a faster motion superimposed on that described by S^2_s, i.e. the NH bond is assumed to oscillate in a cone, which in turn librates inside another cone. For most residues S^2_f will be equal to 1, i.e. the motion is restrained to a single cone in accordance with the simple model-free formalism. The correlation time associated with S^2_f is unable to be measured as it is too short.

One final consideration is that in some instances R_2 can be affected by the pressure of conformational exchange (i.e. the NH is flipping between separate shift environments such as the residues of a loop switching between two differing configurations. This is accounted for by adding an extra term, R_{ex} to the expression for R_2.

The above formulation allows for the possibility of five models to describe the motion of the NH bond.

Model 1: S^2_s

Model 2: S^2_s and τ_e

Model 3: S^2_s and R_{ex}

Model 4: S^2_s, τ_e and R_{ex}

Model 5: S^2_s, S^2_f and τ_e

In this study the iterative fitting of the dynamics data was performed with the program MODELFREE. In this program fitting of a residue for a particular model proceeds as before by varying its parameters to give them best fit to the experimental R_1, R_2 and NOE. The picking of the most appropriate model is based on statistical tests whereby the more complex models are only used if required. Typical fitting usually necessitates several rounds of local fitting (where τ_m is allowed to vary for each residue) and global fitting where τ_m is fixed and stops when a further iteration causes no further change in τ_m.

When many residues appear to require these higher models it can be an indication that there is a more systematic error in the experimental data or that the assumption of isotropic tumbling is invalid (i.e. the protein is not tumbling like a sphere).

Isotopic β-Lg Used for NMR Spectroscopy Analyses

In addition to the recombinant ^{15}N-labelled β-Lg C, which has been the main focus of this thesis, recombinant A and B variant samples were used. These latter samples were prepared by Dr. Komala Ponniah, Massey University. Recombinant ^{15}N-labelled β-Lg C and ^{13}C-^{15}N labelled β-Lg C samples were prepared as described previously (Section 2.3). The recombinant β-Lg variants used in these studies differed by one additional methionine residue at the N-terminus, when compared to each respective wild-type bovine β-Lg variant.

All NMR experiments were recorded on a Bruker Avance 700 MHz NMR spectrometer (Bruker Biospin GmbH, Rheinstetten, GER) equipped with a cryoprobe. NMR spectra were processed with standard parameters using Bruker TopSpin 2.1 (Bruker Biospin GmbH, Rheinstetten, GER) and analysed using CCPNMR ANALYSIS.2.1 (Vranken *et al.*, 2005) downloaded from the website http://www.ccpn.ac.uk/ccpn/software/downloads-v2/. The ^{1}H, ^{15}N and ^{13}C chemical shifts were referenced to the water signal.

2.4.2 Assigning the Backbone of β-Lg C

Introduction to Protein Backbone Assignments

Before investigating ^{15}N protein backbone dynamics, it is essential to assign the chemical shifts of the peaks in the ^{15}N,^{1}H-HSQC spectrum to the appropriate amino acid residue. Although assignments for β-Lg already exist (Uhrínová *et al*., 1998, Kuwata *et al*., 1999) it was found there was sufficient ambiguity of peak identities in crowded regions of the ^{15}N,^{1}H-HSQC to warrant repeating the backbone assignment. This was achieved sequentially *via* a series of three-dimensional NMR experiments that link resonances

associated with backbone amides to resonances associated with the backbone carbons. These experiments require double labelled (^{15}N and ^{13}C) samples. Here the CBCANH and CBCA(CO)NH experiments were used. These are described briefly below.

The **3D CBCANH experiment** bestows a 3D spectrum in which ^{1}H, ^{15}N and ^{13}Cα/^{13}Cβ chemical shifts are viewed in three independent dimensions. It is designed with the purpose of associating the ^{1}H and ^{15}N amide resonances with those of the intra- (residue *i*) and inter-residue (residue *i*-1) ^{13}Cα and ^{13}Cβ resonances. The second experiment, the **3D CBCA(CO)NH,** also gives the chemical shift information for the amide peak (^{15}N,^{1}H) of each residue *i* in the HSQC dimension *but only* provides information regarding the resonances of the ^{13}Cα and ^{13}Cβ carbons of the inter-residue (residue *i*-1). Therefore the information rendered by both CBCANH and CBCA(CO)NH experiments is used in conjunction with one another to assign the resonances for both ^{13}Cα and ^{13}Cβ atoms of residue *i* and residue *i*-1 and then to work progressively down the backbone. Ultimately these 3D-NMR experiments are used to assign the peaks present in the ^{15}N,^{1}H-HSQC spectrum with each residue specific ^{15}N-^{1}H backbone group.

The assignments established in the previous set of experiments can be confirmed with the **HNCO** experiment. In these studies this experiment was used to determine the chemical shifts of the carbonyl residues, which was confirmed with the published assignments (Uhrínová *et al*., 1998)

For some residues it was useful to use a pair of **3D ^{15}N,^{1}H-TOCSY-HSQC** and **^{15}N,^{1}H - NOESY-HSQC** experiments, as these can be used to correlate the backbone amide proton and nitrogen of the intra-residue to the protons of the inter-residue (*i*-1) for confirmation that the appropriate residue had been assigned by identifying the Hα chemical shifts and in most cases the Hβ chemical shifts of the side-chain protons.

Backbone Assignment of β-Lg C

Sequential backbone assignment of the ^{13}C,^{15}N-labelled β-Lg C was achieved using a set of NMR triple resonance experiments: CBCA(CO)NH (Grzesiek & Bax, 1992a), CBCANH (Grzesiek & Bax, 1992b) and HNCO (Grzesiek & Bax, 1992c, Kay *et al.*, 1990, Muhandiram & Kay, 1994). β-Lg C backbone chemical shifts were also verified and assigned by means of ^{15}N,^{1}H-HSQC (Bodenhausen & Ruben, 1980), ^{15}N,^{1}H-TOCSY-HSQC (Marion *et al.*, 1989a) and ^{15}N,^{1}H-NOESY-HSQC (Marion *et al*., 1989a, Marion *et al.*, 1989b) NMR experiments.

All carbon backbone experiments were recorded at 305 K. For carbon backbone assignment experiments of the recombinant β-Lg C variant, NMR spectra were acquired on a 300 μL sample containing 5 % $^{2}H_2O$, in a Shigemi NMR tube (Shigemi, Allison Park, PA). The β-Lg C sample contained 0.3 mM of uniformly ^{13}C-^{15}N labelled protein in 50 mM potassium phosphate buffer (pH 2.6). Under these conditions, β-Lg C is monomeric.

CBCANH, CBCA(CO)NH and HNCO triple resonance spectra were acquired using 32, 12 and eight transients per FID, respectively, employing echo-antiecho (^{13}C) (Kay *et al.*, 1992) and States-time-proportional phase incrementation (States TPPI) (^{15}N) (Marion & Wüthrich, 1983, Marion *et al.*, 1989c). Experimental acquisition parameters are listed in Table 2.8.

Triple Resonance Experiment	1D Dimension	Number of Points	Sweep Width (Hz)
CBCANH	^{1}H	1024	9842
	^{15}N	128	1727
	^{13}C	512	13192
CBCA(CO)NH	^{1}H	2048	9765
	^{15}N	128	1727
	^{13}C	512	13192
HNCO	^{1}H	2048	9765
	^{15}N	256	1727
	^{13}C	512	2817

Table 2.8 Acquisition parameters for CBCANH, CBCA(CO)NH and HNCO 3D experiments. Number of points and sweep widths (Hz) are listed for each ^{1}H, ^{15}N and ^{13}C dimension corresponding to the appropriate triple resonance experiment.

2.4.3 Backbone Verification and Assignment of β-Lg A, B and C Monomeric Variants using 3D ^{15}N,^{1}H-TOCSY-HSQC and 3D ^{15}N, ^{1}H-NOESY-HSQC Experiments

β-Lg A, B and C ^{1}H and ^{15}N chemical shifts were verified and assigned by means of 3D ^{15}N,^{1}H-TOCSY-HSQC and 3D ^{15}N,^{1}H-NOESY-HSQC experiments. Spectra were acquired at 305 K on separate 300 μL β-Lg A, B and C variant samples in Shigemi NMR tubes. Each sample contained 1.0 mM of ^{15}N-labelled β-Lg (A, B or C) protein in 50 mM potassium phosphate buffer (pH 2.6, 5 % $^{2}H_2O$). Under these conditions, the β-Lg A, B and C protein variants are monomeric.

All 3D ^{15}N,^{1}H-TOCSY-HSQC and 3D ^{15}N,^{1}H-NOESY-HSQC spectra were recorded with 12 and 32 transients respectively, 2048 × 64 × 128 (^{1}H × ^{15}N × ^{1}H) complex points employing echo-antiecho (^{15}N) (Kay *et al*., 1992) and States-time-proportional phase incrementation (States TPPI) (^{1}H) (Marion & Wüthrich, 1983, Marion *et al*., 1989c), and sweep widths of 8389.26 Hz (^{1}H) × 1717.33 Hz (^{15}N) × 8403.04 Hz (^{1}H). ^{15}N,^{1}H-TOCSY-HSQC triple-resonance experiments were recorded with a mixing time of 60 ms and a relaxation delay of 1.2 s. 3D ^{15}N,^{1}H-NOESY-HSQC spectra were obtained with a mixing time of 100 ms.

2.4.4 Assigning Backbone β-Lg C ^{1}H-^{15}N Chemical Shifts at 313 K and 320 K *via* a ^{15}N,^{1}H-HSQC Temperature Series

The β-Lg variant C ^{1}H and ^{15}N chemical shifts, at 313 K and 320 K, were resolved by means of a ^{15}N,^{1}H-HSQC temperature series, recorded between 306 and 320 K, using 2 K incremental steps. Experiments were recorded on a 300 μL recombinant β-Lg C sample containing 1.0 mM of ^{15}N-labelled β-Lg protein in 50 mM potassium phosphate buffer (pH 2.6, 5 % $^{2}H_2O$ (v/v)).

2.4.5 ^{15}N,^{1}H NMR Relaxation Experiments Used to Probe Dynamics β-Lg

Standard R_1 (spin-lattice relaxation rate), R_2 (spin-spin relaxation rate), and steady-state (ss)-NOE, 2D heteronuclear ^{1}H and ^{15}N HSQC relaxation experiments were recorded on separate 300 μL recombinant β-Lg A, B and C samples in Shigemi NMR tubes. For the R_1 experiments, NH decoupling was applied during relaxation interval using hard 180° pulses in conjunction with selective flip back pulses on the water resonance to keep it along the length of the +*z* axis. These pulses were applied every 2.5 ms. For the R_2 experiment the CPMG pulse train consisted of 120 μs 180° ^{15}N pulses applied every 900 μs, and 180° pulses were applied every 7.2 ms. All samples contained 1.0 mM of ^{15}N-labelled β-Lg protein in 50 mM potassium phosphate buffer (pH 2.6, 5 % $^{2}H_2O$ (v/v)). Under these conditions, the β-Lg A, B and C protein variants were monomeric.

β-Lg A, B and C ^{1}H-^{15}N relaxation experiments were acquired at 305 K, to analyse and compare the dynamics of the three different β-Lg variants at one temperature. Relaxation experiments of the β-Lg C variant were recorded at 305 K, 313 K and 320 K, to analyse and compare the backbone dynamics of one variant at three different temperatures. A separate β-Lg C R_1 experiment was recorded at 305 K as a control, after the 320 K NMR relaxation data set was acquired. This was necessary to check the integrity of β-Lg C

over the time period and temperature range that the β-Lg C NMR relaxation experiments had been conducted. Highly similar R_1 values showed that the β-Lg C sample had not degraded over the time course of the experiments and at the temperatures sampled.

^{15}N R_1 and R_2 rate measurements (Kay *et al.*, 1989) were acquired with the following relaxation delays: R_1, 51.1, 80.1, 151.1, 215.1, 401.1, 801.1, 1401.1 and 2001.1 ms; R_2, 16.32, 32.64, 48.96, 81.60, 97.92, 114.24, 130.56, 163.20 and 195.84 ms. To estimate noise levels, duplicate spectra were recorded for R = 51.1, 151.1 and 801.1 ms (R_1 spectra) and R = 32.64 and 97.92 ms (R_2 spectra). R_1 and R_2 two-dimensional spectra were attained with the following settings: 2048 × 256 complex points (^{1}H × ^{15}N) and spectral widths of 11261.3 Hz (^{1}H) and 1717.3 Hz (^{15}N), respectively. R_1 and R_2 spectra were recorded with 16 and 8 transients per increment, respectively.

Ss-NOE ^{1}H-^{15}N HSQC spectra (Kay *et al.*, 1989) were attained using 48 transients, 2048 × 256 complex points (^{1}H × ^{15}N) employing echo-antiecho (^{15}N) (Kay *et al*., 1992), and spectral widths of 11261.3 Hz (^{1}H) and 1717.3 Hz (^{15}N). The steady-state-^{1}H-^{15}N NOE values were determined from spectra recorded in the presence and absence of a proton pre-saturation period. A reference experiment was recorded with a 5 second relaxation delay. The NOE experiment was recorded with saturation of the ^{1}H signal for the last 3 seconds of the 5 second relaxation delay.

2.4.6 Data Processing and Analysis

Both R_1 and R_2 relaxation rates were determined from the decay of the intensity of each assigned ^{1}H-^{15}N cross peak for each respective experimental series (Section 2.4.5). The intensities of the peaks in the 2D spectra were described by peak heights as determined by the peak picking software program ANALYSIS.2.1 (Vranken *et al.*, 2005). Values were derived, using ANALYSIS.2.1 software, by fitting the measured peak heights obtained by experimental data to an exponential function:

$$I(t) = I_0 \exp\left(-tR_{1,2}\right) \quad (10)$$

where $I(t)$ is the intensity after a delay time of t and I_0 is the intensity at time t = 0.

The ss- NOE intensities were calculated in ANALYSIS.2.1 from the following ratio:

$$\frac{I_{sat}}{I_{ref}} \tag{11}$$

where I_{sat} and I_{ref} are the heights of an assigned resonance measured in the NOE spectra, with and without proton saturation, respectively.

The uncertainty for NOE intensities was determined from the noise level of the spectra.

2.4.7 Model-Free Analysis

Initial estimates for the molecular rotational correlation times, τ_m, were derived at each temperature from R_2/R_1 ratios using the `r2r1diffusion` program (Tjandra *et al.*, 1995) provided on A. G. Palmer's website (http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer/software.html). These estimates were subsequently refined by the model-free fitting procedure.

The local motion parameters of the amide ^{15}N relaxation rates were quantified by means of the Lipari-Szabo (1982a, 1982b) analysis with the extensions of Clore *et al.* (1990a, 1990b) using the MODELFREE 4.15 software suite (Mandel *et al.*, 1995, Palmer *et al.*, 1991). Model-free analysis expresses quantities directly associated to the spatial constraints and time-scales of the molecular motions.

2.4.8 Estimation of Backbone Flexibility Using Tertiary Structure

The Zhang and Brűschweiler method (2002) estimates S^2 order parameters of N-H bonds of the protein's backbone using pdb coordinates from high-resolution X-ray or NMR protein structures. Using this method, flexibility is presumed to correspond to how grid-locked the N-H bond is by the surrounding atoms, i.e. the local density of atoms in a protein determines its local dynamics. The order parameter, S^2 for residue i, is calculated as a scaled summation of the distances from HN_i and CO_{i-1} to all the protein's heavy atoms. Calculation of S^2 is performed by the program *s2predict.py*, which is available from the authors' website (http://spin.magnet.fsu.edu/software/S2/s2predict.htm). The values for each of the structures were calculated then averaged.

2.4.9 Calculation of Protein Flexibility Using the Random Coil Index (RCI)

The Random Coil Index (RCI) based dynamics is a simple, NMR spectroscopy chemical shift based method capable of estimating S^2 using $^1H/^{15}N/^{13}C$ backbone chemical shifts.

Since protein backbone dynamics appear to correlate with structure, and structure also correlates with backbone chemical shifts (Berjanskii & Wishart, 2005), it follows that a relationship may exist between dynamics and chemical shifts. This relationship was investigated by Wishart *et al.* with 18 proteins, varying in length and conformation, to find a correlation between chemical shift index and structure. They then investigated a subset of 14 well-resolved proteins leading to the establishment of an empirical relationship between the chemical shifts and S^2 dynamics (Berjanskii & Wishart, 2005). It has been implemented on the Wishart RCI webserver (http://wishart.biology.ualberta.ca/rci).

The Wishart RCI web server predicts protein flexibility by calculating the so-called 'Random Coil Index' (where the difference of the observed chemical shifts for a residue and its neighbours are compared with those expected for the same amino acids in a random coil) from an input file of backbone chemical shifts (C_α, CO, C_β, N, H_α, N_H) to predict values of S^2. Since the method does not require NMR measurements other than the backbone atom shifts, the time-consuming collection and analyses of NMR spectroscopy relaxation experiments (NOEs, R_1 and R_2) are, at least in principle, not needed to map the protein's flexibility.

3

Results and Discussion

3.1 Rationalised Site-Directed Mutagenesis of the BLG Gene

Site-specific DNA base-pair substitution of the [pETDuet-DsbC-BLG B] construct, created by Ponniah *et al*. (2010), was required to generate the expression vector necessary for the over-expression of bovine ^{15}N and/or ^{13}C β-Lg variant C in *E. coli.*. Even though wild-type β-Lg C can be readily isolated from fresh cows' milk (Manderson *et al.*, 1997), heterologous expression was required to incorporate the isotopic label/s required for residue-specific NMR spectroscopy dynamical investigations. The expression of the recombinant C variant had not been published anywhere until it had been reported by Ponniah *et al*. (2010), as a consequence of this study.

3.1.1 PCR Site-Directed Mutagenesis

Site-directed mutagenesis was performed by means of PCR, using a whole-plasmid mutagenic and amplification system developed by Stratagene, described in Section 2.2.1. This method was chosen for its rapidness as positive transformants could be potentially selected the day following PCR site-directed mutagenesis.

The β-Lg B expression plasmid, [pETDuet-DsbC-BLG B], was mutated and amplified in its entirety using two complementary mutagenic primers targeted towards the region of the proposed substitution site in the BLG B gene. The forward and reverse primers, 35 oligonucleotides in length, complemented the BLG B gene at positions 160 to 194, with the exception of the rationalised base-pair mismatch at position 177. This mismatch was required to produce the Gln59His substitution upon protein translation (Table 2.6), which is identical to the polymorphism in variant C.

3.1.2 Verifying Plasmid Amplification *via* DNA Agarose Gel Electrophoresis

DNA agarose gel electrophoresis (Section 2.2.6) was used to evaluate the success of PCR amplification (Section 2.2.8) of the entire [pETDuet-DsbC-BLG C] expression vector and used to estimate the size of the linear amplicon prior to transformation into TOP10 *E. coli* cells. The PCR product showed an electrophoretic mobility on an agarose gel corresponding to about 6400 bp (Figure 3.1), the approximate size of the expression plasmid (6370 bp), indicating that the plasmid had been successfully amplified.

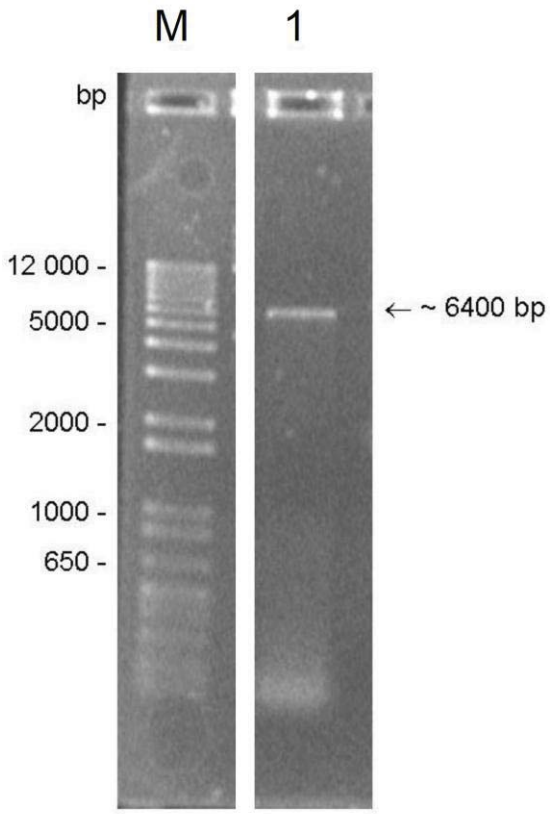


Figure 3.1 Amplicon from PCR Site-Directed Mutagenesis.

The site-directed mutagenesis of the pETDuet-Dsbc-BLG B construct targeting the BLG B gene insert was used as template DNA to create the BLG C gene variant. DNA was separated by agarose gel electrophoresis. Lane M: 1 kb+ ladder as listed in Appendix A.4; Lane 1: PCR amplicon following PCR site-directed mutagenesis of the pETDuet-Dsbc-BLG B template construct.

3.1.3 Transformation of Amplicon into *E. coli* Top10 Hosts

Following DPN1 digestion of methylated and hemi-methylated parental DNA template in the PCR reaction mixture, the amplicon was transformed into *E. coli* Top10 competent cells, where the cells endogenous machinery is able to 'circularise' the construct by ligating the 5' and 3' ends. The transformation mixture was plated onto sterile LB-Amp petri dishes for the selection of positive transformants. Following overnight growth at 37 °C, positive transformants were isolated by aseptically selecting a colony, which was then transferred into LB-Amp broth and grown aerobically overnight at 37 °C to propagate the plasmid. After cultivation the cells were centrifuged and the constructs were harvested.

3.1.4 DNA Sequencing of the BLG C Open Reading Frame (ORF)

The nucleotide sequence of the BLG open reading frame (ORF) within the purified construct was determined using the Big Dye Terminator method at the Alan Wilson Centre. It confirmed the presence of a nucleotide base substitution (CAG→CAT) at position 177; all other bases were homologous to the BLG B template gene. This confirmed that the revised DNA sequence encoded a protein with a Gln to His residue substitution at position 59, which verified that the mutated gene sequence was in alignment with the gene encoding the bovine β-Lg C variant**.** The pETDuet-DsbC-BLG construct containing the mutated gene was used for the heterologous co-expression of β-Lg C and DsbC isomerase proteins in the *E. coli* Origami B (DE3) strain (Sections 2.3.3 & 3.2.2).

3.2 Expression and Purification of β-Lg Variants

3.2.1 Introduction

To study backbone dynamics of β-Lg C, by means of ^{15}N NMR relaxation studies, it was essential to express and purify the double (^{15}N and ^{13}C) and single (^{15}N) labelled β-Lg C variant for NMR spectroscopy backbone assignments and ^{15}N NMR spectroscopy dynamics respectively. As NMR spectroscopy dynamics studies require large quantities of pure protein, it was important to choose an efficient expression and purification strategy to minimise expenses, particularly when ^{13}C-D-glucose was being exploited as the isotopic carbon source. The method employed was one previously used to successfully produce adequate yields of soluble and correctly folded β-Lg variants A and B in *E. coli* (Ponniah *et al*., 2010). This procedure overcame solubility issues in *E. coli* by co-expressing β-Lg A and B with the chaperone protein DsbC isomerase, which is capable of aiding correct disulfide formation in the oxidising environment of the *E. coli* Origami B (DE3) cytoplasm. It also had the advantage of not being co-purified with significant quantities of carbohydrate as with the case of the *Pichia pastoris* system (Section 1.6.1). In this study the protocol was optimised to attain the yields of those published and an extra polishing step was added to increase purity of the final product.

β-Lg C expression and purification were monitored by reduced SDS-PAGE analyses (Section 2.3.2). During SDS-PAGE β-Lg is expected to migrate monomerically as the interactions between the monomers at the dimeric interface are disrupted by the SDS detergent. The protein ladder used to determine protein size contained a marker with a molecular mass of 18.4 kDa. This marker is in-fact native β-Lg and was used as a positive control to ensure that the protein with an equal mass was being over-expressed and purified.

3.2.2 Expression of Recombinant Isotopically Labelled Bovine β-Lg C

Recombinant β-Lg was successfully produced as a soluble protein by means of an *Escherichia coli* IPTG inducible expression system (Sections 1.6.2 & 2.3.3), as monitored by reduced SDS-PAGE analysis (Figure 3.2). The *E. coli* Origami B (DE3) [pETDuet-DsbC-BLG C] cells overproduced recombinant β-Lg C when expression was induced with IPTG at 25 ˚C, as revealed in Figure 3.2. Reduced SDS-PAGE expression analysis showed the absence of a band in the uninduced whole-cell sample lane (Lane 1)

and the presence of a large band in the IPTG-induced whole-cell sample lane (Lane 2), with an approximate electrophoretic mobility of 18.4 kDa. This was a strong indicator that β-Lg was over-expressed, as the band was thick, had approximately the same molecular mass as the β-Lg marker in Lane M, and was not present pre-induction. The co-expression of DsbC (23 kDa) was also deemed successful as a band with an approximate electrophoretic mobility of 23 kDa was absent prior to induction (Lane 1), but present post induction (Lane 2).

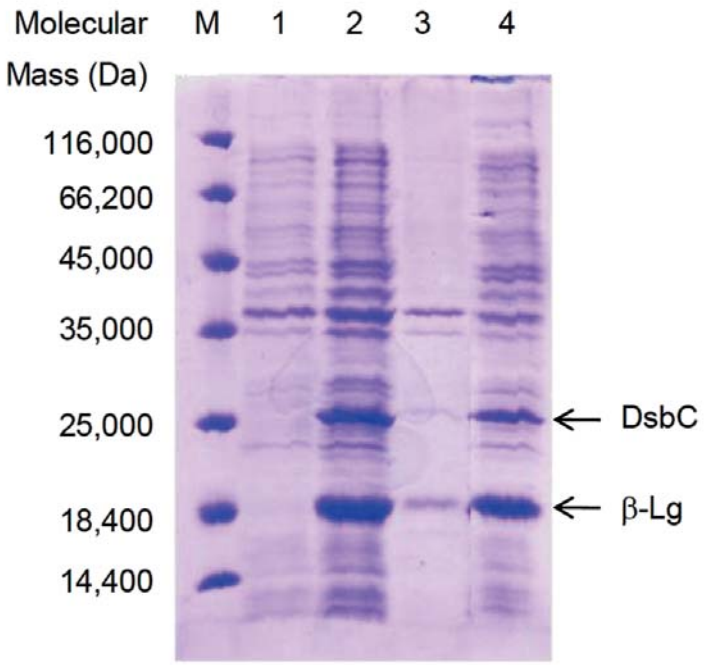


Figure 3.2 Analysis of IPTG-Induced Expression and Solubility of Recombinant β-Lg C. Proteins were separated by SDS-PAGE and stained as described in Materials and Methods (Section 2.3.2). Lanes M: molecular mass markers as in Appendix A.4; Lane 1: uninduced; and Lane 2: IPTG-induced cell extract; Lane 3: insoluble fraction and; Lane 4: soluble fraction of IPTG-induced cell extract.

The presence of a band in all four lanes, which has an approximate electrophoretic mobility of 37 kDa, is expected to be an endogenous *E. coli* protein and not the β-Lg dimer (36.8 kDa) caused by too little SDS detergent in the PAGE system. The rationale behind this is that the band is present in both pre-induction and post-induction whole cell samples (Lanes 1 and 2, respectively), whereas monomeric β-Lg is only detected in observable amounts after induction with IPTG, with quantities being larger in proportion than the 37 kDa band. The expression profiles of the whole-cell extracts are also similar to those published by Ponniah *et al.* (2010), which had similar banding patterns.

As expression of both the BLG and DsbC genes were under control of the T7 promoter, addition of IPTG triggered expression by displacing the *lac* repressor from the *lacUV5* promoter upstream of the gene encoding T7 RNA polymerase. The T7 RNA polymerase then bound to the T7 promoter upstream of the BLG and DsbC genes to initiate transcription of both genes. The Origami B (DE3) [pETDuet-DsbC-BLG] system provided a tightly regulated method for the over-expression of β-Lg C. Leaky expression of the protein was not observed prior to induction in the whole cell sample as indicated by the absence of a band with an electrophoretic mobility equal to that of the β-Lg monomer (Lane 1).

Native β-Lg and DsbC isomerase are both secretory proteins; however, in this investigation they were both targeted to the cytoplasm to collect as both lacked the signal sequence, which would have otherwise led them to the periplasmic space of the *E. coli* host. SDS-PAGE analysis of the soluble and insoluble fractions of the IPTG-induced whole cells showed that over-expressed β-Lg produced in *E. coli* was over 80 % soluble (Figure 3.2). A small quantity of β-Lg was present in the insoluble fraction, whereas a more substantial amount was present in the soluble fraction, post induction (Lanes 3 and 4 respectively). DsbC isomerase was also over-expressed upon induction and was also primarily found in the soluble cell lysate.

During the initial stages of protein expression, a large number of viable host cells, harbouring the expression construct, was produced. This was achieved by inoculating the cells into rich medium (Luria broth) containing antibiotics for the selection of both plasmid (ampicillin) and the *E. coli* strain (Origami B (DE3)) (kanamycin and tetracycline). The initial culture was left to grow aerobically overnight as this cell strain is typically slow growing (Bessette *et al*., 1999).

The M9 minimal medium used in these studies contained all the salts and trace elements required for Origami B (DE3) to grow but contained no carbon or nitrogen sources. After overnight growth the cells were transferred to minimal media for a period of one and a half hours so that the expression hosts would use any remaining metabolites sourced from residual LB medium. This was important to maximise the incorporation of isotopic label into β-Lg C. The appropriate label (^{13}C and/or ^{15}N) was added to the medium after this period, with the remaining unlabelled metabolites if required, together with ampicillin for plasmid selection, to create uniformly isotopically labelled protein. The antibiotics kanamycin and tetracycline were not added for Origami B cell selection

at this stage in order to create a less pressured environment for the slow growing cells. DsbC isomerase and β-Lg protein co-expression was induced with IPTG when the OD_{600} of the culture reached approximately 0.5. At this stage the cell culture was transferred to a cooler environment to enhance soluble protein production.

To improve the productivity of the microbial expression of recombinant β-Lg C, the volume of inoculum transferred into M9 medium was investigated to assess whether or not the time period between inoculation and induction could be reduced. Initially the time period between cell transfer and IPTG induction was 11 hours. This period was shortened to five hours, by increasing the amount of the inoculum from two percent to ten percent, into M9 minimal medium (v/v), making the purification process more workable. The cell densities of the two inocula, prior to transfer into M9 medium, were essentially the same, being 2.4 at OD_{600}. After induction with IPTG, the cells continued to grow for a period of twenty hours and were then harvested.

Once the cells were separated from the media, they were resuspended in lysis buffer and then lysed by French press. Cellular debris was separated from the soluble fraction by centrifugation. Any remaining particulate matter was removed from the soluble supernatant fraction through filtration, to protect the anion exchange column from damage. Fractions from the pellet and supernatant were then analysed by SDS-PAGE for solubility of β-Lg C (Figure 3.2).

Using this procedure and the Origami B [pETDueT-DsbC-BLG] expression system, IPTG effectively induced the over-expression of both DsbC isomerase and β-Lg C. More than 80 % of over-expressed β-Lg C was soluble.

3.2.3 Purification of Recombinant Isotopically Labelled Bovine β-Lg C

The purification protocol chosen for the isolation of the β-Lg variants was based on a procedure previously established for the purification of native β-Lg whey retentate (Mailliart & Ribadeau-Dumas, 1988), which had been modified by Ponniah *et al.* (2010) for the purification of recombinant β-Lg from the cell lysate of *E. coli*. Purification was achieved by means of ion exchange chromatography, acid and salt precipitation, and size-exclusion chromatography, exploiting different characteristics, particularly the extremely acid-stable nature of the protein.

Optimisation of the protocol was necessary to improve sample purity, reduce expenses and obtain adequate yields required for multidimensional NMR studies. Notably, the final protein yields in this study were identical to those previously described (Ponniah *et al.*, 2010), however published yields were not attained when using the published procedure. Changes in protein concentration before acid precipitation and the addition of a final polishing stage were necessary to attain the yields and purity of the final protein solution, thereby allowing more time for NMR spectroscopy analyses. Characterisation of the protein's backbone dynamics by means of NMR spectroscopy requires six milligrams of protein for one series of relaxation experiments.

The purification process was monitored at each step by reduced SDS-PAGE (15 % acrylamide). SDS-PAGE analyses of protein samples at different stages of the purification process are shown in Figure 3.3. β-Lg has no enzymatic function; therefore protein purification could not be examined specifically through any known activity assays. Nonetheless, total protein was determined at the end of each purification step to monitor protein loss and final yield and therefore was not specific to β-Lg C. The purification table summarising the purification of bovine β-Lg C to sufficient purity for NMR studies is shown in Table 3.1. A standard purification resulted in an average yield of 5.3 mg/L of ^{15}N or ^{13}C-^{15}N labelled Met-β-Lg C of >98 % purity (Figure 3.3; Lane 6).

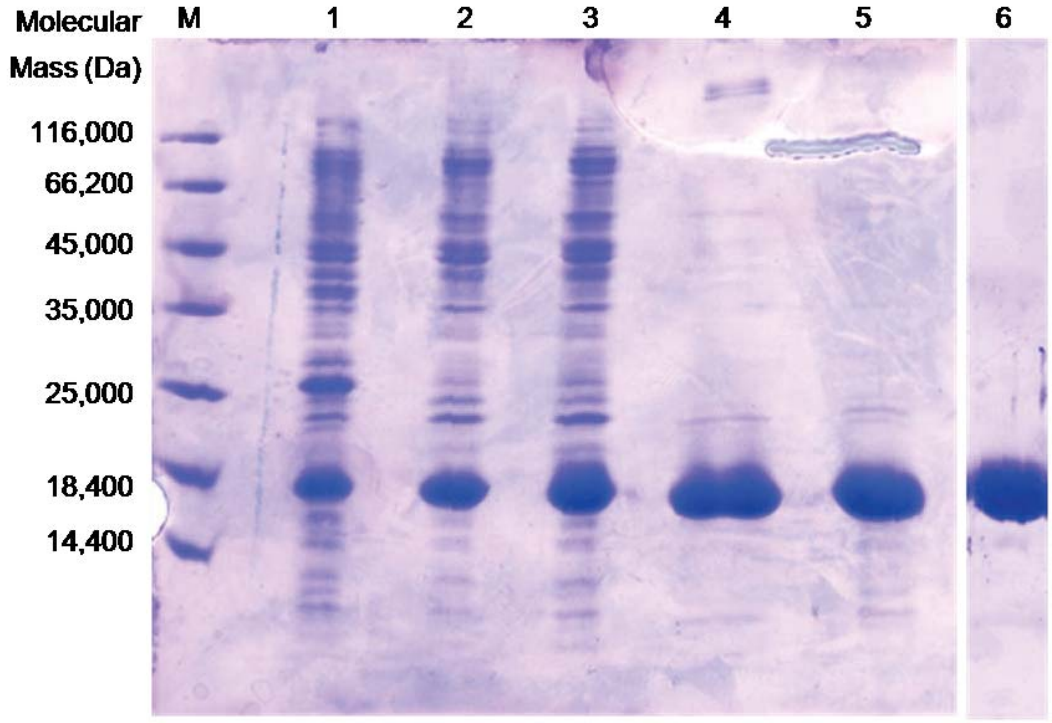


Figure 3.3 Analysis of the Purification of Recombinant Bovine β-Lg C.

Proteins were separated by SDS–PAGE and stained as described in Materials and Methods (Section 2.3.2). Lane M: molecular mass markers shown in Appendix A.4; Lane 1: cell lysate; Lane 2: anion exchange chromatography purified pooled fractions; Lane 3: supernatant fractions from HCl precipitation (pH 2.6); Lane 4: supernatant fractions from NaCl precipitation (7 % w/v); Lane 5: pellet fractions from NaCl precipitation (30 % w/v) after solubilisation in KH_2PO_4 buffer (pH 6.5); Lane 6: size-exclusion chromatography purified pooled fractions. Protein samples (2 mg mL^{-1}) were loaded as 15 μL aliquots.

Fraction	Volume (mL)	Protein Concentration (mg mL^{-1})	Total Protein* (mg)	Yield (%)
Crude lysate	11.00	167.00	1837.00	100
Anion exchange chromatography	22.00	4.65	102.30	5.6
HCl (pH 2.6) supernatant.	27.00	2.00	54.00	2.9
7 % NaCl (w/v) supernatant	35.00	14.50	50.75	2.7
30 % NaCl (w/v) pellet fraction resuspended in KH_2PO_4 (pH 6.5)	3.50	5.76	20.13	1.1
Size-exclusion chromatography	0.65	16.3	10.60	0.58

Table 3.1 Purification of β-Lg C from *E. coli.*

*Total protein from an IPTG-induced 2 L *E. coli* Origami B (DE3) [pETDuet-DsbC-BLG C] culture, grown in minimal media.

In accordance with previous purifications of the A and B variants in our lab, the β-Lg C variant was first partially purified using anion-exchange chromatography (HR 10/30 column packed in-house with SOURCE 15Q resin). The anion-exchange chromatography (AEX) was first used as a capture step from the harvested cell culture fluid.

According to the OD_{280} absorption, fractions were collected and then analysed using SDS-PAGE. The AEX SOURCE 15Q packed column was beneficial as it quickly isolated the bulk of the major contaminant, over-expressed DsbC isomerase (Figure 3.3; Lanes 1 and 2) from the crude cell lysate. DsbC isomerase bound to the column at pH 6.5 and 0 M NaCl, and typically appeared in the 0.05 M NaCl eluted fractions (results not shown). In comparison, β-Lg C also bound to the column (pH 6.5; 0 M NaCl) but was routinely found in the 0.1 M NaCl elutes, separating it from DsbC. The protein fractions containing β-Lg C were pooled, dialysed and concentrated to between 3 and 5 mg mL^{-1}.

The extreme acid-stable nature of β-Lg allowed the protein to be purified effectively by acid (HCl) precipitation coupled to sodium chloride precipitation, as this was a distinguishing characteristic not typically shared by many other proteins. Previous reports of β-Lg purification using the acid/salt precipitation method have typically presented the high acid and 7 % (w/v) NaCl stage as one step (Mailliart & Ribadeau-Dumas, 1988, Ponniah *et al*., 2010), not monitoring purification as a result reducing the pH to 2.6 plus the addition of NaCl separately. In this study, protein purification was monitored just after the centrifugal removal of the precipitates from the acidic protein solution and once again from the soluble fraction after 7 % NaCl (w/v) precipitation.

Pooled and concentrated fractions underwent acid precipitation by reducing the pH to 2.6 with the drop-wise addition 0.1 M HCl. This step, analysed by SDS-PAGE, produced no significant difference in terms of purification, even though there was a significant loss of total protein (~50 %; shown in Table 3.1), including β-Lg. Nonetheless, by reducing the pH to 2.6 in conjunction with increasing sodium chloride concentration to 7 % (w/v), the bulk of the protein contaminants were removed, making this an effective intermediate purification step. Lane 4 in Figure 3.3 showed that in addition to β-Lg, several proteins with comparatively faint Commassie Blue-stained bands were still present in the supernatant, after the combination of these two procedures.

Sodium chloride precipitation and fractionation, commonly referred to as salting-out, is a purification technique used to separate proteins where there is an excess of salt. A sufficient increase in sodium chloride concentration means that the effects of salts on water structure causes protein to precipitate *via* entropic phenomena.

Precipitation of β-Lg, by the addition of 30 % (w/v) total sodium chloride, was advantageous because the procedure was reversible since the precipitant could be dissolved. After re-solubilisation in a neutral buffer (KH_2PO_4; pH 6.5), the protein solution was centrifuged and filtered to remove any remaining precipitates and β-Lg aggregates present in the sample. SDS-PAGE analysis (Figure 3.3; Lane 5) showed no significant increase in protein purity, although there was a 60 % drop in protein yield from the previous step (Table 3.1). However, this process provided an efficient means of concentrating the protein by about 4 fold, reducing the volume of the protein solution by 10 fold, and removing sodium chloride by exchange into a neutral buffer (pH 6.5) prior to size-exclusion chromatography.

During the initial purification runs, when following the prescribed protocol, final total β-Lg yields only amounted to approximately 1 mg of purified β-Lg from a 2 litre culture. To establish if protein concentration prior to HCl precipitation had an effect on final yields, protein concentration was increased from ~1 mg mL^{-1}, as published, to ~3-5 mg mL^{-1}, prior to precipitation with HCl. An increase in the final yield of up to ~10.6 mgs of purified protein per 2 litre culture was observed, dramatically reducing the cost and the time necessary to purify bovine β-Lg. Due to this marked escalation in the final yield, increasing protein concentration to 3-5 mg mL^{-1}, after dialysis and prior to treatment with acid, became an established step during purification. These results observed at lower protein concentrations and higher protein concentrations before acid precipitation were consistently reproducible, with respective protocols giving similar yields.

Ponniah and co-workers (2010) suggested that acid precipitation coupled to NaCl treatment was a method of culling β-Lg species that were not correctly folded, due to disulfide bond mismatch, as native-like β-Lg is extremely acid stable. However, this is not verified, but may explain the substantial drops in yield after the addition of 7 % and 30 % (w/v) NaCl at low pH, as not all of β-Lg C may have gone through the appropriate disulfide bond isomerase and chaperone processes of DsbC.

Size-exclusion chromatography was added as a polishing step to the purification protocol already established in the laboratory as dynamical NMR spectroscopy analyses demand highly pure protein. The purpose of this final stage was to remove residual contaminants and remaining aggregates present in the solution. The Superdex 75 10/300 size-exclusion column was used as β-Lg's molecular weight is within the optimal separation range of the column resulting in high resolution. Protein purification was monitored by OD_{280} absorption and analysed purity using SDS-PAGE (results not shown). Protein eluted as one major peak, and the fractions embodying the peak were pooled then dialysed overnight against the low pH buffer (KH_2PO_4 buffer; pH 2.6). Following dialysis, the protein solution was centrifuged and filtered to remove protein aggregates produced from a change in pH.

By adding a size-exclusion chromatography step as a polishing stage to the provided purification procedure, it was shown that majority of the residual contaminants were removed. Even though the final yield was compromised, the quality of the protein sample was improved, making it more amenable for NMR spectroscopy analyses.

3.2.4 Conformational Analyses by Means of NMR Spectroscopy

Neither solubility analysis of recombinant β-Lg C nor a purification protocol similar to one used to express and purify other β-Lg variants guarantees that the protein has been folded in its wild-type form. A protein's $^{15}N,^{1}H$-HSQC spectrum provides a fingerprint of its conformational state. Therefore, the structural integrity of the recombinant β-Lg C could be assessed by comparison of its HSQC spectrum with that of a β-Lg variant sample known to be correctly folded.

NMR spectroscopy analyses required uniformly ^{15}N-labelled protein that had been grown in M9 minimal media supplemented with ^{15}N ammonium sulfate. The $^{15}N,^{1}H$-HSQC spectrum of recombinant ^{15}N-labelled β-Lg C displayed good line-width and dispersion of the ^{1}H and ^{15}N resonances, indicating that the protein was well-folded (Figure 3.4). The overlay of the β-Lg C spectrum had similar peak dispersion to that of ^{15}N-labelled β-Lg B. This also implied that the C variant had been folded properly. The ^{15}N-labelled recombinant B variant had been determined to be folded correctly in previous studies (Ponniah *et al*., 2010), as it displayed a highly similar fingerprint to that of the wild-type β-Lg B HSQC, which relied on the natural abundance of ^{15}N at pH 2.6 for its $^{15}N,^{1}H$-HSQC spectrum,

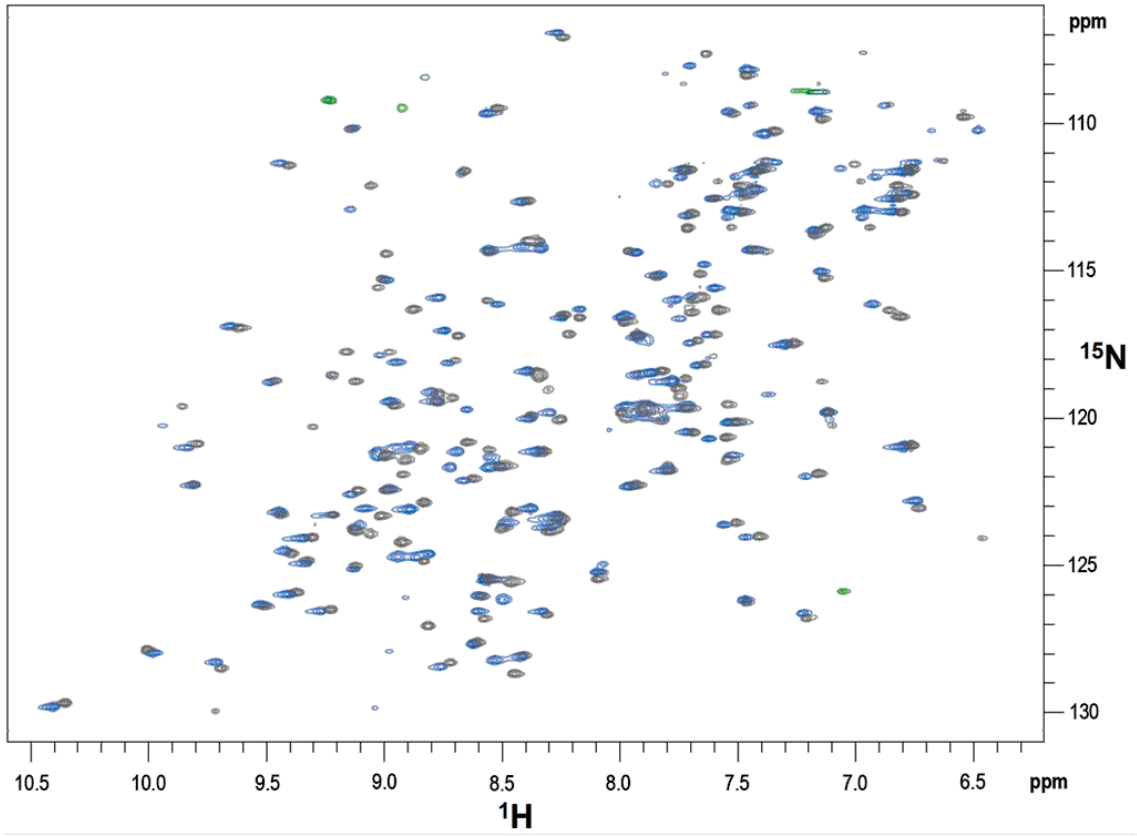


Figure 3.4 Overlay of the ^{15}N-Labelled β-Lg C HSQC Spectrum with the ^{15}N-Labelled β-Lg B HSQC Spectrum. The β-Lg C spectrum (sky-blue) has similar peak dispersion to that of ^{15}N-labelled β-Lg B (grey). This indicates that the C variant is folded correctly. NMR spectra were sampled at pH 2.6 and 305 K using 1 mM protein samples.

3.3 Assigning the Protein Backbone of β-Lactoglobulin C

In order to relate dynamical information to the primary sequence it is necessary to first assign as many peaks as possible in the $^{15}N,^{1}H$-HSQC spectrum by their respective residues. Although backbone assignments are available for β-Lg, ambiguity still existed for many peaks, particularly in crowded regions. Therefore a full backbone assignment was undertaken.

3.3.4 Backbone Assignments for β-Lg C at 305 K

A double-labelled ($^{13}C/^{15}N$) β-Lg C sample was used to assign the backbone of β-Lg C at 305 K. NMR backbone assignment experiments were performed at pH 2.6, using a 700 MHz spectrometer. At this pH, the signal widths and intensities in the $^{15}N,^{1}H$-HSQC spectra were homogenous, making it amenable to NMR spectroscopy analysis.

Sequential assignment of the polypeptide backbone atoms was achieved using NMR 3D CBCA(CO)NH, CBCANH, HNCO and HN(CA)CO pulse sequences. Ambiguities were resolved in some cases using the $^{15}N,^{1}H$-TOCSY-HSQC and $^{15}N,^{1}H$-NOESY-HSQC experiments. Assignments were checked for consistency with those published by Uhrínová *et al* (1998) for β-Lg A. The ^{1}H and ^{15}N chemical shifts of the backbone atoms are presented in Appendix B.1 and the assigned $^{15}N,^{1}H$-HSQC spectrum is shown in Figure 3.5.

Sequence-specific ^{1}H and ^{15}N resonance assignments were completed for 127 of 163 total residues (78 %) for recombinant β-Lg C at pH 2.6. Unassigned residues included the additional N-terminal methionine residue at position 0, nine proline residues at positions 38, 48, 50, 79, 113, 126, 144 and 153 and two cysteine residues located at positions 66 and 160. Other unassigned residues included those located in the C- and N-terminal regions of the protein molecule. It is likely that the undetected terminal residues have signals broadened beyond detection by conformational exchange. Also, other unassigned residues included those located in crowded or merged areas within the $^{15}N,^{1}H$-HSQC spectrum, regions where ambiguities could not be resolved.

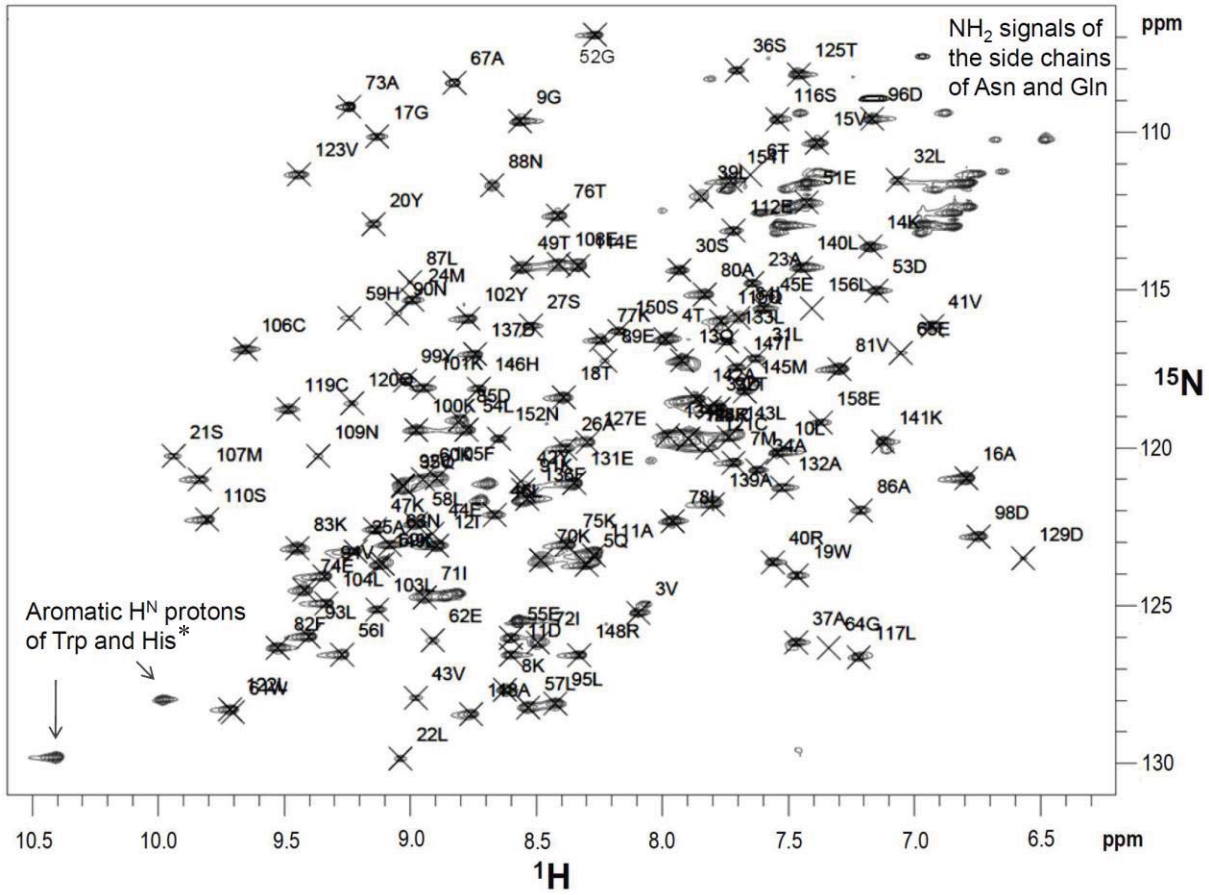


Figure 3.5 Assigned ^{15}N,^{1}H-HSQC Spectrum of Monomeric ^{13}C,^{15}N-β-Lg C.

The sample conditions were 50 mM KH_2PO_4 buffer, pH2.6, 5 % D_2O at 305 K.

* Side-chain atoms had been assigned previously (Uhrínová *et al.*, 1998).

3.3.5 Backbone Assignments for β-Lg C at 313 K and 320 K

Double-labelled ($^{13}C/^{15}N$) and single-labelled (^{15}N) β-Lg C samples were used to assign the backbone of monomeric β-Lg C at 313 K and 320 K. NMR backbone assignment experiments were performed at pH 2.6, using a 700 MHz spectrometer. Starting from the assigned $^{15}N,^{1}H$-HSQC spectrum at 305 K a series of spectra was recorded at 2 K intervals to trace the trajectory of each peak. Overlaid spectra 305 K, 313 K and 320 K are shown in Figure 3.6. However; due to some peaks separating or merging as the temperature changed, it was not possible to reliably identify all the peaks using this method. For these cases the $^{13}C,^{15}N$-β-Lg C sample was used to assign the peaks in the $^{15}N,^{1}H$-HSQC spectra, using the same combination of experiments described for 305 K. Sequence-specific ^{1}H and ^{15}N resonance assignments were completed for 116 of 163 total residues (71 %) for recombinant β-Lg C at 313 K, and for 112 of 163 total residues (68 %) at 320 K. Differences in the proportion of assigned residues, as the temperatures increased, were generally due to peaks merging making the ^{1}H and ^{15}N resonances more difficult to assign.

The ^{1}H and ^{15}N chemical shifts are presented in Appendix B.1. Temperature change did not alter the $^{15}N,^{1}H$-HSQC significantly, indicating that the averaged conformation of the structure is similar to the structure at 305 K.

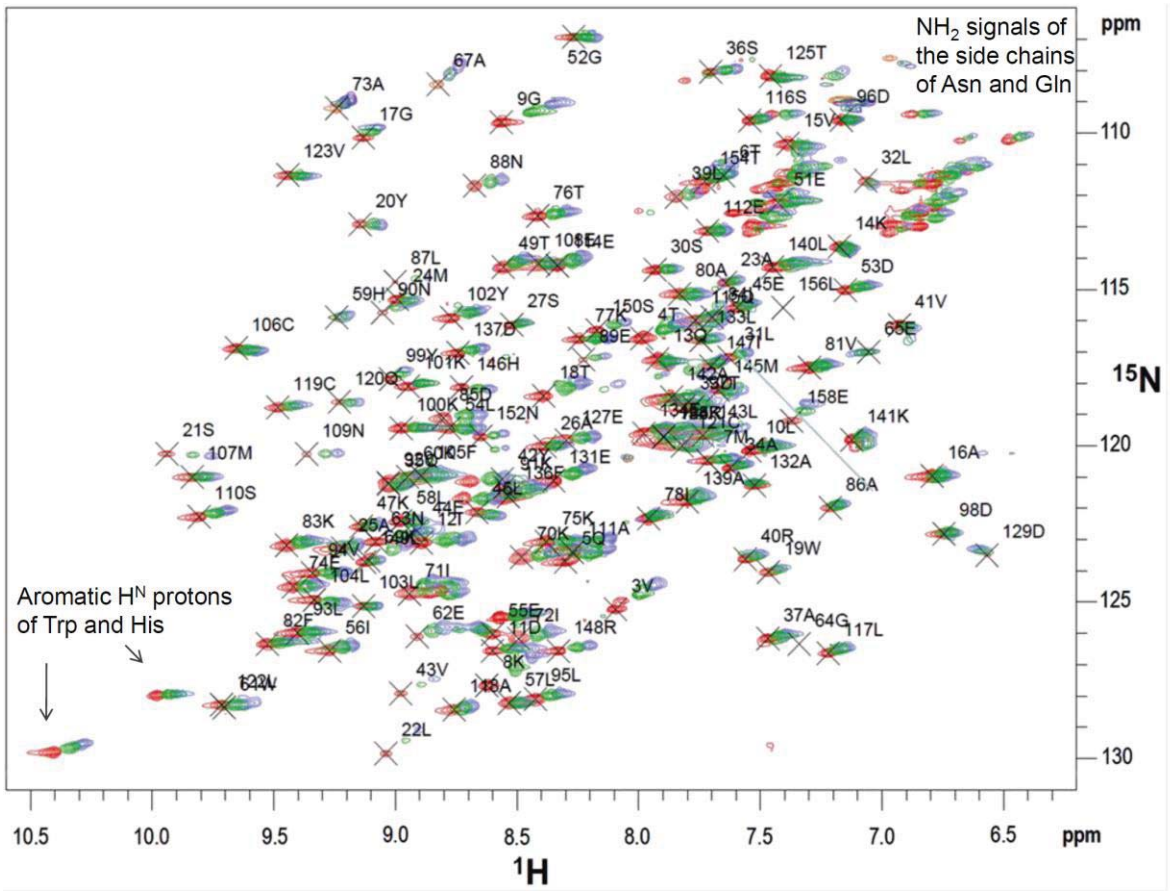


Figure 3.6 An overlay of Three $^{15}N,^{1}H$-HSQC Sampled at 305 K (red), 313 K (green) and 320 K (blue).

This series of HSQCs were collected on ^{15}N-labelled β-Lg C (1 mM) in 50 mM KH_2PO_4 buffer; pH 2.6.

3.4 ^{15}N NMR Backbone Dynamics of Bovine β-Lg C at 305 K

^{15}N relaxation data were collected on monomeric β-Lg C samples at 305 K and at one field strength (700 MHz). ^{15}N-HSQC relaxation spectra were collected as described in Section 2.4.5. Of the 154 ^{15}NH expected backbone resonances, peak height measurements for 110 were considered reliable for relaxation analysis. Backbone amides with no relaxation results corresponded to unassigned residues, or to residues with cross-peaks considered to be too overlapped. Unfortunately, for the case of Gly64 (one of the variants substitution sites) the intensity of the cross-peak was too low for a reliable measurement of peak height.

3.4.1 Results for ^{1}H-^{15}N Relaxation of Monomeric Bovine β-Lg C

The values of the measured ^{15}N relaxation parameters, R_1, R_2 and NOE, for β-Lg C measured at 305 K, were each plotted separately as a function of residue number. These plots are featured in Figure 3.8 and Figure 3.9, and the rates are listed in Appendix C.1. Examples of relaxation delay curves, used to determine R_1 and R_2 relaxation rates independently, are presented in Figure 3.7.

R_1 and R_2 relaxation rates are sensitive to different motional frequencies. R_1 relaxation rates give information in regards to motional properties of the backbone amides, with a frequency of about 10^8-10^{12} s^{-1}, whereas R_2 values are sensitive to these motions occurring at these frequencies, as well as being dependent on dynamics on the μs to ms time-scale. Therefore, by measuring these two rates, motions can be detected over a range of time-scales. The variation in relaxation rates over the entire protein molecule ranges from 1.01 s^{-1} (Ala80) and 1.39 s^{-1} (Gly9) for R_1 and from 7.59 s^{-1} (Ala34) to 22.97 s^{-1} (His59) for R_2. The average R_1 for all residues is 1.14 s^{-1}, with an average R_1 error for an individual residue of 0.04 s^{-1} and an estimated standard deviation for scatter about the mean value of 0.07 s^{-1}. The average R_2 rate for all residues is 12.59 s^{-1} with an average R_2 error for an individual residue of 0.49 s^{-1} and an estimated standard deviation for scatter about the mean value of 3.01 s^{-1}.

Sequential variation of R_1 rates is observed through several regions of the protein backbone. R_1 values are particularly high throughout the A/B loop, when compared to the rest of the structure, which could point towards relatively high order (S^2) and/or relatively long internal correlation times (τ_e) for the NH bonds in this region (increases in both of these parameters can potentially increase the spectral density at relaxation frequencies

(see Figure 2.3), as indicated by significantly higher than average R_1 values (> 1.21 s^{-1}) displayed by residues Ser27 (1.24 s^{-1}), Leu31 (1.31 s^{-1}), Leu32 (1.30 s^{-1}), Ser36 (1.23 s^{-1}), Ala37 (1.31 s^{-1}) and Arg40 (1.27 s^{-1}). These values are interesting for a comparatively large loop, but could be explained by the presence of a 3_{10}-helix (residues 29-32), which may provide a more rigid structure to this region. Other residues displaying significantly higher than average R_1 rates, possibly suggesting increased rigidity, include Ser116 (1.22 s^{-1}) and Leu117 (1.25 s^{-1}), which are positioned in the hydrogen bond-stabilised helical structure within the G/H loop, and Asp137 (1.24 s^{-1}), positioned in the major α-helix (α-2). Although the N-terminal regions display significantly higher than average R_1 values, these could be attributed to a longer τ_e, for reasons just previously discussed. Interestingly, nine residues display R_1 rates significantly lower than the average (< 1.07 s^{-1}). Most of these residues that show increased flexibility are located in loops and other sites that link secondary structural elements (Figure 3.8 (A)), which are expected to be more mobile regions of the protein. Gradual variations were observed for backbone amides positioned in the secondary structural elements, with values progressively dipping for N-H bonds of residues positioned in the β-B and β-H strands and progressively rising for N-H bonds for residues sitting in the β-C, β-D and β-F strands, which could point towards patterns in respect to changes in flexibility progressing through the secondary elements. However, as R_1 is sensitive to the contributions of the correlation time of the internal motions, interpretations into these trends, with respect to β-Lg's structure have been withheld until the more quantitative analysis using the model-free approach.

Like R_1, ^{15}N R_2 relaxation is sensitive to higher frequency (ns) motions but is also dependent on motions on a slower μs to ms time-scale, which could be contributed to by conformational exchange. Spanning the length of β-Lg C's backbone, the R_2 values are relatively uniform, although some variation is observed (Figure 3.8 (B)). There are a number of residues that display significantly higher than average R_2 rates (> 15.60 s^{-1}), well above the mean. These residues include Leu22 (22.85 s^{-1}), Leu39 (21.39 s^{-1}), His59 (22.97 s^{-1}), Glu62 (21.91 s^{-1}), Ala67 (22.35 s^{-1}), Ala73 (17.97 s^{-1}), Asn88 (22.24 s^{-1}), Asn109 (18.78 s^{-1}), Glu127 (16.81 s^{-1}), Phe151 (19.62 s^{-1}) and Glu158 (17.96 s^{-1}). Six of these residues are found in loop regions and links (Figure 3.8 (B) and three sit at or near the edges of the secondary structural elements. Ala62 and Ala73 are both sited in the β-D strand, whose hydrogen bonding potentials are not completely fulfilled, and Leu22 sits at an important position in the 90˚ bend in the middle of the β-A strand. Therefore, the

suggestion of conformational exchange for these residues is not unreasonable due to the positions they hold within β-Lg's structure. At the other end of the R_2 scale, five residues sited at the N-terminal end of β-Lg C display lower than average R_2 values (< 9.58 s^{-1}): Val3 (9.51 s^{-1}), Thr4 (9.32 s^{-1}), Gln5 (8.15 s^{-1}), Gly9 (8.58 s^{-1}) and Leu10 (9.26 s^{-1}). In addition, Glu51 (8.96 s^{-1}), located in the B/C loop between two prolines and interestingly, Ala34 (7.59 s^{-1}), which is located the more rigid region of the A/B loop, also have significantly low R_2 values, pointing to a high degree of flexibility.

The {^{1}H}-^{15}N NOE enhancements of backbone amides are sensitive to motions of the backbone that are on a ns to ps time-scale. Values greater than 1.0 can indicate an increase in flexibility. Values of {^{1}H}-^{15}N NOE enhancement across the backbone are relatively uniform, although some variation is observed (Figure 3.9 (A)). The average NOE enhancement value is 0.73, with an average NOE error for an individual residue of 0.03, with an estimated standard deviation for scatter about the mean of 0.07. NOE enhancement values spanned the range of 0.41 (Val3) to 0.81 (Asp98). Of the total NOE enhancement values, twelve values were significantly less than average (< 0.66). Six of the twelve belong to backbone amides of residues positioned in the N-terminal region (Val3 (0.41), Thr4 (0.42), Glu5 (0.43), Lys8 (0.53), Gly9 (0.63) and Leu10 (0.55)), providing evidence for flexibility in this region, with five having significantly low R_2 rates. A further five of the twelve NOE enhancement values correspond to residues Ala34 (0.61), Glu51 (0.60), Asn63 (0.58), Glu127 (0.64) and Ala142 (0.60), sited in loops and links, and Asp129 (0.65), which is positioned at the N-terminal end of the major α-helix. These results point towards flexibility in these regions.

^{15}N R_1 and ^{15}N R_2 relaxation rates change by approximately the same amount with increasing amplitude of motion; however, R_2 is also sensitive to the presence of conformational exchange. The ratio R_2/R_1 remains reasonably constant with changes to the order parameter, S^2. However, since R_2 is strongly affected by presence of conformational exchange, examining the R_2/R_1 ratio is helpful for spotting parts of the backbone experiencing relatively slow types of motions. The R_2/R_1 ratio ranges from 6.16 to 20.88 with a mean value of 11.06. The average R_2/R_1 error for an individual residue is 0.67, with an estimated standard deviation for scatter about the mean value of 2.69 (Figure 3.9 (B)). Six residues (Val3 (7.56), Thr4 (8.28), Glu5 (7.33), Gly9 (6.16), Leu10 (7.27) and Ala34 (6.91) showed significantly low values of R_2/R_1 (< 8.37), while twelve residues: Leu22 (17.91), Leu39 (18.19), His59 (20.04), Glu62 (18.99), Ala67

(20.88), Ile71 (13.81), Ala73 (15.13), Asn88 (19.51), Asn109 (17.35), Glu127 (15.93), Phe151 (17.32) and Glu158 (15.88), display significantly high ratios (> 13.75). Interestingly, the majority of residues sited at the N-terminal end of the protein and residue Ala34, found in the comparatively rigid A/B loop, display R_2/R_1 below the average, indicative of a greater degree of mobility. Residues with higher than average R_2/R_1 ratios (Figure 3.9 (B)), except for Ile71 (also positioned in the β-D strand), have significantly high R_2 values, which have been commented on previously. These data indicate that the backbone amides corresponding to residues Leu22, Leu39, His59, Glu62, Ala67, Ile71, Ala73, Asn88, Asn109, Glu127, Phe151 and Glu158 maybe be undergoing slow conformational motions on a μs to ms time-scale.

The experimental parameters R_1, R_2 and NOE can provide information on internal motions to an extent; however, R_1 and NOEs have a strong dependence on the effective correlation time, τ_e. Therefore, these rates can only provide an indication of low S^2 if τ_e is assumed to be short and making minimal contribution to the relaxation. R_1 and R_2 change by the same amount with increasing S^2; however, R_2 is also sensitive to the presence of conformational exchange. In order to compose a more quantitative assessment of β-Lg's internal dynamics, the parameters need to be analysed in concert. This was achieved using model-free analyses.

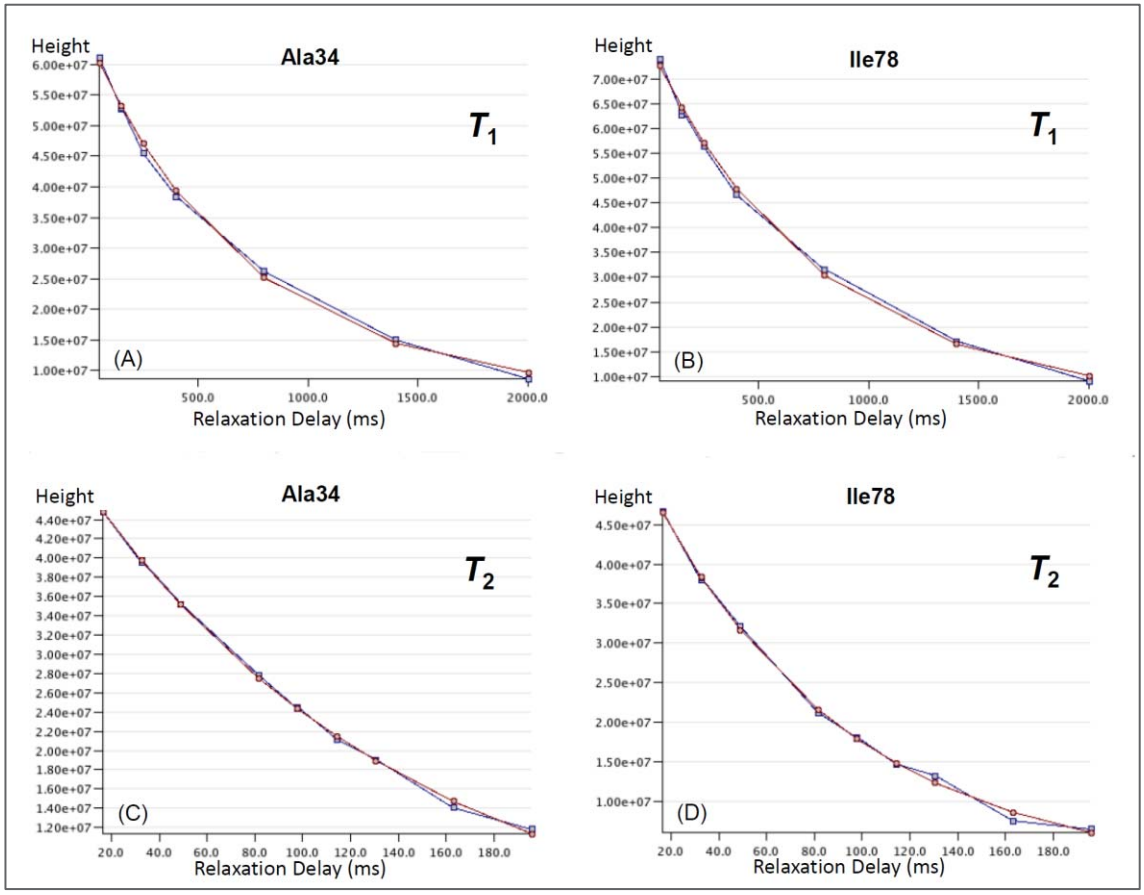


Figure 3.7 Examples of Plots Used to Determine Relaxation Rates.

R_1 and R_2 relaxation rates are the inverse of T_1 and T_2 relaxation times, respectively. Example plots of T_1 and T_2 relaxation delay series, required to determine relaxation rates, are shown for the N-H bonds of Ala35 and Ile78, (A) T_1 for Ala34 (B) T_1 for Ile78 (C) T_2 for Ala34 and (D) T_2 for Ile78. Using ANALYSIS 2.1 (Vranken *et al*., 2005), experimental points were fitted with a spline curve, together with the fitted points from the best fit exponential. These figures are formed with screenshots from ANALYSIS 2.1.

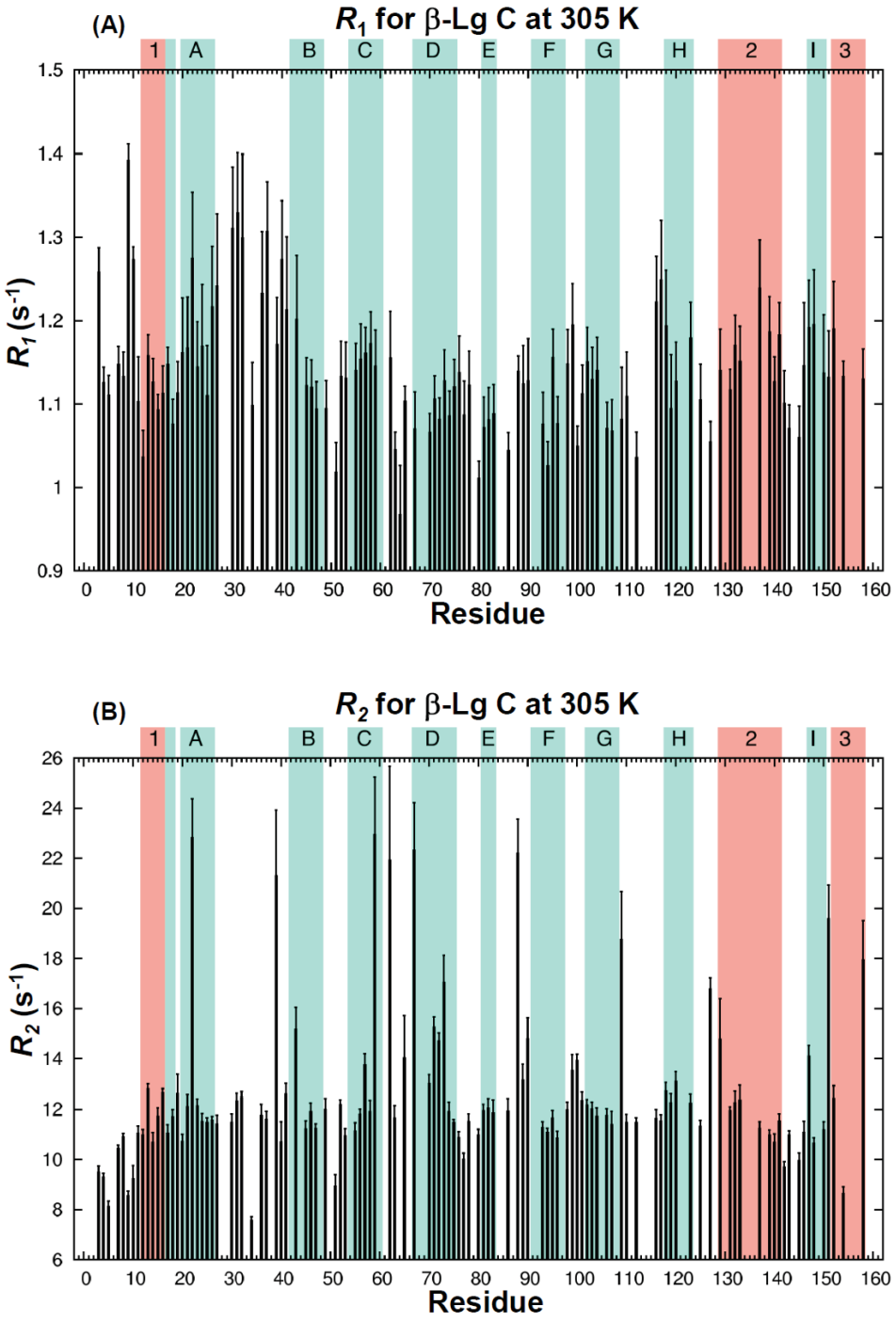


Figure 3.8 A Summary of ^{15}N R_1 and R_2 Relaxation Rates for Monomeric β-Lg C.
Summary of relaxation rates; (A) A plot of the measured ^{15}N longitudinal (R_1) relaxation rates *vs*. residue number, and (B) a plot illustrating ^{15}N transverse (R_2) relaxation rates *vs*. residue number. Experimental parameters were derived at 305 K and at pH 2.6. In plots (A) and (B) the positions of the nine β-strands (labelled A-I) and α-helices (labelled 1-3) are highlighted with teal and salmon, respectively.

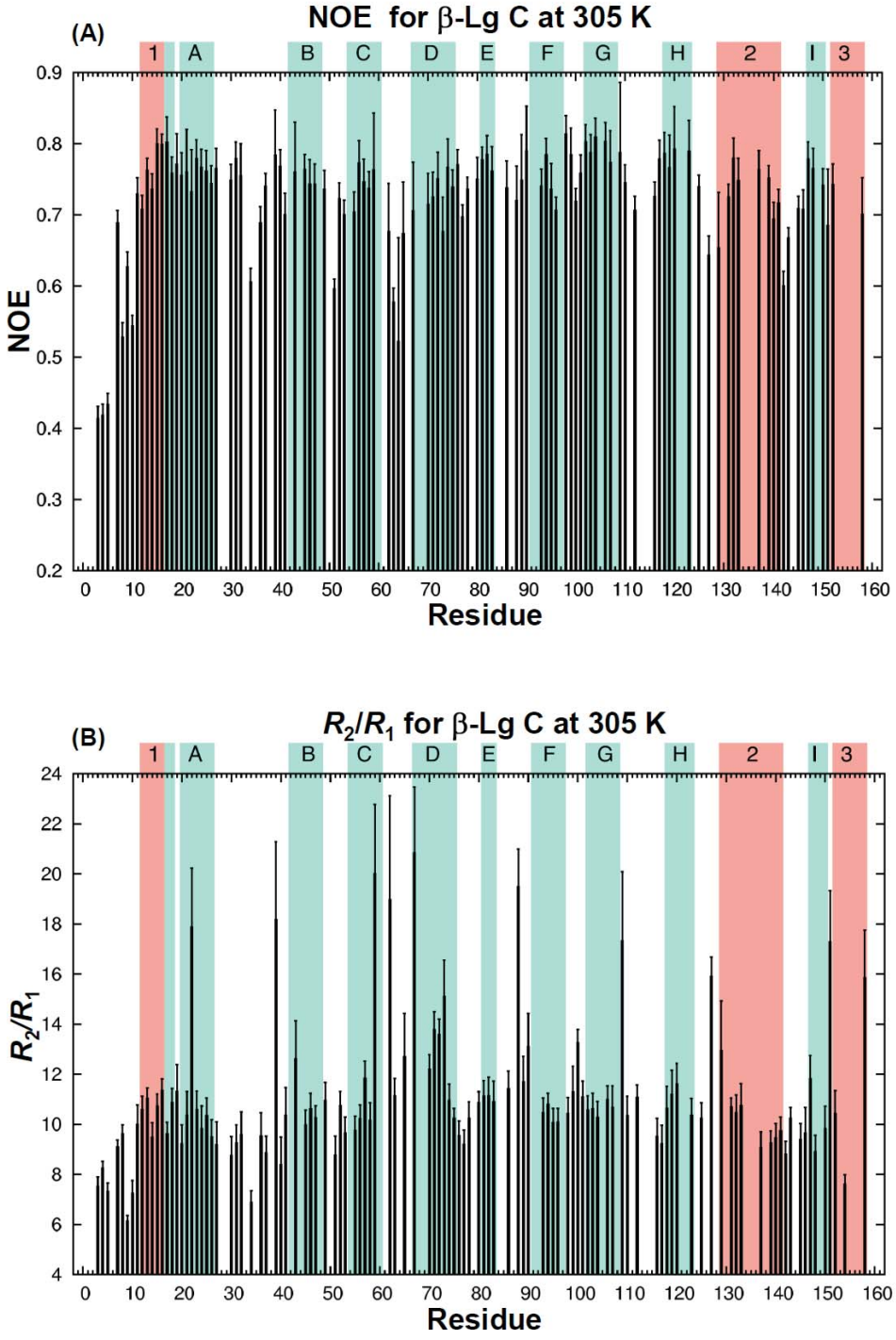


Figure 3.9 Summaries of {^{1}H}-^{15}N NOE Enhancement Values and the Ratios R_2/R_1 for Monomeric β-Lg C.

Summary of relaxation values; (A) a plot of {^{1}H}-^{15}N NOE enhancement values *vs*. residue number, and (B) a plot of the ratio R_2/R_1 *vs*. residue number. Experimental relaxation parameters were derived at 305 K and at pH 2.6. In both plots the positions of the nine β-strands (labelled A-I) and α-helices (labelled 1-3) are highlighted with teal and salmon, respectively.

3.5 Model-Free Analysis of Dynamics at 305 K

3.5.1 The Determination of the Overall Correlation Time for β-Lg C at 305 K

The backbone ^{15}N relaxation data were interpreted with the Lipari-Szabo model-free formalism. Model-free analysis still uses models to account for the N-H bond motion, but the exact nature of this motion is undefined. The first step of the model-free analysis is to estimate the molecular rotational correlation time, τ_m. If τ_m can be defined, the parameters that define the internal motion can be optimised. Note, β-Lg is assumed to behave as a spherical molecule, tumbling isotropically (Section 2.4.1).

The overall correlation time was determined experimentally from the ratio of tranverse (R_2) and longitudinal (R_1) relaxation rates for each residue from a subset of residues with rigid NH bonds i.e. those residues where relaxation was due entirely to molecular rotation with no contribution from internal dynamics. Eighteen residues (Val3, Thr4, Gln5, Gly9, Leu10, Leu22, Ala34, Leu39, Gln59, Glu62, Ala67, Ile71, Ala73, Asn88, Asn109, Glu127, Phe151 and Glu158) were omitted from the estimate as their respective R_2/R_1 values were further than one standard deviation from the mean value, indicating flexibility (Figure 3.9 (B)). The averaged R_2/R_1 value was 11.06 with an average error of R_2/R_1 for an individual residue of 0.67 and an estimated standard deviation for scatter about the mean value of 3.05. A further nine residues (Met7, Lys8, Ser36, Glu51, Asn63, Glu65, Val129, Ala142 and Leu143) were excluded because their NOE values were less than 0.7, again indicating flexibility (Figure 3.9 (A)). An initial τ_m value was calculated (Section 2.4.1) separately for each backbone ^{15}N-^{1}H bond of the remaining 82 residues, and the mean calculated as a preliminary estimate of the overall correlation time of the protein molecule, which was equal to 9.0 ns.

3.5.2 Results for Selection and Distribution of Model-Free Motional Parameters

The optimal value of τ_m for β-Lg C at 305 K was iteratively refined using the model-free protocol and was finally determined to be 8.54 ns. Model-free analysis was used for the quantitative interpretation of ^{15}N NOE, R_1 and R_2 data for each residue. In this analysis, up to five types of model-free spectral density functions were used to derive the motional parameters for each selected backbone ^{15}N-^{1}H bond. The possible models included the following combination of parameters S^2_s; S^2_s and τ_e; S^2_s and R_{ex}; S^2_s and τ_e and R_{ex}; and a two time-scale model, S^2_s, τ_e and S^2_f, (Clore *et al*., 1990a, Clore *et al*., 1990b), as described in Materials and Methods (Section 2.4.1). The most appropriate model was

chosen based on statistical tests (Mandel *et al*., 1995). Fitting was performed using MODELFREE4.15 (Palmer *et al*., 1991, Mandel *et al*., 1995) under the iterative control of FAST-Modelfree (Cole & Loria, 2003).

Of the 110 residues that were reliable for model-free analysis, all could be fitted into one of the five models. The distribution of fits is listed in Table 3.2. In general, for each backbone amide group, the model exploiting the fewest parameters, but capable of reproducing relaxation parameters, was selected.

Model One	Model Two	Model Three	Model Four	Model Five	No fit
S_s^2	S_s^2 - τ_e	S_s^2-R_{ex}	S_s^2 - τ_e-R_{ex}	S_s^2 - τ_e- S_f^2	
39	37	5	13	16	0

Table 3.2 Distribution of fits for model-free analyses of β-Lg C (305 K) and at a magnetic field strength of 16.4T (corresponding to the resonance frequency of 700.25 MHz for ^{1}H).

The spectral density models used to describe internal motions of β-Lg C and the obtained motional parameters, derived from measured R_1, R_2 and NOE relaxation data, for each of the 110 residues are listed in Appendix C.1.

3.5.3 Results for the Model-Free Analysis of Backbone Dynamics

In model-free analysis the parameters S^2, R_{ex} and τ_e provide information on how much a backbone ^{15}N-^{1}H vector is moving and how fast. The direction and distribution of motions is unknown. The optimised values of the model-free parameters for β-Lg at 305 K are listed in Appendix D.1 and are plotted in Figure 3.10, Figure 3.11 and Figure 3.12. The following results point to regions and residues within β-Lg C that have been described by the motional parameters.

Amplitude of Motion (S^2)

Of the derived parameters S^2 is the most important as it describes the amplitude of the internal motion. An S^2 value of 1.0 defines rigidity of the backbone amide whilst sequentially lower S^2 values points to progressively increasing mobility. Note that $S^2 = S^2_s S^2_f$ and for residues fit with models one to four $S^2_f = 1$ (i.e. the overall order parameter is the product of two order parameters, one for relatively slow motions, the other for relatively fast). The order parameters for the backbone amide of each residue included in model-free analysis are illustrated in Figure 3.10 and listed in Appendix D.1. The average S^2 value of the 110 residues is 0.82, with an average error of S^2 for an individual residue of 0.02 and an estimated standard deviation for scatter about the mean

value of 0.08. For 95 % of the residues with measured relaxation data, S^2 ranged from 0.74 - 0.90, indicating that the internal motions of β-Lg at 305 K are essentially restricted, consistent with values for a densely hydrogen-bonded globular protein. However, S^2 values range from 0.50 (Ala34) to 0.99 (Val43), indicating that some regions of the protein backbone are highly flexible, while other regions are markedly rigid.

Looking across the length of the β-Lg C's backbone, gradual variations of S^2 are observed along the lengths of some of the structural elements (Figure 3.11), in some areas dipping, showing a progressive increase in flexibility, and in other regions rising, progressively becoming more rigid. Not surprisingly, the loop regions and N-terminus appear to be more flexible than their neighbouring secondary structural elements. Starting with the N-terminus (Leu1 - Leu10), this region shows pronounced reductions in S^2, due to its lack of rigid structure. Five residues: Val3 (0.65), Thr4 (0.64), Gln5 (0.55), Gly9 (0.55), and Leu10 (0.62), display order parameters more than two standard deviations (i.e. $S^2 < 0.66$) from the mean value, whereas Met7 (0.73) is more than one standard deviation and Lys8 (0.78) have close to average flexibility. Met7 forms an H-bond with Val94 (Qin *et al*., 1998b, Uhrínová *et al*., 2000), which is a plausible explanation for its relative rigidity. Relaxation data collected for the C-terminal region (residues beyond strand β-I) are sparse; therefore a trend in motional parameters is not observed for this region.

Moving on to the loops and links, many of the ^{15}N spins of residues showing significantly low order parameters (< 0.74) are positioned in these regions. These regions are where increased flexibility is expected. In general the loop regions are more flexible than their respective neighbouring secondary structural elements. Residues displaying significantly lower than average order parameters include Glu51 (0.62), Lys77 (0.70), Ala142 (0.68) and Met145 (0.70). Glu51 and Lys77 are located in loop regions, and Ala142 and Met145 are both situated in the link between the major helix (α-2) and the β-I strand (Figure 3.11). The lowest order parameter (0.50 ± 0.01) is observed for residue Ala34, positioned in the A/B loop (residues Ser27 - Val41), signifying a high degree of flexibility for its ns – ps time-scale motions. At neutral pH the residues Ile29 - Gln35 are involved in forming the dimer and have a distinct well ordered conformation. At pH 2.6, residues close to Ala34 display relatively high S^2 values, especially Leu31 (0.90) and Leu32 (0.90), which form part of the 3_{10}-helix (residues Ile29 - Leu32). Unfortunately, ^{15}N relaxation data could not be acquired for the immediate neighbours of Ala34. However, at low pH, where β-Lg C is monomeric, the loop could be more disordered, providing an

explanation for Ala34's high mobility. The relative rigidity of Leu39 (0.89) in the A/B loop may be a consequence of its side-chain protruding into an area of the cavity that is densely packed. At low pH the E/F loop is in its closed conformation, serving to block the opening of the cavity. The side-chain of Glu89 (0.89) in this loop, is buried, which gives rise to a relatively high S^2.

β-Lg's high content of rigid β-sheet structure means that β-Lg is stabilised by a considerable network of inter-strand H-bonds (Figure 1.2). Residues with ^{15}N spins displaying significantly high order parameters (> 0.90) include Leu22 (0.97), Val43 (0.99), Leu57 (0.92), Ala118 (0.91), a site of variation between variant A and both C and B variants, Gln120 (0.91) and Ile147 (0.98). These residues are all positioned in the β-strands. Leu22 is positioned at the hinge point of strand β-A and forms one intra- and two inter-strand H-bonds, which could account for its rigidity. Leu57 is located in the middle of strand β-C and associates with both Glu44 and Glu45, situated in strand β-D. Gln43, sitting next to residues Glu44 and Glu45, displays the highest order parameter, which is interesting as it does not participate in any H-bonding; however, this could be explained by the kink in the β-strand at Gln43. Ile147, located in strand β-I, shares one H-bond with Ala26, located towards the end of the β-A strand.

Ala118 and Gln120 are positioned in a highly ordered region within the β-H strand, both neighbouring the buried Cys119-Cys106 *cis*-disulfide bond. Gln120 is also located next to the free thiol, Cys121, which forms three inter-strand H-bonds. This extensive inter-strand H-bond network formed by these residues and their neighbours, coupled to the presence of a stabilising disulfide bridge, are possible factors that may contribute to the significantly high rigidity of Ala118 and Gln120. ^{15}N relaxation data were not obtained for the free cysteine.

S^2 trends were observed along the lengths of many of the secondary structural elements (Figure 3.11 (A)), pointing towards an increase in flexibility across the backbone of some β-strands, as indicated by gradual reduction in S^2 values, and across others, a progressive increase in rigidity, as shown by an increase in S^2. Some changes in flexibility are also observed across the length of the major α-helix (α-2), as seen by the reduction in values for S^2. The reduction in S^2 along the helix reflects a decrease in order, from the bottom of the β-barrel towards the top. S^2 in the α-2 helix decreases from 0.89 to 0.75. Other S^2 trends, which could arise from the strands' orientation within the β-barrel, are shown in

Figure 3.11 (B). Strands β-A and β-G become more flexible as they reach from the bottom of the barrel (more confined end) towards the open end of the barrel, whereas strands β-D and β-H, become more rigid as they protrude down into the relatively closed end of the barrel. The increase in flexibility towards the top of the barrel, for four of the β-strands, may help with the ingress of ligands into the barrel of this lipocalin.

When comparing strands, the lower order of the β-D strand is possibly a reflection of its hydrogen-bonding potentials not being completely fulfilled, whereas the extensive bonding network formed by residues in strand β-H could provide a plausible explanation for its relatively higher order. S^2 averages were not calculated across the secondary structural elements for comparison, as ^{15}N relaxation data were not available at 305 K for a number of key residues that could contribute significantly to rigidity. For example, Phe105 (β-G), whose H_N signal persists at temperatures as high as 80 ˚C (Edwards *et al*., 2002), and Tyr42 (β-B), whose side-chain is buried and is unavailable for chemical modification (Uhrínová *et al*., 2000), have not yielded ^{15}N relaxation data.

Residues that Show Conformational Exchange (R_{ex})

R_{ex} accounts for the presence of low frequency motions (Palmer, 1997). Significantly large values of R_{ex} identify residues that experience conformational exchange on a micro to millisecond time-scale. In these studies R_{ex} is only indicative of conformational exchange as R_2 relaxation was measured only at 700 MHz. In theory, conformational exchange can be confirmed by measuring the R_2 rate at least at two different field strengths.

The optimised model-free fitting of the dynamics data suggests the presence of a slow conformational exchange for 18 of the 110 residues for which relaxation data were collected at 305 K. Collectively, 5 and 13 residues fit into model three (S^2_s-R_{ex}) and model four (S_s^2 - τ_e-R_{ex}) respectively (Table 3.2; Appendix D.1). Model three differs from model four in that the contribution by the effective correlation time (τ_e) is too short to affect the relaxation data. However, both models reveal residues that may be undergoing conformational exchange.

Of the 11 residues that had significantly higher than average R_2 rates, ten of these residues produce the largest R_{ex} terms (Figure 3.12 (A) & Appendix D.1). These residues include Leu22 (positioned at the 90˚ bend of β-A), Leu39 (A/B loop), His59 (near edge of β-C), Ala67 (β-D), Ala73 (β-D), Asn88 (E/F loop), Asn109 (G/H loop), Glu127 (link between

β-H and major α-helix), Phe151 (link between β-I and α-3) and Glu158 (α-3). The other additional residues that required the R_{ex} term, are Lys70, Ile71 and Ile72 (clustered in strand β-D), Gln13 (α-1), Asn63 (C/D loop), Lys100 (F/G loop) and Glu112 (G/H loop).

Leu22 is critically positioned at the 90˚ bend of β-A, where it occupies an important position between the two β-sheets. Its high conformational exchange rate points towards Leu22's involvement with both β-sheets. The R_{ex} term is possibly due to the slight expansion and contraction of the β-barrel, which involves a change in Leu22's conformation. This has been documented before for Ser21 in studies published by Uhrínová *et al*. (2000). However, these studies point to residue Leu22 occupying this hinge point.

Out of the residues positioned in loop regions whose dynamics indicate a significant R_{ex} term, only those displaying significantly high R_2 rates were located in loops at the open end of the β-barrel (loops A/B, E/F and G/H). Other residues whose motions have been defined with smaller R_{ex} contributions, and which are positioned in loops at the open end of the β-barrel, include Asn63 (loop C/D) and Glu112 (preceding a 3_{10}-helical turn in the G/H loop). Lys100 is the only residue located in a loop region at the bottom end of the β-barrel with a detectable R_{ex} term.

Five residues positioned on the β-D strand display dynamical parameters that are associated with slow conformational motions on a micro to millisecond time-scale. This comparatively long β-strand (Ala67 – Lys75) undergoes interesting dynamics, which stems from its hydrogen-bonding potentials not being completely fulfilled. Progressing along the strand, from the open end of the barrel to the closed end, motions of backbone amides are described with the conformational exchange constant, R_{ex}, indicating that this region of the protein, down to Ala73, is undergoing motions suggestive of conformational exchange. Glu74 does not appear to have a contribution from R_{ex}. This residue displays the highest S^2 value for this region. This is not surprising as Glu74 is the only residue in strand β-D that participates in inter-strand H-bonding with residues located in strand β-E (Qin *et al*., 1998b), providing an explanation for its rigidity on nano to picosecond time-scale scale. Slower micro to millisecond motions have not been detected for residues located in the shorter β-E strand (Val81 - Lys83) at 305 K, which is probably due to its short length.

His59 is positioned at the B→C substitution site in the area of the β-C strand that is next to residues that form a frustrated inter-strand H-bonding network with other residues positioned in strand β-B. Like Leu22, this residue has a relatively high order parameter (0.90) and comparatively high R_{ex} constant, indicating that its backbone amide is fairly rigid on a ns-ps time-scale, but undergoes slower ms-μs movements that point towards conformational exchange. Two explanations are provided for this slower motion: 1) similar to Leu22, His59 may be positioned in a kink region formed by irregular H-bonding, which could contribute to conformational exchange when the β-barrel expands or contracts slightly (Figure 1.2) and/or 2) the external disulfide bond formed between Cys66 in the C/D loop and Cys160 in the C-termini could produce a stretching in this region of β-Lg, contributing to correlating slower motions in this region and the region of the β-D strand, which is not particularly tethered well to strand β-E (Figure 1.2). This proposed segmental motion caused by the Cys66-Cys160 disulfide bridge has been commented on previously by Uhrínová *et al.* (2000).

Leu151 is the only residue in a link region that separates strand β-I (forming part of the dimeric interface) from the α-3 helix near the C-terminal region. It is possible that its position at this bend is responsible for its conformational exchange term. Other residues that are possibly experiencing motions on a slow milli to microsecond time-scale include Gln13 and Glu158, which are both positioned in the minor helices, α-1 and α-3, respectively.

Correlation Times (τ_e)

The effective correlation time defines the rate of the internal motion, whose amplitude is defined by S^2. Typically, values of τ_e are imprecisely determined, and are usually not analysed in detail (Palmer, 1997). However, correlation times are generally proportional to mobility. The correlation time for the internal dynamics, τ_e, could be extracted for 65 of the 110 backbone amides (approximately 60 %), whereas for the others it is too short to effect relaxation (Figure 3.12 (B)). Residues with significant τ_e contributions are distributed throughout the protein molecule, but are absent from two stretches of residues encompassing the highly-ordered β-G and β-H strands. These two regions are stabilised by an extensive network of H-bonding and a disulfide bridge formed by residues Cys109 (β-G) and Cys119 (β-H). Backbone amides in these strands are fitted with the simple one-parameter model (model one). Residues fit with a τ_e > than 500 ps are positioned in the N-terminus, the loop regions, the minor helix

(α-1) and the C-terminal end (open end of barrel) of the major helix (α-2), indicative of a high level of mobility in these locations.

Two Time-Scale Spectral Density Functions (Model Five)

There are residues undergoing librations on two time-scales, one slow, the other fast. The overall order parameter is the product of the order parameters for these two motions i.e. $S^2 = S^2_s S^2_f$. In this study all residues requiring model five fitting have a τ_e contribution greater than 500 ps, pointing towards a high degree of mobility in these regions. In this study, six residues; Val3, Thr4, Gln5, Met7, Gly9 and Leu10, corresponding to the N-terminal region, required the two time-scale spectral density function to describe their dynamics. At 305 K, residues positioned in the loop regions (A/B loop: Ser30, Ala34 and Ala37; B/C loop: Glu51; D/E loop: Lys77, and the link between the major α-helix and the β-I strand: Ala142, Met145), required a two-time-scale spectral density function to account for the measured relaxation data. Interestingly, two residues (Ala139 and Leu140), sitting at the open end of the calyx, in the major α-helix, required a two-time-scale spectral density function as well as one residue (Lys14) positioned in the minor helix (α-1), pointing towards more complicated motions in these regions. These results are typical for disordered regions of proteins (Palmer, 1993).

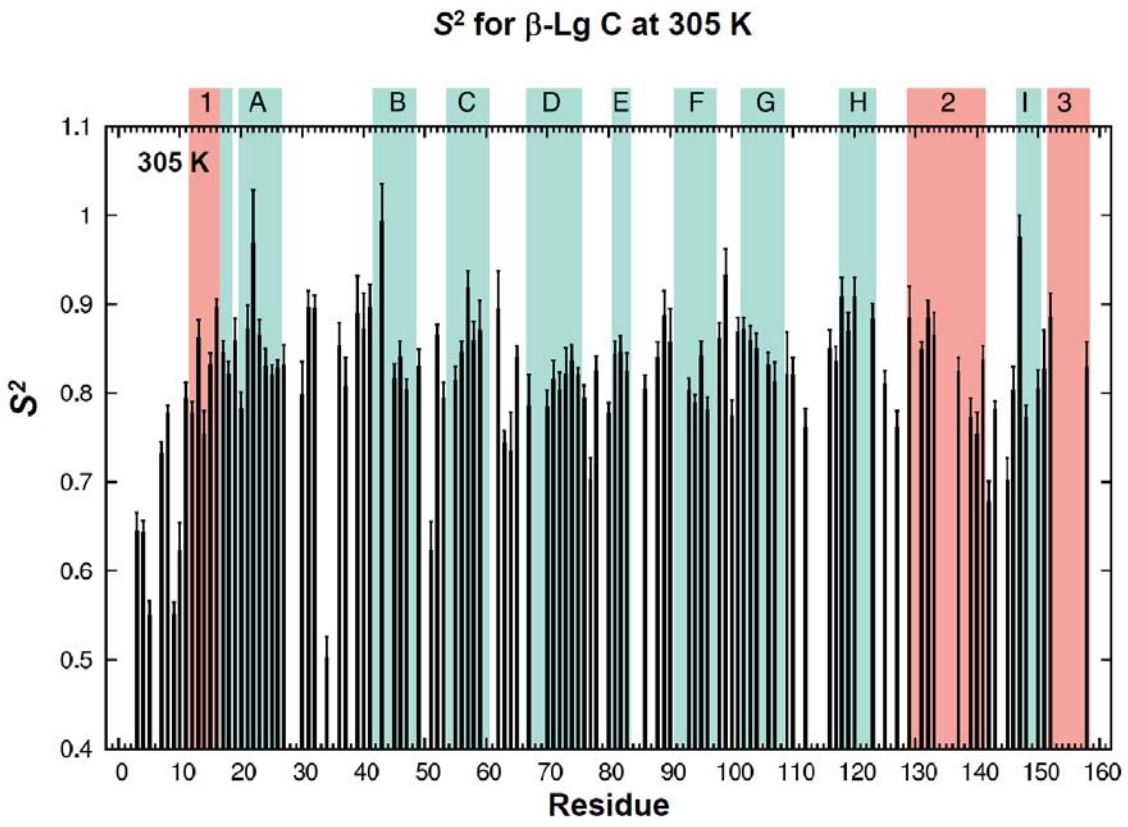


Figure 3.10 Order Parameters (S^2) *vs* Residue for Monomeric ^{15}N β-Lg C at 305 K.
Values are derived from the Lipari and Szabo model-free formalism (1982a, 1982b). In this plot the positions of the nine β-strands (labelled A-I) and α-helices (labelled 1-3) are highlighted with teal and salmon, respectively.

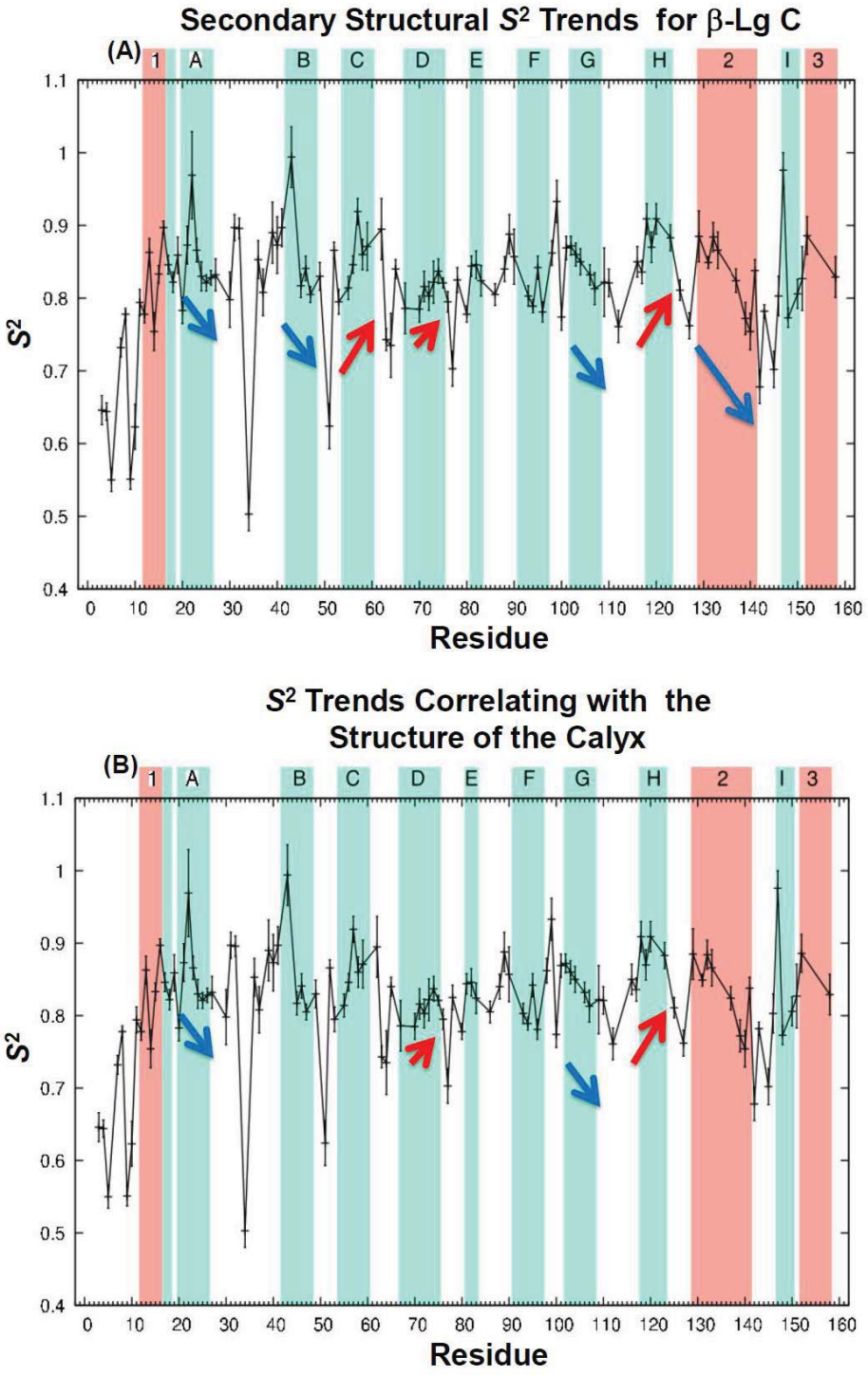


Figure 3.11 S^2 Trends Along Secondary Structural Elements of Monomeric β-Lg C.
Plots showing (A) trends pointing towards gradual increases and decreases in flexibility as defined by S^2 in secondary structural elements, and (B) plots indicating gradual increases and decreases in flexibility through the lengths of the β-strands as they progress into and away from the more open end of the calyx (β-barrel), respectively. Positions of the nine β-strands (labelled A-I) and α-helices (labelled 1-3) are highlighted with teal and salmon, respectively.

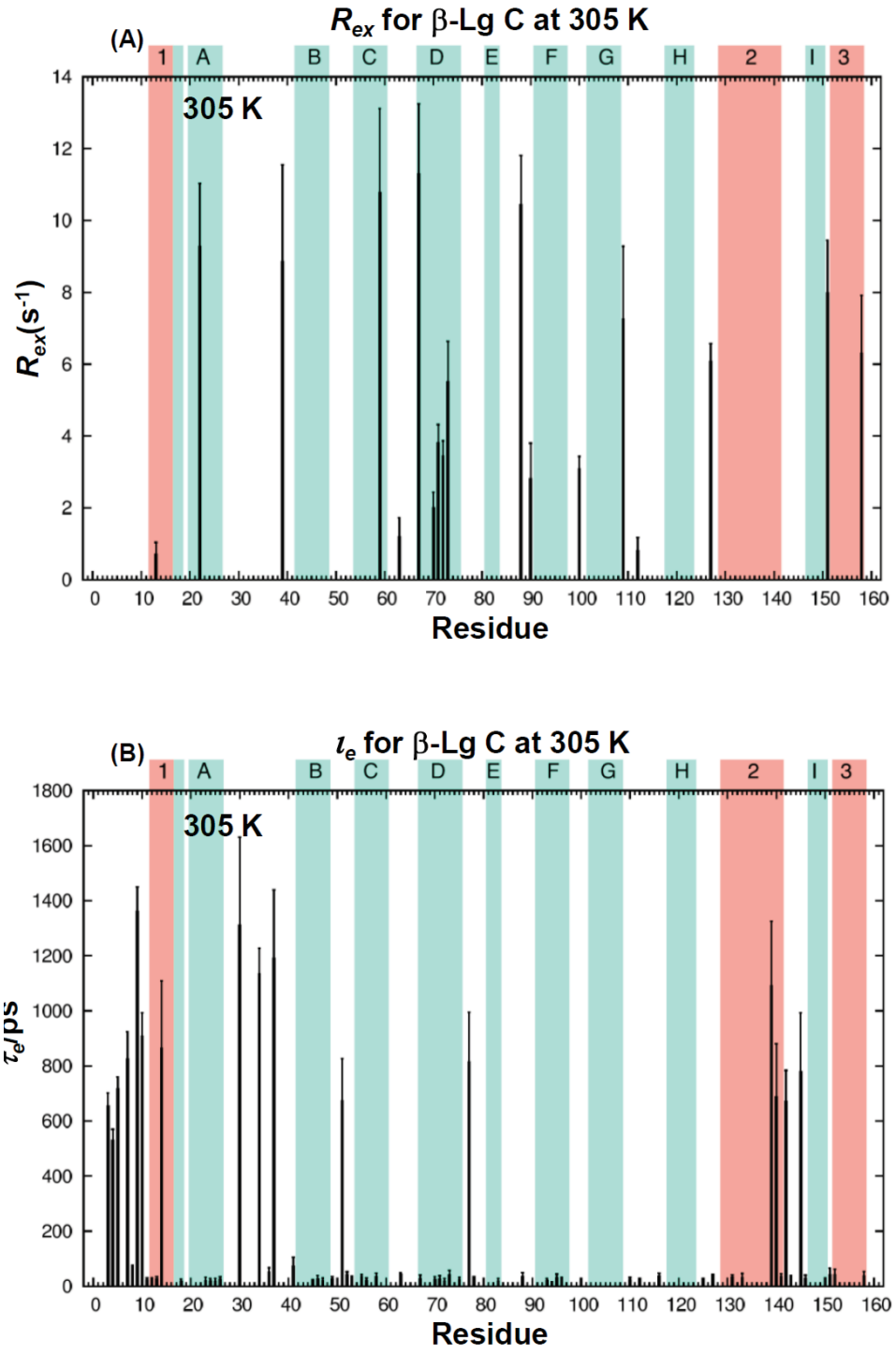


Figure 3.12 Conformational Exchange Terms and Internal Correlation Times for Monomeric β-Lg C.

Values in (A) R_{ex} and (B) τ_e are both plotted as a series of sequential plots. In both plots the positions of the nine β-strands (labelled A-I) and α-helices (labelled 1-3) are highlighted with teal and salmon, respectively.

3.6 Backbone Dynamics of β-Lg C at 305 K, 313 K and 320 K

To investigate the effects of temperature on the dynamics of bovine β-Lg, the ^{15}N relaxation rates for the C variant were recorded at 313 K and 320 K to complement the measurements described in section 3.4 at 305 K. All data were collected at a single magnetic field strength (700 MHz) and at pH 2.6 (where β-Lg is monomeric). As before, data were then fitted to derive the motional parameters of the protein backbone. Although it would have been interesting to record dynamics data at higher temperatures, 320 K is the upper limit of the cryoprobe and furthermore β-Lg tends to aggregate irreversibly if held at temperatures close to the denaturation point for the extended periods necessary to collect the relaxation data.

Relaxation data for 113 and 105 backbone amides of the 154 expected ^{15}N-^{1}H backbone resonances were obtained at 313 K and 320 K, respectively, to complement those for 110 residues obtained at 305 K, using the same methods as described in Section 2.4.5. The differing numbers reflect the changes in the number of overlapping peaks as the temperature is changed. Relaxation data were fitted to give the backbone dynamics parameters using the model-free method as described in Section 2.4.7. An overlay of ^{15}N relaxation parameters measured at 305 K, 313 K and 320 K are plotted in Appendix C.2 (A), (B) and (C). The model-free derived motional parameters at 305 K, 313 K and 320 K are listed in Appendices D.1-3, respectively. Initial estimates for the molecular rotational correlation time, τ_m, were derived at each temperature as described in Section 2.4.1. These estimates were subsequently refined by the model-free fitting procedure. The average relaxation rates and subsequent order parameters together with the optimised τ_m are listed in the following table (Table 3.3).

Temperature	R_1 (s^{-1})	R_2 (s^{-1})	NOE	τ_m (ns)	S^2
305 K	1.14 ± 0.07	12.71 ± 3.39	0.73 ± 0.07	8.54	0.82 ± 0.08
313 K	1.23 ± 0.06	11.60 ± 2.57	0.72 ± 0.07	7.97	0.81 ± 0.08
320 K	1.31 ± 0.06	10.83 ± 2.17	0.71 ± 0.10	7.24	0.83 ± 0.09

Table 3.3 Summary of the average relaxation and model-free dynamical parameters for β-Lg C measured at 305 K, 313 K and 320 K. The average R_1, R_2 and NOE experimental parameters, the rotational correlation times (τ_m), and average S^2 values were sampled at 305 K, 313 K and 320 K. An estimated standard deviation of scatter about the mean value was calculated and listed in the table.

As expected, the average R_1 rate increases, the average R_2 rate decreases and the average NOE enhancement value decreases slightly as the temperature is raised to 320 K. This is due to an increase in the tumbling rate of β-Lg C at higher temperatures, which is caused by the lowering of the solvent viscosity. As a function of residue number (Appendix C.2), similar peaks and troughs are observed, pointing towards similarities in dynamics, except for Gly64 in the NOE enhancement plot (C). Further investigation into the dynamics of this residue has been ruled out because an accurate measurement of its NOE enhancement could not be obtained, due to its weak signal at all three temperatures.

The overall correlation times were determined for recombinant β-Lg C using 305 K, 313 K and 320 K relaxation data sets independently. The initial τ_m estimates, derived from the relaxation experiments, were determined to be 9.0 ms, 8.0 ms and 7.2 ms at 305 K, 313 K and 320 K respectively. The model-free iteratively refined τ_m, at the three temperatures, are listed in Table 3.3 and plotted in Figure 3.13.

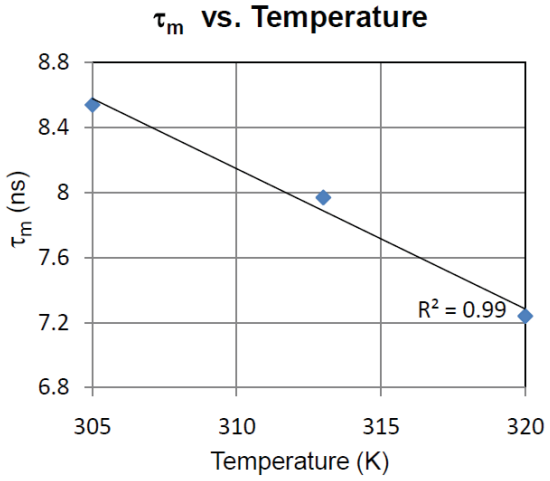


Figure 3.13 Overall Correlation Times for β-Lg C at Different Temperatures.
The values are derived from the ^{15}N relaxation data of β-Lg C at 305 K, 313 K and 320 K, using the model-free approach.

As the temperature increases from 305 K to 313 K and then to 320 K, τ_m decreases linearly over a range of 1.3 ns, showing that the increase in the tumbling rate is most likely due to the lowering of solvent viscosity and not to major changes in the protein's tertiary and quaternary structure. The overall order of the internal motions (S^2) does not change significantly, indicating that, in general, the backbone amides remain relatively rigid and that the internal motions are independent of the global tumbling rate of

β-Lg C, with a change in the latter accounting for most of the changes in the relaxation parameters.

3.6.1 Model Selection for Bovine β-Lg C at 305 K, 313 K and 320 K

The number of residues for each of the optimal models selected by the model-free fitting procedure for β-Lg C is shown in Table 3.4.

Temperature	**Model One** S^2_s	**Model Two** $S^2_s\ \tau_e$	**Model Three** $S^2_s\ R_{ex}$	**Model Four** $S^2_s\ \tau_e R_{ex}$	**Model Five** $S^2_s\ \tau_e\ S^2_f$	**No fit**
305 K	39 residues	37	5	13	16	0
313 K	7	46	1	19	37	2
320 K	18	24	3	24	23	13

Table 3.4 Distribution of the model fits for β-Lg C at 305 K, 313 K and 320 K.

Differences were identified for the distribution of the model-free fits at the three temperatures sampled. At the lowest temperature, simpler models (models one and two), were predominant in fitting the relaxation data. At increased temperatures, an increase in models incorporating the conformational exchange term R_{ex} (models three and four) were selected as best fits for the backbone amides, indicating that although the overall flexibilities of the ^{15}N-^{1}H bonds remain similar (Table 3.3), the additional slower motion is becoming more prevalent with increasing temperatures. An increase in the number of no fits as the temperature progresses to 320 K, suggests that the internal motions within β-Lg C are becoming more complicated as the temperature is raised. Dynamics highlighted by the two time-scale spectral density function are not analysed in the following comparison. Reasons for this are discussed in the following section.

3.6.2 The Model-Free Parameters at 305 K, 313 K and 320 K

Order Parameters (S^2)

The model-free motional parameters for β-Lg C, derived from ^{15}N relaxation values measured at 305 K, 313 K and 320 K are listed in Appendices D.1-D.3, respectively. The model-free derived order parameters for these experiments are plotted separately in Figure 3.14 (A), (B) and (C), and a comparison of these parameters is shown in Figure 3.15.

Looking across β-Lg C's backbone S^2 variations are very similar when comparing trends among the three model-free data sets (Figure 3.15). These results generally indicate that the amplitudes of motions of the ^{15}N-^{1}H bonds do not change significantly,

showing similar gradual variations of flexibility across secondary structural elements and increased flexibility for backbone amides positioned in the N-terminus and for backbone amides positioned in loop regions. Ala34, right in the middle of A/B loop, still displays a significantly low order parameter at 313 K and 320 K.

When the S^2 trends of the different temperature data sets are compared, two subtle changes in flexibility are detected. The first being that when the temperature is raised from 313 K to 320 K the ^{15}N-^{1}H vector of Ser30, located in the region of the 3_{10}-helix within the A/B loop, becomes more constrained as indicated by a higher S^2 value (Figure 3.15). This implies that Ser30 becomes more associated with another element, when the temperature is raised to 320 K, restricting its motion. Differences in the values of S^2, among the different temperature data sets, are also observed across the major α-helix (α-2), suggesting that this helix (Asp129 - Lys141) becomes more rigid between 305 K and 320 K. These results also imply that the major α-helix becomes more associated with another nearby region/s of β-Lg, as the temperature is raised. However, ^{15}N relaxation data is not available for the centre region of the helix for a full S^2 assessment, as seen by the absence of values for this region in Figure 3.14 (A), (B) and (C), to confirm that this was happening for the entire region.

These results show the flexibility of internal motions across the backbone of β-Lg C are all remarkably similar at 305 K, 313 K and 320 K, except for two subtle changes in the A/B loop and the α-2 helix, providing evidence that these regions are becoming more associated with other regions in β-Lg C. This lack of large changes in flexibility observed is perhaps not unexpected, when changes in amide protection were only observed for β-Lg A above 55 ˚C (Edwards *et al*., 2002).

Conformational Exchange (R_{ex})

Comparing values of R_{ex} between temperature sets is not considered in this comparison, as R_{ex} is only determined at one static field strength, which can compromise the precision of the resulting values. By measuring R_{ex}, at least at two different field strengths, a more precise value for R_{ex} can be determined (Millet *et al.*, 2000), but was not possible within the time-frame of this project. However, a qualitative discussion of the frequency of R_{ex} terms was still achievable.

The motional parameters of the model-free analysis indicate that β-Lg C displays higher numbers of millisecond time-scale motions as temperatures are increased from 305 K to

320 K. The total number of residues revealing R_{ex} terms increases from 18 at 305 K to 27 at 320 K (Table 3.4). Not surprisingly, at all three temperatures tested, the conformational exchange constant, R_{ex}, is significant for some of the backbone amides that are located in loops and link regions (Figure 3.16). A higher occurrence of these motions are found in loops that are positioned at the open end of the barrel (A/B, C/D and E/F loops), and a lesser occurrence in loops located at the more closed end of the barrel (B/C, D/E, and F/G loops).

The R_{ex} constant for Leu22 shows that its critical position at the hinge point of the β-A strand, is still maintained when the temperature is raised to 313 K and then at 320 K. His59, located at substitution site between variant B and variant C, displays a relatively high R_{ex} value at 313 K, but unfortunately ^{15}N relaxation data measured at 320 K could not be fitted to define its motions at this temperature. The highest density of R_{ex} values observed at 305 K, are concentrated on the β-D strand at 305 K (Figure 3.16 (A)). The model-free parameters at increasing temperatures describe this β-strand as still experiencing these slower motions, which is not unexpected.

In comparison, ^{15}N relaxation data from the β-E strand residues had optimal fits using simpler models at lower temperatures (models one and two for 305 K and 313 K respectively); whereas at 320 K, relaxation data from backbone amides in this region were fit with the more complex model four. These results suggest that the only two H-bonds formed between strands β-D (Glu74) and β-E (Lys83) (Qin *et al*., 1998b, Edwards *et al*., 2002) are disrupted at temperatures between 313 K and 320 K. The hypothetical 'freeing' of the β-E strand would contribute to its conformational mobility at 320 K. The relative length of the long β-D strand, coupled to its hydrogen bonding potentials not being completely fulfilled, are plausible explanations for why this β-strand is undergoing conformational exchange at the lower temperatures, before the possible disruption of the two H-bonds between 313 K and 320 K.

As noted previously, the interactions formed between the β-G and β-H strands are highly ordered, with a dense network of inter-strand H-bonds and a disulfide bridge connecting these strands at positions Cys106 and Cys119. At 305 K, the measured ^{15}N relaxation data from these strands are fitted with the simplest model (model one), which is not uncommon for residues located in secondary structural elements. However, additional slower motions (ms - μs) are detected as the temperature increases to 313 K, and become more prevalent at 320 K. These data suggest that that these two strands

become more conformationally mobile due to an increase in temperature, which possibly contributes to a stretching motion of these β-strands in the regions closest to the open end of the barrel or caused by the disruption of H-bonds formed by residues located in these areas.

Correlation Times (τ_e)

As stated earlier, the effective correlation times (τ_e) are imprecisely determined, but are useful for spotting mobile regions of the backbone (Palmer, 1993). As seen in Figure 3.17, when the temperature is raised to 313 K (Plot B) from 305 K (Plot A), residues with significant τ_e contributions (> 500 ps) become more prevalent in the A/B loop, the B/C loop, D/E loop G/H loop, the link regions around the major α-helix, β-F strand and the N- and C-terminal end of the major α-helix (α-2). These results indicate that many of the loop regions become more mobile when the temperature is raised to 313 K. These trends are still seen at 320 K (Figure 3.17 (C)), but in some instances, ^{15}N relaxation data could not be fitted with a model or they were discarded from dynamical analysis as their assigned peaks were merging with others in the ^{15}N,^{1}H-HSQC spectrum, so therefore, are absent in the 320 K plot.

Previously, when comparing the S^2 values with increasing temperatures, it appears that the α-2 helix (major α-helix) becomes more rigid when the temperature is raised, therefore, the presence of nanosecond time-scale motions with relatively small amplitudes are assumed. In this case, model-free analysis can significantly under-estimate the effective correlation time (Chen *et al.*, 2004), particularly for residues Leu133 and Ala139, whose S^2 values are highest at 320 K, but display longer effective correlation times when temperature is increased to 313 K, but become shorter again at 320 K.

Two Time-Scale Spectral Density Function (Model Five)

A comparison at different temperatures of the extended model-free two time spectral density function is not considered in these studies, as ^{15}N relaxation measurements need to be acquired at a number of magnetic field strengths to assess the quality of the model-free fitting, which is impossible to determine by statistical tests alone because an additional parameter has been introduced (Rule & Hitchens, 2006), making a comprehensive comparison difficult to achieve.

Changes in Chemical Shifts

NMR spectroscopy chemical shifts are very sensitive to the local environment of the residue's atoms. Increases in temperature do not generally change the dispersion of peaks in the $^{15}N,^{1}H$-HSQC spectrum, showing that the average conformation of the protein's structure is similar at all three temperatures (Figure 3.6). However, it is possible that minor conformational changes occur at some sites when the temperature is raised to 320 K, as detected by chemical shift perturbation of peaks in the ^{15}N dimension between 305 K and 320 K (Figure 3.18). Residues positioned in the N-terminus (Val3, Thr4, Gln5, Lys8, and Gly9) show significantly higher than average changes in their backbone amide ^{15}N chemical shifts (> 0.35ppm). However, this is not observed for Met7, which forms an H bond with Val94 (Qin *et al*., 1998b, Uhrínová *et al*., 2000). Significantly higher than average changes in ^{15}N chemical shifts are also observed for Tyr18 and Leu22 (β-A), Glu42 (β-B), Asn63 (C/D), Ala67 and Ile71 (β-D). Notably, Tyr18, Glu42, Ala67 and Ile71 are all positioned in β-strands that form one β-sheet, which makes up half of the β-barrel and Leu22 is positioned at the midpoint of both β-sheets. This β-sheet is less ordered than the second β-sheet, which may account for its higher frequency of residues with ^{15}N chemical shift changes greater than 0.35 ppm, causing subtle changes in conformation as the temperature is increased. The second β-sheet is formed by β-strands that possesses a more extensive inter-strand H-bond network and is further stabilised by a disulfide bridge. Changes in ^{15}N chemical shifts for Glu131 positioned in the α-2 helix and Glu158 positioned in the ill-defined α-3 helix, also suggest subtle changes in conformation when temperature is raised to 320 K.

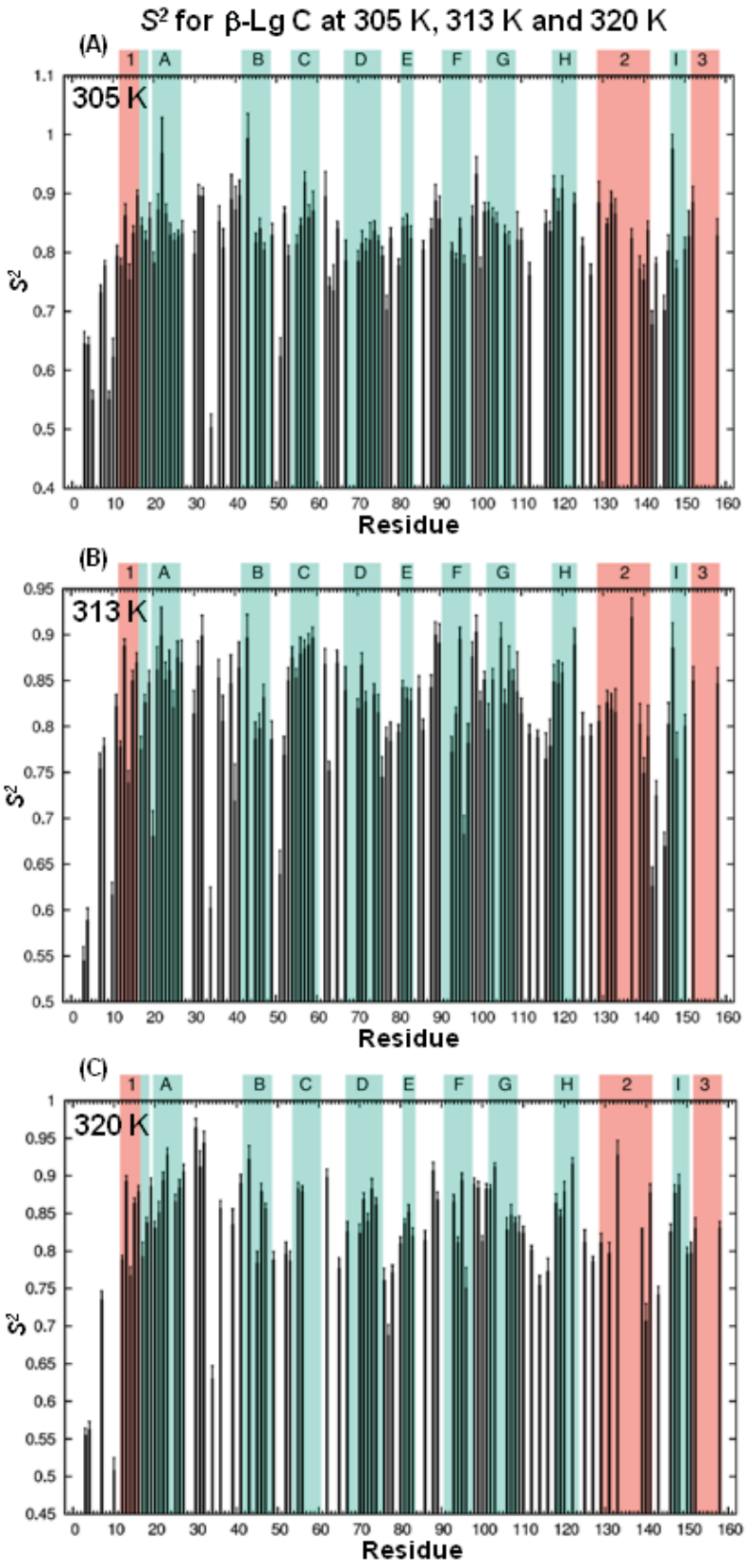


Figure 3.14 S^2 vs. Residue for Monomeric β-Lg C at 305 K, 313 K and 320 K.

S^2 vs. residue for monomeric β-Lg C sampled at (A) 305 K, (B) 313 K and (C) 320 K. The positions of the nine β-strands (labelled A-I) and α-helices (labelled 1-3) are highlighted with teal and salmon, respectively.

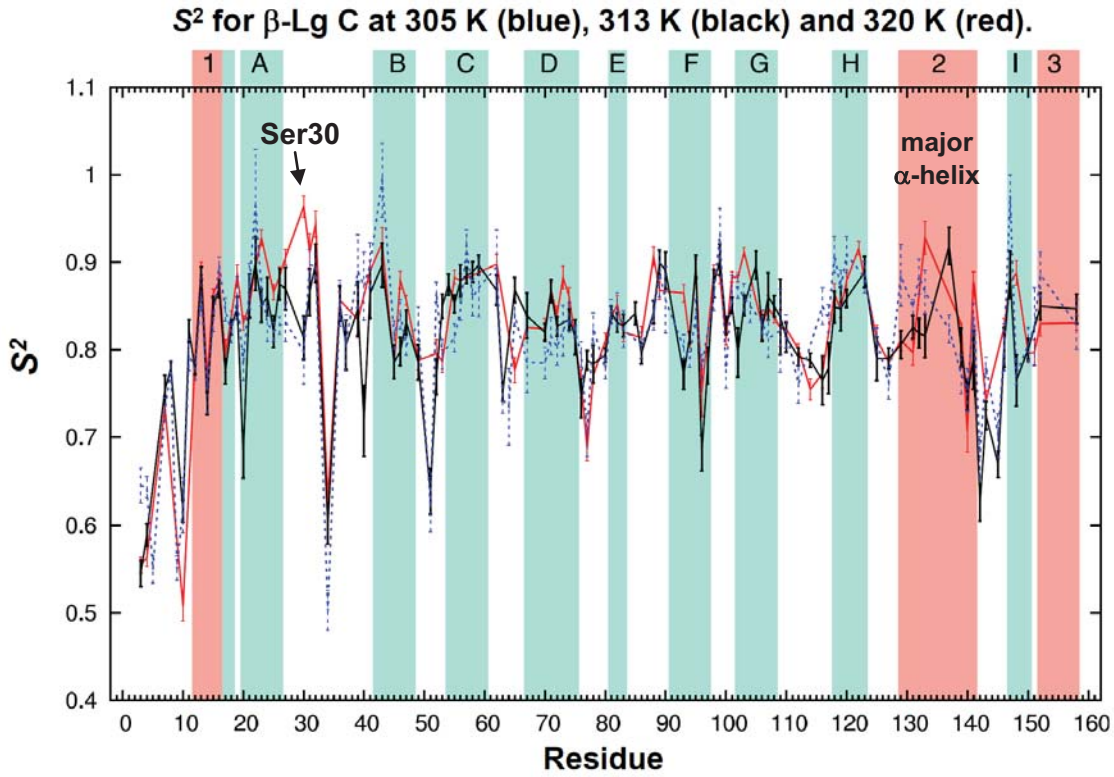


Figure 3.15 An Overlay of S^2 Traces for β-Lg C at 305 K, 313 K and 320 K.

Values are derived from the Lipari and Szabo model-free formalism (1982a, 1982b) from ^{15}N relaxation parameters measured at 305 K (blue), 313 K (black) and 320 K (red). In this plot the positions of the nine β-strands (labelled A-I) and α-helices (labelled 1-3) are highlighted with teal and salmon, respectively.

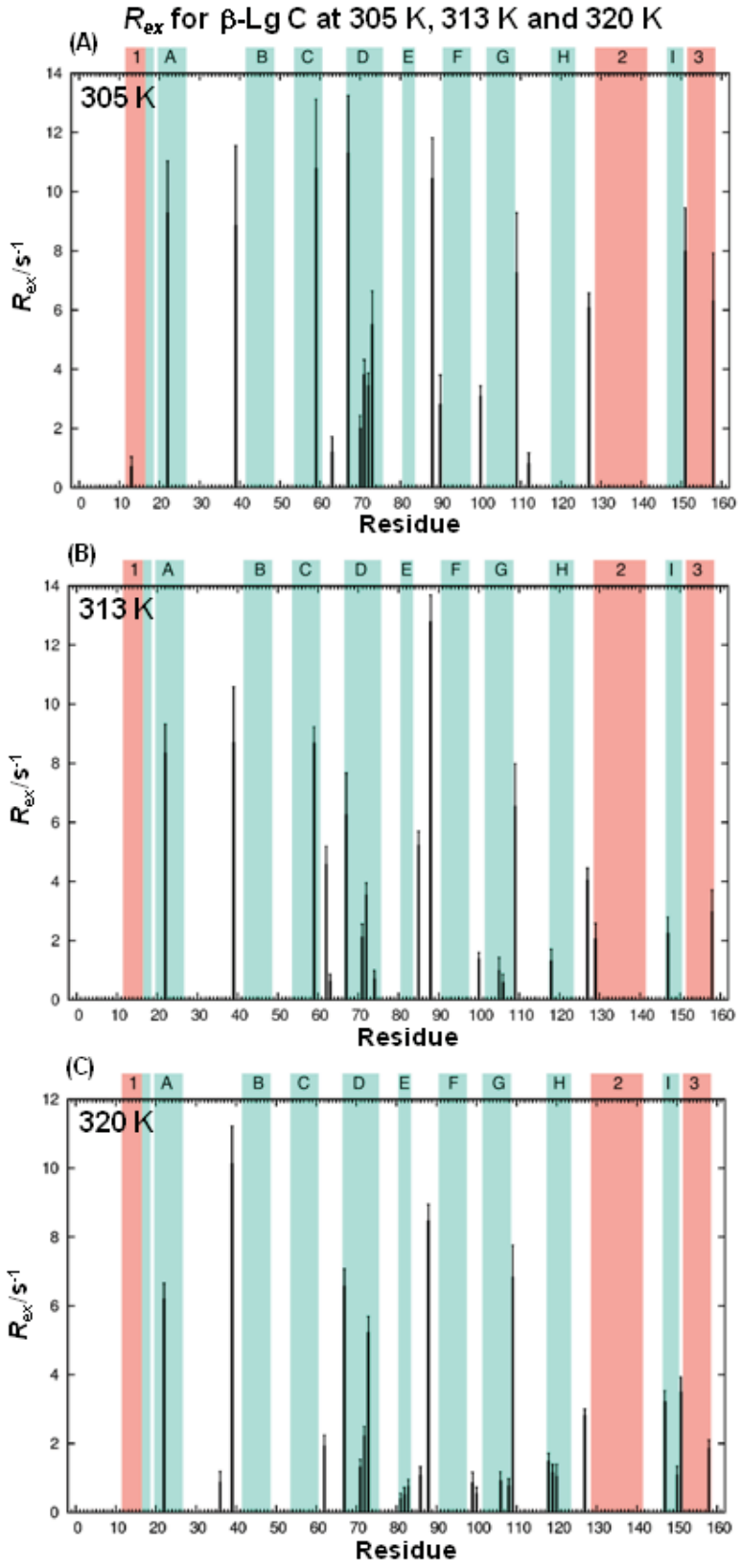


Figure 3.16 R_{ex} vs. Residue for β-Lg C Sampled at 305 K, 313 K and 320 K.

R_{ex} vs. residue for β-Lg C sampled at (A) 305 K, (B) 313 K and (C) 320 K. In both plots the positions of the nine β-strands (labelled A-I) and α-helices (labelled 1-3) are highlighted with teal and salmon, respectively.

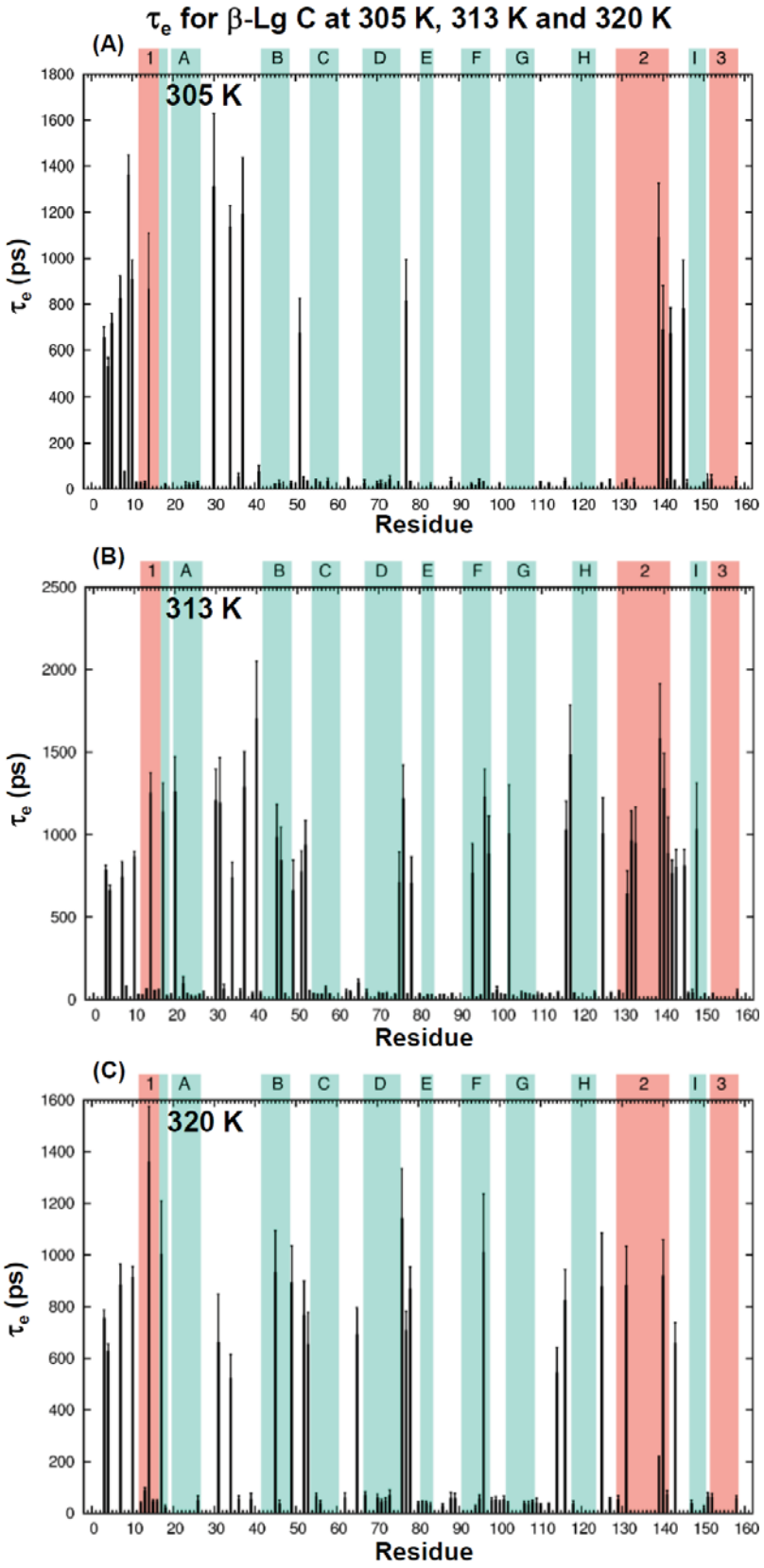


Figure 3.17 τ_e vs. Residue for Monomeric β-Lg C at 305 K, 313 K and 320 K.

The motional parameter τ_e (ps), plotted vs. residue number at (A) 305 K, (B) 313 K and (C) 320 K. In both plots the positions of the nine β-strands (labelled A-I) and α-helices (labelled 1-3) are highlighted with teal and salmon, respectively.

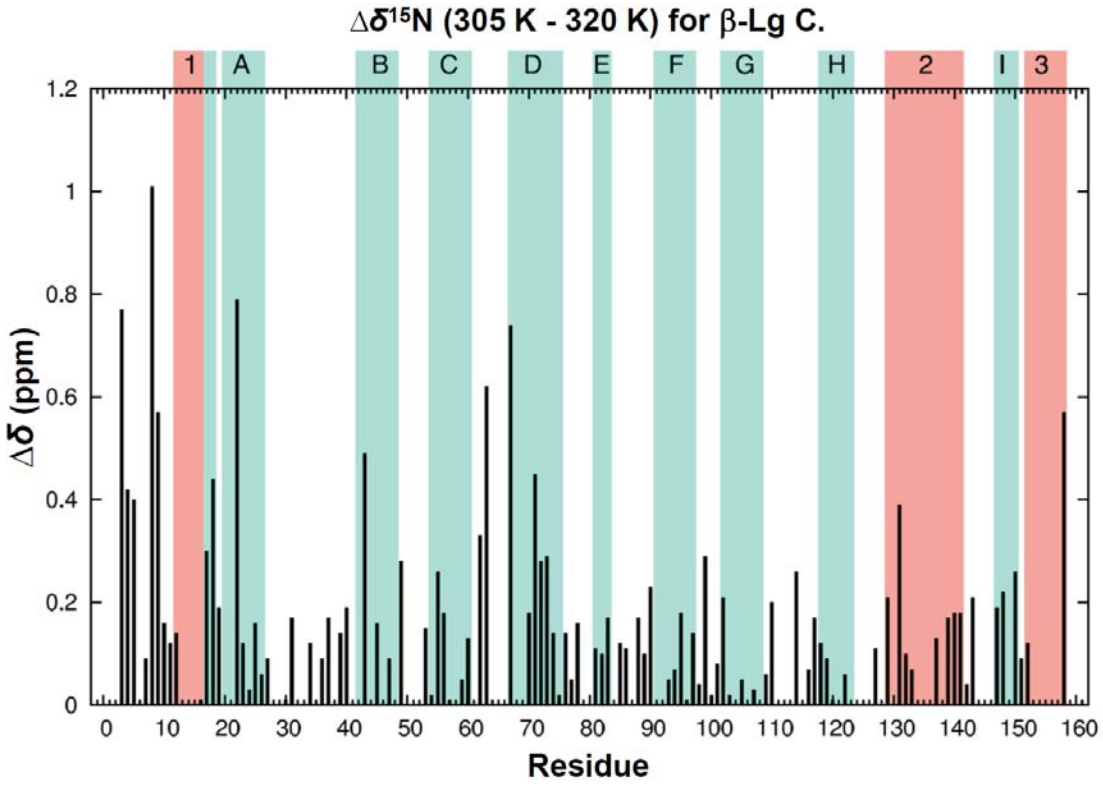


Figure 3.18 Changes in ^{15}N Chemical Shifts for β-Lg C between 305 K and 320 K.

Values are calculated from the differences in chemical shifts of residues in the ^{15}N-dimension between 305 K and 320 K. In this plot the positions of the nine β-strands (labelled A-I) and α-helices (labelled 1-3) are highlighted with teal and salmon, respectively.

3.7 Assigning the Backbone of β-Lg Variants A and B

Single-labelled (^{15}N) β-Lg A and β-Lg B samples were used to assign as many peaks as possible in the respective ^{15}N,^{1}H-HSQC spectra. NMR spectroscopy backbone assignments experiments were performed at pH 2.6, using a 700 MHz spectrometer. Backbone assignments for the A and B variants, achieved by means of ^{13}C/^{15}N triple-resonance experiments, were already available (Dr. Patrick J. B. Edwards; unpublished data). However, these assigned proteins had been expressed in the non-reproducible *E. coli* system (Ariyaratne *et al*., 2002). Recombinant ^{15}N-labelled β-Lg A and B in this investigation had been provided by Dr. Komala Ponniah, which had been over-expressed in the reproducible *E. coli* Origami B (DE3) [pETDuet-DsbC-BLG] system. The assigned ^{1}H and ^{15}N shift lists from the former system, were transferred to the spectra (^{15}N,^{1}H) of the latter system and then the assignments were checked for consistency with those published by Uhrínová *et al*. (1998), using a set of 3D ^{15}N,^{1}H- TOCSY-HSQC and ^{15}N,^{1}H-NOESY-HSQC experiments. The assigned ^{1}H and ^{15}N backbone chemical shifts for the three variants at 305 K are listed in Appendix B.2. The ^{15}N,^{1}H-HSQC spectra for variants A, B and C are shown in Figure 3.19. This figure points to peaks where there are variations due to substitutions in the polypeptide sequence. Sequence-specific ^{1}H and ^{15}N resonance assignments were completed for 137 of 163 total residues for recombinant β-Lg A and for 136 of 163 residues for recombinant β-Lg B (84 % and 83 %, respectively).

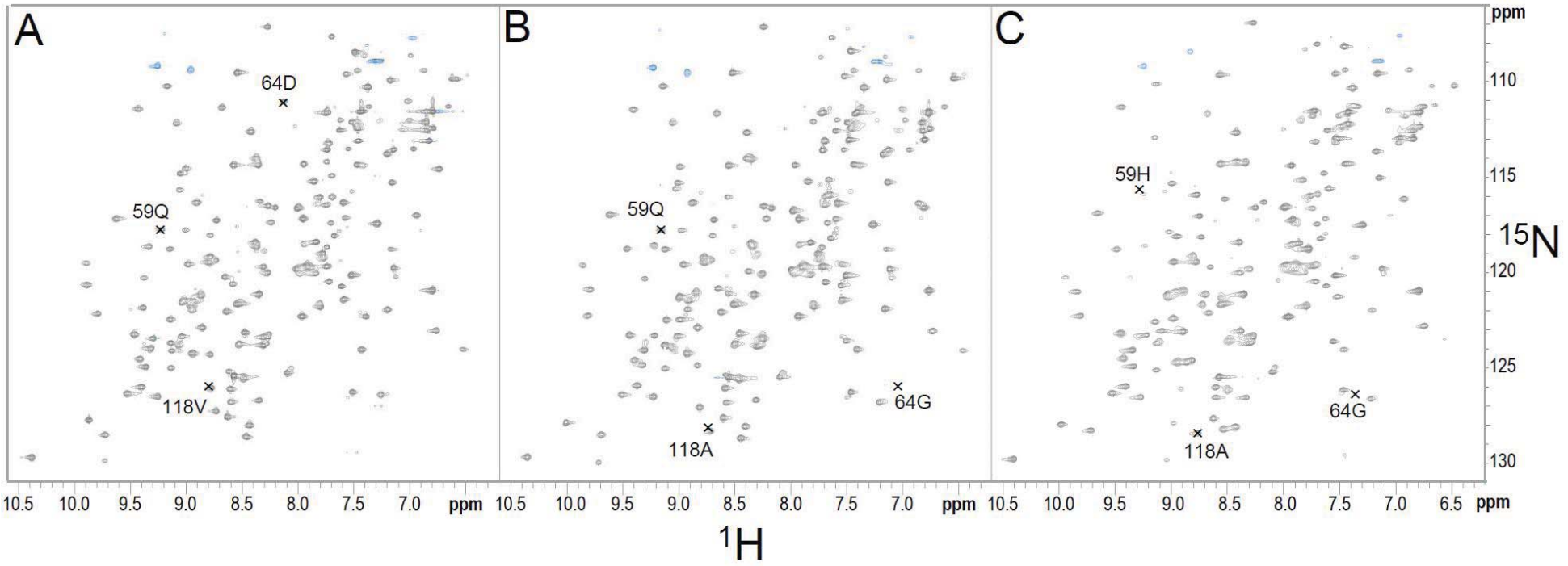


Figure 3.19 ^{15}N,^{1}H-HSQC Spectra of ^{15}N-Labelled β-Lg Variants Sampled at 305 K and at pH 2.6.

(A) β-Lg variant A, (B) β-Lg variant B and (C) β-Lg C. HSQC spectra were collected on 1 mM protein samples in 50 mM KH_2PO_4 buffer. Labelled cross-peaks represent the three polymorphic amino-acid substitutions present among the three variants. Spectra were assigned as described in Sections 2.4.2 and 2.4.3. Grey indicates positive contours whereas sky-blue represents negative contours. ^{1}H and ^{15}N chemical shifts for the assigned residues for the A, B and C variants are listed in Appendix B.2.

3.8 Comparing Dynamics of β-Lg A, B and C at 305 K

To investigate the effect of polymorphisms on the backbone dynamics of purified β-Lg, ^{15}N relaxation data was collected for the backbone amides of β-Lg A, B and C variants separately at 305 K. All data were collected at one magnetic field strength (700 MHz) and at pH 2.6 (where the β-Lg variants are monomeric). After the measured relaxation data were processed, by means of the model-free protocol, motions defined by the parameters S^2 and R_{ex} were used to probe backbone dynamics on the fast ns to ps time-scale and used to detect additional slower motions on the ms to μs time-scale, respectively. In order to detect similarities and differences in dynamics among the natural variants, interpretation of backbone dynamics, using these two model-free parameters, is the primary focus in this section. Of particular interest in these studies are the differences in dynamics, which are due to external and internal amino acid substitutions (see below).

3.8.1 Assessment of Residues in Close Proximity to the Substitution Sites

X-ray crystallography-derived structures of the A, B and C β-Lg variants show that at neutral pH these proteins are structurally homologous except for a few minor alterations around the substitution sites (Section 1.8.2) (Qin *et al*., 1999, Bewley *et al*., 1997). Even though complete 3D NMR spectroscopy derived structures of the B and C variants are not available, the fingerprints of all three ^{15}N,^{1}H-HSQC spectra suggest that these proteins adopt a very similar tertiary conformation at pH 2.6. Upon careful examination of the 3D NMR spectroscopy derived structure of β-Lg A, at low pH (Uhrínová *et al*., 2000), by means of PyMOL (Delano, 2008) using the protein data base coordinates 1DV9, amino acids in close proximity (within 6 Å) to the substitution sites in the A, B and C variants were identified.

In the case of the **Asp64Gly** substitution (**A→B**), residues in close proximity to this position are not clearly defined as it sits in the C/D loop; however, its neighbouring residues include Asn63 and Glu65. The **Val118Ala** substitution site, which sits at the beginning of strand β-H near the more open end of the barrel, is the second site of substitution between variants A and B. Close residues include Ala26 (β-A strand), Ser27 (N-terminal end of A/B loop), Cys106 and Met107 (β-G strand) and neighbours Leu117 (G/H loop) and Cys119 (β-H strand). In the instance of the **Gln59His** substitution (**B→C)**, nearby residues include Tyr42 and Val43 (β-B strand), neighbours Leu58 and Lys60 (β-C strand) and residues Ala67, Gln68 and Lys69 (β-D strand).

3.8.2 Relaxation Measurements for Bovine β-Lg Variants, A, B and C at 305 K

Relaxation data for 131, 115 and 110 backbone amides of the 154 expected ^{15}NH backbone resonances were obtained for variants A, B and C respectively. The following table (Table 3.5) lists average relaxation values, average order parameters and respective average errors of 84 residues, with measured relaxation data in all three variants (i.e. values from Val3 were excluded in the averages, as ^{15}N relaxation data were only available for variants A and C, etc). The model-free iteratively refined overall correlation times, τ_m, are also included.

Variant	R_1 (s^{-1})	R_2 (s^{-1})	NOE	τ_m (ns)	S^2
A	1.14 ± 0.03	11.94 ± 0.29	0.73 ± 0.01	8.70	0.80 ± 0.01
B	1.14 ± 0.05	12.73 ± 0.53	0.76 ± 0.07	8.88	0.82 ± 0.02
C	1.14 ± 0.04	12.64 ± 0.48	0.74 ± 0.03	8.54	0.82 ± 0.02

Table 3.5 Summary of the average R_1, R_2 and NOE values and average order parameters (S^2) for residues in β-Lg A, B and C variants at 305 K.

The average ^{15}N relaxation rates and the average order parameters all fall with the margin of error, pointing towards similarities in ^{15}N relaxation and backbone dynamics among the variants. The model-free iteratively refined overall tumbling rate varies over a narrow range of 0.34 ns, showing slight differences among the variants in their time to rotate around one radian in solution. However, the internal motions are independent of these subtle differences as all the average S^2 values fall within the margin of error. There is not enough information at this stage to state what contributes to this small difference. The dispersion of peaks in the ^{15}N,^{1}H-HSQC are similar, showing that the variants have a similar fingerprint, providing evidence that the variants have an overall similar conformation, under these conditions. Also, the protein samples are highly pure, therefore, effects of contamination can be ruled out.

For each of the analysed β-Lg variants (A, B and C at 305 K), the values and standard errors for each residue with measured ^{15}N relaxation data (R_1, R_2 and NOE) obtained, are presented as plots in Appendix C.3. Looking at the relaxation overlays in Appendix C.3, R_2 relaxation trends (Plot (B)) correlate better with variants A and B, than variant C. Looking at the differences and similarities in this plot, the variants share commonalities, but there are some noticeable differences detected for β-Lg C. Except for Gln59His at the B→C substitution site and for Val43 in close proximity to this substitution site, other sites of substitution or residues in close proximity to substitution sites are not where notable differences are detected. Other points include Leu39 (A/B loop), Glu62

(C/D loop), Ile71 - Ala73 (β-D strand), Asn88 (E/F loop), Asn109 (G/H loop) and Phe151 (link between β-I and α-3). Relaxation data for these residues have been interpreted in Section 3.4.1, and are all located in regions where conformational exchange could be expected. Relaxation rates were measured from protein made by two different people, the C variant was made in these investigations and the A and the B variants were made two years prior. It is not known whether the fresh stock of protein and/or slight differences in the sample preparation increased the sensitivity for R_2 relaxation in β-Lg C. However, R_1 relaxation (Plot (A)) correlates well for all three variants indicating that the motions on a ns to ps time-scale and the correlation times for the internal motions are similar. The $\{^1H\}$-^{15}N NOE enhancement overlay (Plot (C)), indicates there may be differences in dynamics in the G/H loop (Asn109 – Leu117). Although it may have been preferable to prepare fresh A and B samples, the model-free analyses still proceeded with the available samples due to time constraints on the project. Even so, shared relaxation trends and average R_2 values were within the margin of error across the three variants. Note that Ala111 has been removed from analyses due to inaccuracies of its NOE value.

3.8.3 Model-Free Fits for Bovine β-Lg A, B and C at 305 K

As before, the relaxation data for each variant were fitted with one of the five standard mathematical models of the extended Lipari-Szabo analysis. A significant similarity is present in the number of best-fit models for the three variants (Table 3.6), except that the number of no fits is 8 for variant B and 12 for variant A.

Variant	Model One	Model Two	Model Three	Model Four	Model Five	No fit	Total
	S^2	S^2 - τ_e	S^2 - R_{ex}	S^2 - τ_e - R_{ex}	S^2 - τ_e - S_f^2		
A	41 residues	39	4	10	25	12	131
B	43	37	1	9	17	8	125
C	39	37	5	13	16	0	110

Table 3.6 Distribution of the model-free fits for monomeric β-Lg A, B and C variants sampled at 305 K.

3.8.4 Model-Free Parameters for β-Lg A, B and C at 305 K

Order Parameter (S^2)

To spot similarities and differences in flexibility of backbone motions (ns – ps) across the lengths of the variants, S^2 distributions of variants A, B and C were analysed. The order parameters of the structurally homologous monomeric β-Lg A and B variants display

very similar variations across the peptide backbone compared to β-Lg C at 305 K (Figure 3.20 (A)). No significant S^2 differences were detected in the G/H loop, positioned next to the Val118Ala substitution site, which had been suggested with preliminary dynamical investigations using slightly impure protein and with the ^{15}N NOE enhancement data in these investigations. In this region, differences in S^2 values all fall in with the margin of error. These results show that the insertion of a shorter side chain at the Val118Ala substitution site (A→B) into the hydrophobic cavity does not seem to have an effect on the ns – ps backbone dynamics of β-Lg. These studies also show that an altered charge distribution at the Asp64Gly substitution site does not have a significant effect on the internal flexibility of backbone motions displayed by variants A and B (no data for C), nor does the Gln59His substitution (B→C) near the edge of the β-C strand. It should be noted that this carboxylate side chain, along with all other carboxylate side chains are protonated and neutral carboxylic acids at pH 2.6. However, for the Gln→His substitution, there is an increase in protein charge at pH 2.6. The S^2 values for ^{15}N spins of residues positioned at these substitution sites or in close proximity to the substitution sites do not differ significantly. For other residues, where there looks like there are differences in S^2 (Glu51, Gln115 and Thr154), ^{15}N relaxation rates have not been measured for all of the three variants, producing differences in these regions of the plot, due to lack of data points. In general S^2 trends of the variants look similar with the overall correlation of the A, B and C variants' ^{15}N R_1 and NOE relaxation data supporting this. It must be noted that the effects of substitutions on β-Lg's backbone flexibility were only assessed at one temperature (305 K), which does rule out differences in dynamics due to amino acid substitutions at higher temperatures.

Conformational Exchange (R_{ex})

To establish if there are variations of additional slower motions (ms to μs), R_{ex} frequencies for all three the variants were compared. Commonalities in the frequency for R_{ex} were observed for Leu22, positioned at the hinge region of the β-A strand, and Ala67, positioned in the β-D strand, whose H-bond potentials are not completely fulfilled, showing that conformational exchange at these sites is shared by the variants. Interestingly, the R_{ex} value at the site of the Gln69His substitution in variant C suggests that this residue is subject to conformational exchange, whereas at this position in variants A and B this is not observed. However, these results do not confirm that additional slower motions at this substitution site are specific to the C variant, which have been observed at 305 K and 313 K (no relaxation data available at 320 K), because of the

variability of the ^{15}N R_2 relaxation data noted previously (Appendix C.3 (B)). These results only show that further investigations are needed to draw a comparison. Conformational exchange is not detected for Val43, which had been inferred tentatively from its ^{15}N R_2 relaxation data. Unsurprisingly, other differences in the number of R_{ex} terms (Figure 3.20 (B)) are not reflected at positions in close proximity to the substitution sites, but at sites where there is variability in the R_2 relaxation data among the variants, which has been noted, as the detection of conformational exchange in most cases contributes to higher values of R_2 (Section 2.4.1). The additional R_{ex} terms detected in β-Lg C do not seem unreasonable, due to the positions these residues hold within β-Lg's structure (Section 3.5.3), which include the loops at the more open end of the β-barrel and the β-D strand, whose hydrogen bonding potentials are not fulfilled. It is more likely that subtle changes in the purification procedure and/or the freshness of the protein solutions enhance the sensitivity for conformational exchange. However, this has not been verified.

It might be conjectured, that being stored in the -20 ˚C freezer for two years may have affected the quality of the A and B samples. However, the ^{15}N,^{1}H-HSQC spectra in Figure 3.19 does not support this idea, as no additional peaks appear that point to degradation of these proteins. It is unclear whether or not subtle changes in the purification protocol, caused the differences for R_2 relaxation and hence conformational exchange in β-Lg C. However, to come to more conclusive statements in regards to differences in dynamics, fresh stocks of purified β-Lg A and β-Lg B proteins need to be tested again, this time using the exact same purification protocol. It would also be interesting to repeat these experiments at 320 K and higher, temperatures where near-UV-CD investigations have previously implied a difference in configuration of the Cys66-Cys106 disulfide bridge due to the Gln59His substitution found in the C variant (Manderson *et al*., 1999a).

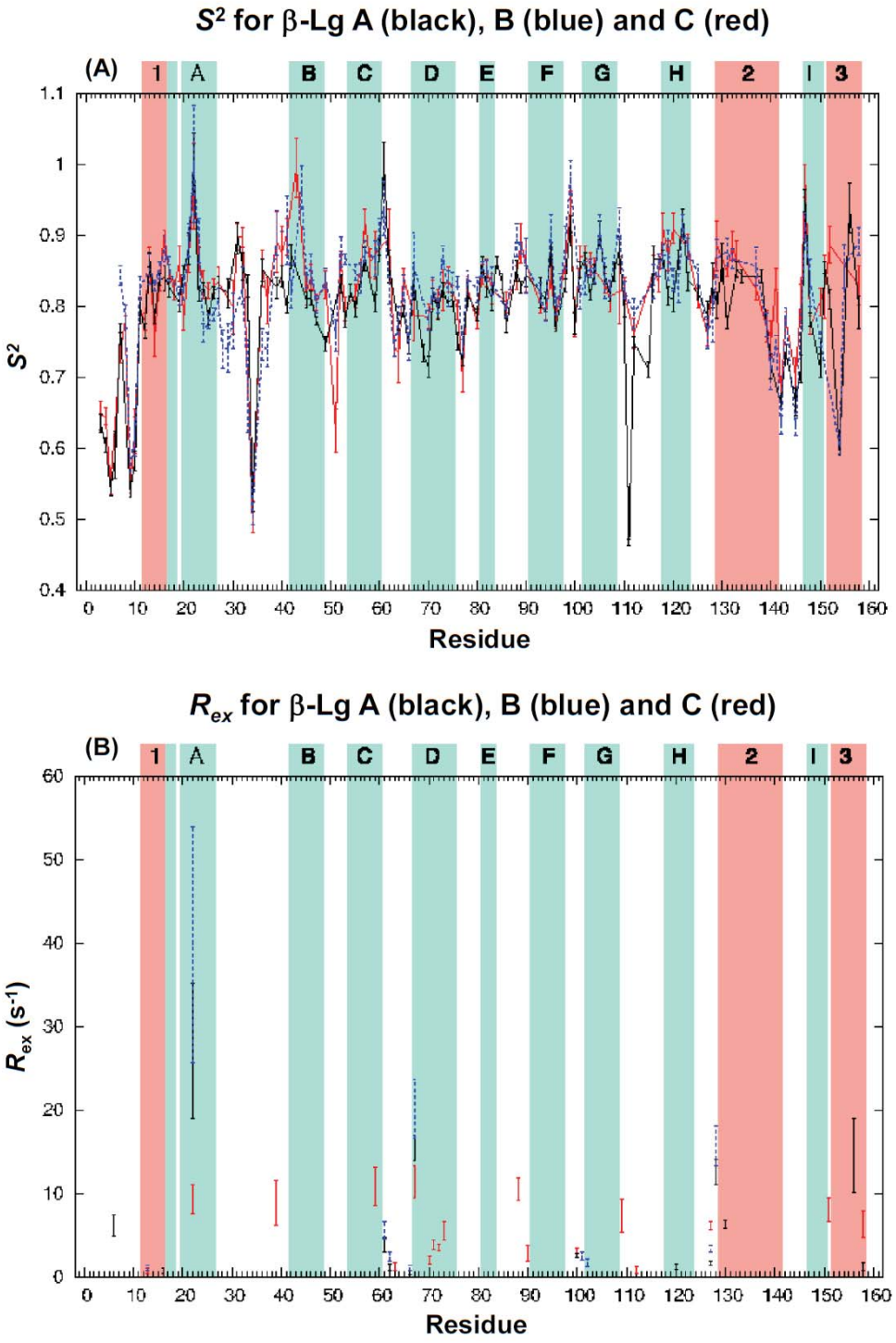


Figure 3.20 Comparison of S^2 and R_{ex} for β-Lg Variants A, B and C.

Comparison of the model-free derived motional parameters for β-Lg A (black), B (blue) and C (red). (A) Plot of S^2 values *vs*. residue and (B) plot of R_{ex} values *vs*. residue. In these studies β-Lg variants were monomeric. In both plots the positions of the nine β-strands (labelled A-I) and α-helices (labelled 1-3) are highlighted with teal and salmon, respectively.

3.9 Comparing the Model-Free Derived Order Parameters with those Estimated Using Two Alternative Methods

Since the adoption of the model-free protocol as the standard method for measuring protein backbone dynamics, two potential methods have been proposed which potentially give S^2 values with significantly less experimental time and with simpler analyses (Section 2.4.8). Here, the results of the two alternative methods for estimating S^2 values for β-Lg A are compared with those derived using ^{15}N relaxation measurements in conjunction with the model-free protocol (Lipari & Szabo, 1982a, Clore *et al*., 1990b). The comparison between the β-Lg A model-free S^2 values at 305 K were compared with the structure-derived S^2 values (Zhang & Brűschweiler, 2002) using the β-Lg A NMR ensemble of structures described in the pdb file 1DV9 (Uhrínová *et al*., 2000). For comparison with the Random Coil Index (RCI)-based dynamics, the chemical shifts for β-Lg A of Uhrínová *et al.* were used (1998). The experimental parameters and predicted S^2 values for β-Lg A, using the structure and chemical shift-based methods are plotted in Figure 3.21, (A) and (B) respectively.

3.9.1 The Zhang and Brűshweiler Structure-Based Method

The order parameters estimated by means of the Zhang-Brűschweiler structure based model show that the backbone ^{15}N-^{1}H vectors positioned in loop regions are more flexible on a ns-ps time-scale than those estimated from residues positioned in the neighbouring secondary structural elements (Figure 3.21 (A)). Specifically, S^2 values are seen to dip at the A/B, C/D, D/E and E/F loops plus at the linkages surrounding the major α-helix and the β-I strand. The N- and C-termini are shown to be highly flexible with the exception of the higher ordering of Met7, mirroring model-free values. Met7 is the only residue in the N-terminus which forms a hydrogen bond, in its case to Val94 positioned in the centre of the β-F strand (Qin *et al*., 1998b, Uhrínová *et al*., 2000). On the other hand, the estimated S^2 values derived from the α-helices and β-sheets, predict that these regions are more rigid than that observed *via* model-free analysis. These trends are similar to those observed using the model-free method, except that the dynamics of the C-terminus was not well-defined using the model-free method, due to the lack of ^{15}N relaxation measurements for this region. Both methods are sensitive enough to detect the more rigid nature of Met7 in the N-terminus.

However, though the Zhang-Brűschweiler structure-based model shows that the loops and terminal regions are more flexible than other regions of β-Lg, estimated S^2 values from the large A/B loop and the G/H loop regions appear to be overestimated, and order parameters estimated from the D/E loop region, and from the link that connects the major α-helix and β-I strand, appear to be underestimated,. The S^2 values are also remarkably high and uniform across the β-sheets and α-helices, indicating that the Zhang-Brűschweiler structure based model is not sensitive to intra-strand variations in flexibility and seems to overestimate rigidity through the secondary structural elements.

3.9.2 The Random Coil Index (RCI) Chemical Shift Based Dynamics

The S^2 trace, predicted from backbone chemical shifts by means of the Wishart RCI web server, looks superficially similar in comparison to the S^2 trace projected by NMR spectroscopy model-free analyses (Figure 3.21 (B)). However, a closer look at the data reveals that the overall order of β-Lg A is estimated to be higher using the chemical shift based model, with smaller amplitudes of motions being predicted for the ^{15}N-^{1}H vectors throughout most of the backbone. Similar to the model-free trace, the N-terminal region is estimated to be highly flexible, with a relatively higher order predicted for Met7. Gradual increases and decreases in flexibility along the β-strands and α-helices are often (but not always) observed when estimating order parameters, showing that this method is slightly sensitive to changes in flexibility within the secondary structural elements.

3.9.3 Comparison of Methods Estimating β-Lg Order Parameters

It was determined that of the methods analysed, the extraction of model-free order parameters from ^{15}N NMR spectroscopy relaxation experiments proved to be more sensitive than both the structure-based model; and the RCI method, deeming it more applicable in detecting subtle changes in motion across the backbone. However, both alternative methods have their merits as both can give a qualitative description of a protein dynamics prior to relaxation analyses, and both give information for regions where ^{15}N relaxation data is sparse. It could be argued that the order parameters derived from the model-free method are less accurate than these shortcut methods, but as the S^2 values for the A, B and C variants were determined independently and yet show very similar trends, it was concluded that the model-free analyses are more reliable.

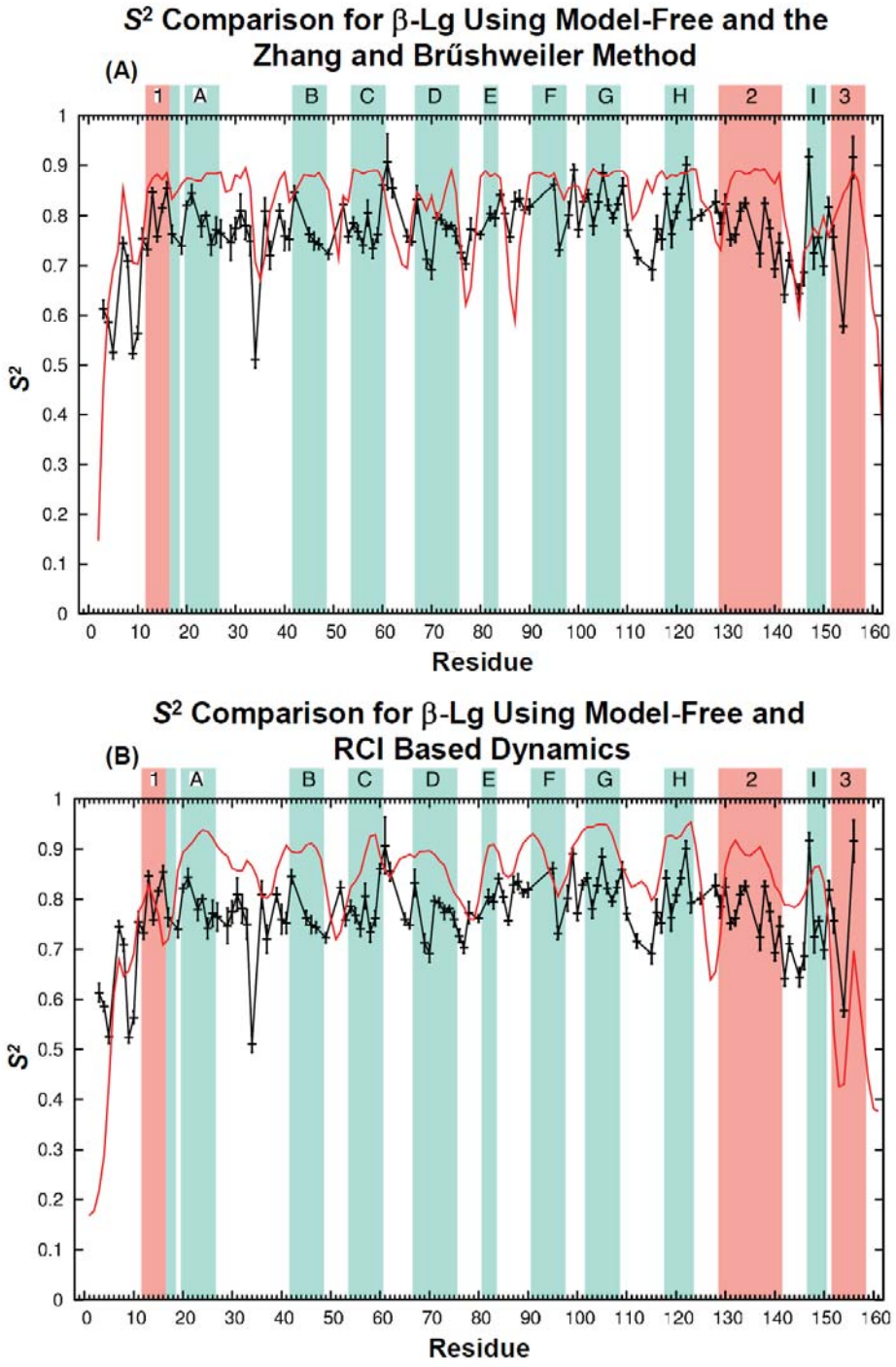


Figure 3.21 Comparison of Methods Estimating β-Lg Order Parameters.

Comparison of the model-free derived order parameters of β-Lg A with (A) order parameters estimated by means of the Zhang-Brűschweiler structure-based model using the pdb file 1DV9 (Uhrínová *et al.*, 1998), and (B) order parameters predicted from β-Lg A backbone chemical shifts (Uhrínová *et al.*, 1998). In these studies, β-Lg A was monomeric. In both plots the positions of the nine β-strands (labelled A-I) and α-helices (labelled 1-3) are highlighted with teal and salmon, respectively.

3.10 β-Lg Covalently Linked Mutant Dimers

3.10.1 Introduction

β-Lg variant A/B ΔAla34Cys mutants were created to suppress signal broadening during NMR spectroscopy analyses at neutral pH. This is required so that β-Lg's motional behaviour can be studied close to the physiological pH of cows' milk, assuming that signal broadening stems from the dynamics of the monomer-dimer equilibrium, which causes conformational exchange. The positions of the mutations were chosen to situate the two cysteine residues, located in the A/B loop at position 34, in proximity to each other, with the purpose that the monomers link through the formation of a disulfide bridge (Figure 3.22). This had been achieved previously by Sakurai and Goto (2006) using the *Pichia pastoris* yeast expression host, but not with a prokaryotic *E. coli* expression system.

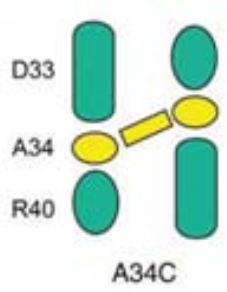


Figure 3.22 Ala34Cys Mutation Designed to Engineer an Artificial Covalently Linked Dimer.

Artificial homodimers were engineered to suppress signal broadening during NMR spectroscopy analyses at neutral pH. Covalent dimers were generated by the formation of disulfide linkages between Cys34 residues, positioned at the dimeric interface. Yellow circles represent introduced cysteine residues, whereas yellow lines show where disulfide bonds are expected to form. Figure rendered from studies published by Sakurai and Goto (2006).

3.10.2 Site-directed mutagenesis

In these studies site-specific base-pair substitutions were critical for the generation of vectors required for the expression of the mutant β-Lg homodimers. These plasmid constructs were created by PCR site-directed mutagenesis separately (Section 2.2.8), with a set of complementary mutagenic primers; 39 oligonucleotides in length (Table 2.6), pETDuet-DsbC-BLG A/B constructs (template DNA) and other necessary PCR components as listed in Section 2.2.8. The oligonucleotides were designed to complement the BLG gene insert within the DNA template, at positions 82 through to

120, with the exception of three targeted base-pair mismatches at positions 100 to 102. The pETDuet-DsbC-BLG A and B constructs (Table 2.5) were both mutated and amplified separately to produce expression vectors, with these primers to produce the mutated β-Lg variants.

PCR amplifications of the constructs were confirmed by agarose gel electrophoresis, separately. Results are not shown but were typical to those described in Section 3.1.2. Site-directed substitutions of three nucleotide bases (GCG→TGT), at positions 100 to 102 within the BLG A and B genes, were confirmed with nucleotide sequencing by means of the Big Dye Terminator V3.1. All other bases were homologous to the respective parental BLG A and B gene sequences. DNA sequence alignment of both mutated genes, showed the desired A34C residue substitution within the β-Lg A and B translated protein sequences. The pETDuet-DsbC-BLG A and B mutated constructs were used separately for the heterologous expression of respective β-Lg variant A Ala34Cys and variant B Ala34Cys, and the molecular chaperone DsbC isomerase in the *E. coli* Origami B (DE3) strain.

3.10.3 Expression

Initially, β-Lg expression and covalent homodimer formation were investigated in whole-cell samples using the *E. coli* Origami B (DE3) [pETDuet-DsbC-BLG] system. IPTG-induced and uninduced samples were analysed by means of reduced and non-reduced SDS-PAGE (15 % acrylamide), to identify whether or not the expression system exploited was capable of producing dimeric protein molecules; and if so, to detect if this system produced mixtures of monomeric and dimeric β-Lg, or completely dimeric protein.

In this study β-mercaptoethanol (BME), a component of the sample loading buffer, was used as a reducing agent. BME works by denaturing proteins through the reduction of disulfide bonds, which can destroy the native conformation required for the protein's biological function. A reducing agent, such as BME, is typically added to the sample loading buffer along with other components, so that the proteins separate more linearly on a SDS-PAGE gel and, hence, more accurate molecular weights (MW) can be determined. In these investigations BME (reduced SDS-PAGE) was used to reduce the disulfide bonds within the tertiary (intra-molecular) and quaternary (inter-molecular) structure of the covalently linked mutant, enabling the protein to migrate uniformly on an SDS-PAGE gel with an electrophoretic mobility corresponding to ~18.4 kDa, the

size of the β-Lg monomer. Also, omitting BME from the sample loading buffer (non-reduced SDS-PAGE) is beneficial in these studies, as it showed whether or not β-Lg had formed disulfide-bonded dimers, higher-order covalently linked oligomers plus monomers that had not formed dimers.

Native dimers, those that form because of the wild-type monomer-dimer equilibrium (Section 1.4.1), are not present in the reduced and non reduced sample mix because of the presence of SDS, which disrupts any non-covalent interactions at the dimeric interface and within the protein molecule. Therefore all non-covalently dimerised β-Lg separates into monomers. SDS detergent works by denaturing non-covalent interactions within the protein, destroying its native conformation, thereby eliminating differences in shape as a factor for separation on a PAGE gel.

Studies with whole-cells established that the *E. coli* expression system used in these studies is capable of producing recombinant β-Lg ΔAla34Cys protein (Figure 3.23). Comparisons between uninduced cells (Lanes 1 and 3) and IPTG-induced cells (Lanes 2 and 4) showed the presence of two additional bands in both IPTG-induced whole-cell samples, verifying that IPTG-induced expression was successful. An extra band in Lane 3, corresponding to the size of monomeric β-Lg (~18.4 kDa), is evidence for the expression of β-Lg in the host cells, whereas the extra band in Lane 4, corresponding to the size of a protein approximately two times the MW of β-Lg (~36.8 kDa), confirmed that the mutant had successfully dimerised through a covalent linkage. The second band, present in both Lanes 2 and 3, had an electrophoretic mobility corresponding with a protein with a similar to the MW of DsbC (23 kDa), showing that the chaperone protein had been successfully co-expressed.

These results verified that the mutant protein had completely dimerised through the formation of a covalent disulfide bridge. Mixtures of monomer and dimer were not apparent on the non-reduced SDS-Page gel. These investigations showed that β-Lg had not formed any higher order oligomers as there were not any visible additional bands Commassie-Blue stained bands with molecular weights corresponding to 55.2 kDa (tetrameric β-Lg), 73.6 kDa (quatermeric β-Lg) or higher (up to 116 kDa).

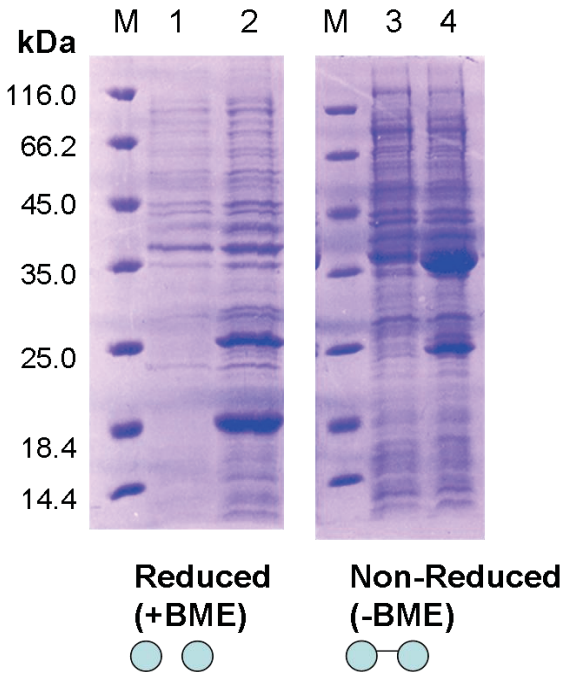


Figure 3.23 Reduced and Non-Reduced SDS-PAGE Analysis of IPTG-Induced Expression of β-Lg A Ala34Cys.

Proteins were separated by SDS-PAGE and stained as described in Materials and Methods (Section 2.3.2). BME was either present (+) or absent (-) in the sample loading buffer. Lanes M: molecular mass markers as in Appendix A.4; Lane 1: whole-cell uninduced +BME; Lane 2: whole-cell IPTG-induced +BME; Lane 3: whole-cell uninduced - BME; Lane 4: whole-cell IPTG-induced - BME. Linked circles represent conditions in which covalent dimers could be detected, whereas unlinked represent conditions were β-Lg would be totally monomeric.

Soluble and insoluble samples, sourced from uninduced and IPTG-induced expression hosts, were collected for β-Lg Ala34Cys solubility analysis and separated using reduced SDS-PAGE (results not shown). Both recombinant mutants were predominantly found in the soluble cytoplasmic fractions of the expression host (over ~80 %) and displayed a similar expression profile to that of monomeric recombinant β-Lg C (Section 3.2.2; Figure 3.2).

3.10.4 Purifying the β-Lg A34C Mutant

Purification of both the β-Lg A and B Ala34Cys mutant dimers follows similar trends to the purification of β-Lg C (Section 3.2.3). The only distinguishing feature is that during anion exchange chromatography, the β-Lg mutant dimers eluted in the 10 % NaCl fractions and not the 5 % fractions where β-Lg C was consistently found (results not shown). Yields and purification progression were similar to those for observed for β-Lg C (Section 3.2.3). Following purification, the dimers were not reduced to form monomers as assessed with SDS-PAGE analyses shown in Figure 3.24. Lane 1

illustrated that a protein similar to the size of the β-Lg monomer was purified. Lane 2 illustrated the purified protein was dimeric, by means of omitting the reducing agent, BME from the sample loading buffer. NMR spectroscopy conformational analyses were then used to assess if the protein purified was β-Lg and if the mutant had been folded correctly.

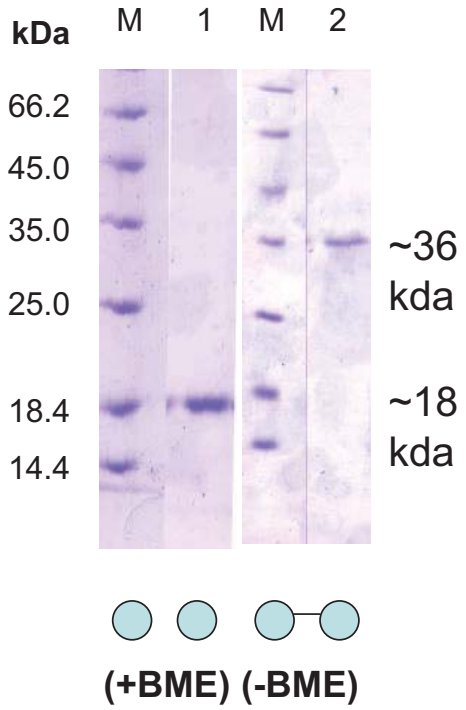


Figure 3.24 Reduced and Non-Reduced SDS-PAGE Analysis of Pooled Samples Containing Purified β-Lg A34C Mutants.

Fractions were pooled following *size*-exclusion chromatography. Proteins were separated by SDS–PAGE and stained as described in Materials and Methods (Section 2.3.2). BME was either present (+) or absent (-) in the sample loading buffer, Lanes M: molecular mass markers as in Appendix A.4. Lane 1: SEC purified fractions +BME; Lane 2: SEC purified fractions -BME.

3.10.5 NMR Spectroscopy at Neutral pH

To assess if the dimeric mutant had been folded correctly a ^{15}N,^{1}H-HSQC spectrum was collected at 305 K on a ^{15}N-enriched sample of the artificially-linked dimeric β-Lg variant A Ala34Cys protein. The overlay of the ^{15}N-labelled mutant with monomeric β-Lg A showed that the mutant displayed good-line width dispersion of the ^{1}H and ^{15}N resonances, indicating that the mutant was well-folded. The overlay of the dimeric mutant had similar peak dispersion to that of the monomeric ^{15}N-labelled β-Lg A sample, implying that the mutant had been folded in a 'native-like' state.

This study implies that the Ala34Cys mutant had been folded correctly; however, 3D NMR spectroscopy analyses need to be conducted with double-labelled protein ($^{13}C/^{15}N$) to verify if the mutant structure is able to be assigned at neutral pH. If assignments are achievable, then a better assessment of the chemical shift perturbations between the different data sets, could point to differences in conformation at the dimeric interface, which is due to β-Lg A being monomeric at pH 2.6 and the artificially linked mutant being completely dimeric at neutral pH, due to the suppression of the monomer-dimer equilibrium. The ^{15}N variant B Ala34Cys mutant is yet to be tested. If such studies are feasible, then the next step would be to assess its backbone dynamics at pH 6.5.

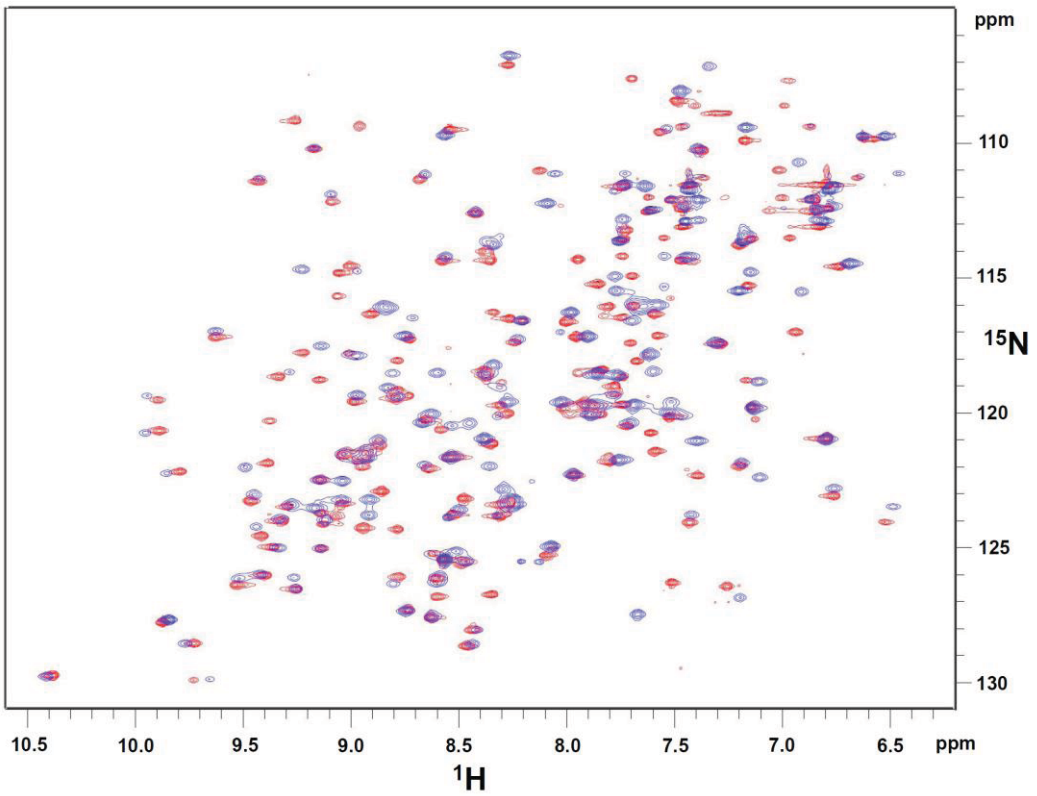


Figure 3.25 Overlay of the ^{15}N-Labelled β-Lg A Ala34Cys HSQC Spectrum with the 'Native-Like' Monomeric ^{15}N-Labelled β-Lg A HSQC Spectrum.

The sample conditions were 50 mM KH_2PO_4 buffer, 5 % D_2O at 305 K. The mutant spectrum (blue) was acquired at pH 6.5 and the spectrum for the β-Lg A (red) variant was collected at pH 2.6, where β-Lg is monomeric.

3.11 ^{15}N Backbone Dynamics of β-Lg

To provide a greater understanding of how bovine β-lactoglobulin behaves and how its behaviour is affected by heat, ^{15}N backbone dynamics of monomeric β-Lg C was assessed and compared at 305 K, 313 K and 320 K. Also, to provide a greater understanding of factors that differentiate the structurally similar A, B and C variants from one another, ^{15}N backbone dynamics of the three variants were compared at 305 K. All measurements were made at pH 2.6, where β-Lg is monomeric. The following relates the major findings of the ^{15}N backbone dynamical investigations to β-Lg's solution structure, in concert. This discussion is made with reference to the results for β-Lg C at 305 K. Notable differences and similarities in ^{15}N backbone dynamics among the different temperature sets and among the variants are also discussed.

Flexibility of the Terminal Regions

Significantly low order parameters and longer effective correlation times (> 500 ps) are detected with relaxation measurements for six of the seven N-H bonds of residues (Val3, Thr4, Gln5, Met7, Gly9 and Lys10), positioned in the N-terminal region (residues Leu1-Lys10). Additionally, internal motions on more than one time-scale contribute significantly to relaxation, with all six of these backbone amides being fitted with the two time-scale model, model five. Unsurprisingly, these results show that the N-terminus is highly mobile and undergoes complicated dynamics due to its lack of rigid structure.

Met7 is associated with a low S^2 value, but it is still more rigid than the five other residues in this region with significantly low order parameters. Its relative rigidity has been commented on previously from studies that suggest that this is due to the presence of weak hydrogen bond formed between Met7 and Val94 (β-E) (Qin *et al*., 1998b, Uhrínová *et al*., 2000). In these investigations, this phenomenon has also been detected with a method estimating S^2 from the empirical relationship between NMR spectroscopy chemical shifts and flexibility (Uhrínová *et al*., 1998, Berjanskii & Wishart, 2005) and a method predicting S^2 from its solved solution structure (Uhrínová *et al*., 2000, Zhang & Brűschweiler, 2002). The more rigid nature of Lys8 in this region (S^2 = 0.78), as prescribed by the model-free protocol, might be a consequence of its association to Met7.

The solution structure (pdb entry 1DV9) suggests that residues Ile12 through to Ala16 form a 3_{10}-helix (Uhrínová *et al*., 2000). The data in these investigations indicate that the

amide bonds of residues in this region and Asp11 have near average S^2 values and are generally more rigid than residues Leu1 to Lys10.

Linkages and Loops

As expected, the S^2 values show that the amplitude of motions of the backbone ^{1}H-^{15}N vectors positioned in the linkages and loop regions are greater than those of their neighbouring β-strands and α-helices. In addition, relaxation data for the N-H bonds of residues positioned at these sites were often fitted with at least two parameters to describe motions. In some cases, dynamics were described with a motional parameter, which takes into account the correlation times of internal motions that are not very fast ($\tau_e > 500$ ps). This trend is noted for the N-H bonds of Ala34 (A/B loop), Glu51 (B/C loop) and Lys77 (D/E loop) at 305 K. The internal motions of these residues are also described by model five, the two time-scale fitting model, which points to the presence of an additional faster internal motion. These complicated dynamics reflect a level of disorder in these regions, which is not uncommon for backbone amides positioned in loops and linkages (Palmer, 1993, Fushman, 2003).

In other cases, the motions of the N-H bonds of residues located at these sites are also described with R_{ex}, which takes into account the presence of additional slower motions (ms – μs), indicative of conformational exchange. Interestingly, residues Leu39 (A/B loop), Asn88 (E/F loop) and Asn109 (G/H loop) are each associated with a relatively large R_{ex} value that was indicated by a significantly higher than average ^{15}N R_2 relaxation rate. Notably, these loops are located at the more open end of the barrel, and not the more closed end, which provides support for Uhrínová and coworkers' (2000) speculation that these slower motions, at the more open end of β-Lg, may help with the regulation and ingress of ligands into the hydrophobic cavity. Loop F/G contains a residue with a relatively small contribution of R_{ex}, but this is located at the more closed end of the barrel. The absence of longer effective correlation times, for at least some of these residues listed, maybe due to underestimations of τ_e, which may result from the model-free selection technique (d'Auvergne & Gooley, 2003).

At neutral pH, the A/B loop forms part of the dimeric interface. However, these studies were conducted at pH 2.6 where β-Lg is monomeric. The S^2 values through this region are relatively high compared to those seen across the rest of the backbone, particularly for the N-H bonds of residues Leu31, Leu32, Leu39 and Val41, whose S^2 values range from 0.89 to 0.90. The short 3_{10}-helix occupying residues Ile29 through to Leu32 (Uhrínová *et*

al., 2000) may account for the higher S^2 values seen in this region. The notable increase in S^2 for Ser30, positioned in the 3_{10} helix at 320 K, shows that this region becomes more associated with another element in β-Lg as the temperatures are raised, contributing to its increased rigidity. However, the S^2 value for Ala34 is the lowest observed in the protein, implying its backbone amide is significantly flexible, due to the loop's long length at low pH.

At low pH, the E/F loop, which is thought to control ligand entry, is in its 'closed' position with the protonated side-chain of Glu89 tightly buried within hydrogen bonding distance of Ser116 (Uhrínová *et al*., 2000, Sakurai & Goto, 2006). The relative rigidity of libration across this region, as compared to other loop regions correlates reasonably well with this loop's position. The relatively high S^2 value for Glu89 (S^2 = 0.89) also supports this. However, at low pH the equilibrium between 'closed' and 'open' position still exists, which could provide an explanation for the presence of conformational exchange observed for residues Asn88 and Asn90.

β-Lg's Barrel and Secondary Structural Elements

The higher order seen throughout the eight β-strands, which make up the β-barrel, is a reflection of the dense network of inter-strand H-bonding and other stabilising factors such as the presence of an internal disulfide bridge. There regions are generally more rigid on a ns to ps time-scale than the disordered loops, which unsurprisingly, is similar to NMR spectroscopy studies describing the fast internal motions of another lipocalin, mouse major urinary protein 1 (MUP-1) (Křížová *et al.*, 2004). S^2 trends are observed along the lengths of many of the β-strands pointing towards gradual increases in flexibility along three strands (β-A, β-B and β-G), and, in other strands, S^2 values reflect a gradual increase in rigidity along the strands (β-C, β-D and β-H). S^2 trends, which may stem from the strands' orientation within the barrel and which show a gradual increase in flexibility as the strands reach towards the open end of the barrel, which may help with the ingress of ligands into the hydrophobic core. This trend is observed for strands β-A, β-D, β-G and β-H. A gradual increase in flexibility is also observed across the length of the major α-helix (α-2). However, it is possible that the α-helix becomes more associated with other nearby regions as the temperature is increased, as seen by an increase of rigidity, on the ns to ps time-scale, for Leu133.

Leu22 is critically positioned at the hinge point of β-A, between the two β-sheets. Its position is conserved in all three variants tested and at all temperatures tested. The backbone amide is very restrained (305 K S^2 = 0.97), but its high R_{ex} value coupled to a significantly high R_2 rate signifies its participation at the hinge point of both sheets, which is possibly due to the slight movements causing the expansion and contraction of the barrel that involves a change in Leu22's conformation. This has also been noted for Ser21, in a study published by Uhrínová *et al*. (2000), but not Leu22, as its peak could not be not assigned. However, these results show that Leu22 holds this critical position in the recombinant β-Lg molecules that have been expressed for these studies, and not Ser21, as motions for this residue could be accounted for without a R_{ex} contribution.

The β-C strand contains the Gln59His site of substitution, which is the site that differentiates the C variant from both the A and B variants. The R_{ex} values measured at 305 K and 313 K (model-free data not acquired at 320 K), shows that this His59 in the C variant is likely to undergo conformational exchange at the temperatures sampled. However, because of the variability in the R_2 data observed between the variants, which has been discussed in Sections 3.8.2 and 3.8.4, the absence of the R_{ex} term for the A and B variants is not conclusive, but strongly indicates that these finding needs to be further investigated. If the repeated studies at this site reveal that its backbone dynamics are significantly different, it would contribute to a greater understanding of factors that differentiate these structurally homologous variants, which is of great importance to the dairy industry. No significant differences in backbone dynamics have been detected at the two A→B substitution sites (Asp64Gly and Val118Ala).

This long β-D strand (Ala67 – Lys75) is located next to the comparatively short β-E strand (Val81 - Lys83). Because of the strand's relative length, its H-bonding potentials of β-D are not completely fulfilled. An increase in flexibility (ns – ps) is detected along the β-D strand, which is shown by a gradual increase in S^2 values, starting from the more closed end of the barrel to the more open end. At 305 K, two H-bonds are formed between Ala63 (β-D) and Lys83 (β-E), which helps to tether the β-E strand to its surrounding secondary structural elements, as observed by the simple model fitting (models one and two) of β-E's three residues. In contrast, the relatively long length of the β-D strand allows for additional slower motions (ms - μs) to occur for backbone amides beyond the H-bond, up until the more open end of the barrel. Due to an increase in temperature, localised changes in model selection are detected in the β-E strand

showing that the β-E strand becomes more conformationally mobile between 313 K and 320 K. This is possibly caused by the disruption of the two H-bonds formed between β-E and β-D that tethers the short β-E strand in place at 305 K and 313 K, causing conformational exchange,

At 305 K, ^{15}N relaxation data from two stretches of residues that encompass the β-G and β-H strands were only fitted with the simple one-parameter model, showing that the correlation times (τ_e) of the internal motions were too short to contribute significantly to relaxation. An extensive network of inter-strand H-bonds formed between neighbouring secondary elements plus a stabilising Cys106-Cys119 *cis*-disulfide bridge that links these two strands provides a high level of order to this region, which contributes to rigidity of movements, as detected by significantly high order parameters near this disulfide bond. However, the highly ordered β-G and β-H strands become more conformationally mobile as the temperature is raised to 320 K, in regions that are located at the more open end of the barrel. This may be due to a stretching motion of these strands caused by increased temperatures, or to the possible disruption of inter-strand H-bonds in these areas.

The Free Thiol and the External Cys66-160 Disulfide Bridge

The model-free motional parameters could not be obtained for the free cysteine residue (Cys121) and the two cysteine residues that form a covalent bridge (Cys66-Cys160) for the C variant; however, preliminary dynamical investigations of the A variant indicates that the backbone amide of Cys66 (C/D loop, positioned at edge of strand β-D) is relatively flexible (S^2 = 0.74), which is not uncommon for residues located in loop regions, and the free cysteine (strand β-H) is more rigid (S^2 = 0.87), which is unsurprising as it participates in the formation of three inter-strand H-bonds. However, data from these studies do not show if the Cys66-Cys160 disulfide bridge contributes to rigidity on the ns to ps time-scale, as ^{15}N relaxation data acquired from the C-termini were sparse.

.

4

Conclusions and Future Work

4.1 Conclusions

4.1.1 Generating Isotopically Labelled β-Lg

The creation of vectors for the expression of β-Lg C variant and the covalently linked β-Lg A and β-Lg B Ala34Cys dimeric mutants were required for the analysis of ^{15}N backbone dynamics of β-Lg at low pH (monomeric β-Lg C), and the intention of generating β-Lg protein (β-Lg mutants), which is more amenable to NMR spectroscopy analyses at pH 6.5, through the suppression of signal broadening during NMR spectroscopy analysis, which is caused in part by the dynamics of the monomer-dimer equilibrium at this pH. Following the successful overproduction of pure β-Lg A and B (Ponniah *et al*., 2010), rationalised site-directed mutagenesis of the BLG gene, within the previously studied constructs, was achieved by means of the Stratagene mutagenesis protocol.

In this study, isotopically labelled 'native-like' β-Lg C and β-Lg A and B dimeric mutants were successfully overexpressed and purified using an expression and purification system developed by Ponniah *et al*. (2010) and Mailliart & Ribadeau-Dumas (1988). However, the published purification protocol had to be adjusted to attain the published yields by means of increasing protein concentration to between 3 and 5 mg mL^{-1}, just before acid and salt precipitation. Following purification, SDS-PAGE analysis showed that all β-Lg samples purified were sufficiently pure for NMR spectroscopy analyses and that the Ala34Cys mutants had completely dimerised through the formation of an introduced disulfide link. Following these studies, NMR spectroscopy analyses revealed that the C variant and mutants had been folded in a 'native-like' conformation and the narrow peaks displayed by β-Lg C in the $^{15}N,^{1}H$-HSQC spectrum, at low pH, showed that this protein is amenable to NMR spectroscopy dynamical analyses. NMR spectroscopy conformational analyses at neutral pH showed that the mutant had been folded correctly. However, 3D studies are needed to see if the protein is assignable, under the conditions sampled. In these investigations, the conditions of the mutants had not been studied extensively to assess whether these proteins are amenable to NMR spectroscopy studies. To our knowledge this is the first time that recombinant techniques have been used to express the β-Lg C variant, whereas the Ala34Cys mutant has been successfully expressed before, but in the *Pichia pastoris* host (Sakurai & Goto, 2006).

4.1.2 ^{15}N Backbone Dynamics of β-Lg C at 305 K

In these studies, model-free motional parameters showed that the dynamics and flexibility of β-Lg C, at 305 K, correlate well with the solution structure solved by Uhrínová *et al.* (2000). The N-terminus is highly mobile and experiences more complicated dynamics on more than one time-scale, due of its lack of rigid structure. Not surprisingly, the internal motions for many of the backbone amides positioned in loops and link regions are more flexible than the more rigid neighbouring secondary structural elements. The A/B loop, which forms part of the dimeric interface at neutral pH, is highly flexible, but the region of the loop occupying the 3_{10}-helix is relatively rigid. Additional motions pointing towards conformational exchange are also observed in loop regions, with these studies showing that significantly high R_2 values, which correlate with high R_{ex}, values, define dynamics for a number of backbone amides of residues positioned in loops at the more open end of the barrel (Loops A/B, E/F and G/H), whereas such dynamics are not detected at the more closed end. Uhrínová and co-workers (2000) have previously speculated that these slower motions at the open end of β-Lg may help facilitate the ingress of ligands into internal cavity. The presence of conformational exchange is also defined for many residues in the β-D strand, the β-strand in β-Lg whose H-bonding potentials are not completely fulfilled. In contrast, the internal dynamics of the highly ordered regions of the β-G and β-H are only described by S^2 at 305 K. These studies have also shown that some gradual changes in flexibility are observed along the lengths of many of the secondary structural elements, which includes a large change in flexibility along the major α-helix. For four β-strands, flexibility is shown to increase towards the top of the β-barrel, which may help with the ingress of ligands, into the hydrophobic cavity. Other striking features observed in the β-strands include the presence of conformational exchange for Leu22, which is situated at a critical position at the hinge point of the β-A strand, occupying both β-sheets and His59, a residue positioned at the B→C substitution site.

4.1.3 Effects of Temperature on ^{15}N Dynamics of β-Lg C

These investigations have shown that increases in temperature affect the dynamics of β-Lg C. Temperatures in these studies only varied between 305 K and 320 K. However, β-Lg C is still seen to become more conformationally mobile as the temperature is raised to 313 K and then 320 K. The addition of slower motions in the β-E strand at 320 K potentially arises from the disruption of an H-bond, which tethers the short β-E strand to

the longer β-D strand, between 305 K and 313 K, and the stretching motion of the highly ordered β-G and β-H strands may also contribute to conformational exchange as temperatures increase.

Even though the amplitudes of internal motions across the core structure of β-Lg are highly similar at all temperatures sampled, two subtle changes have been observed for internal motions on a ns-ps time-scale. These two changes suggest that at least parts of the major α-helix, and residue Ser30 in the A/B loop become more associated with other parts of β-Lg as temperatures are raised, shown by an increase in S^2 when the temperature is raised to 320 K.

4.1.4 Effects of Polymorphisms on ^{15}N Backbone Dynamics of β-Lg

The differences in ^{15}N backbone dynamics, among the A, B and C variants at 305 K, which are due to site-specific internal and external amino acid differences, are not conclusive due to the variability in the R_2 relaxation and the frequencies of the R_{ex} term observed at some points in β-Lg C, when compared to the corresponding data for the A and B variants. However, these preliminary investigations show that that these experiments need to be repeated again, to document changes in dynamics, especially at the Gln59His (B→C) substitution sites, where R_{ex} has been noted for C, but not for variants A and B. Other differences for the frequency of the R_{ex} term in β-Lg do not occur at the substitution sites, or in close proximity to these sites, but are at sites where conformational exchange is not unreasonable, such as the loop regions and the less ordered β-D strand. The commonalities among the variants, in the measured R_1 relaxation data and the model-free iteratively refined S_2 data, suggest that the differences in backbone dynamics, across the ns to ps time-scale, are minor.

4.1.5 Methods of Interpreting Backbone Dynamics

These investigations have shown that the model-free method is also more reliable than two other methods trialled for estimating flexibility of β-Lg A's backbone, as defined by S^2. The model-free protocol is able to detect more subtle variations in flexibility along β-Lg's secondary structural elements than the Zhang and Brűshweiler method (2002), which uses the resolved protein structure to estimate S^2, and is also able to provide a more accurate estimate of the overall flexibility of internal motions, when compared to the Wishart RCI web server, which predicts S^2 based on an empirical relationship between chemical shifts and flexibility (Berjanskii & Wishart, 2005).

4.2 Future Directions

4.2.1 Alternative Testing for Model Selection

Studies published by Chen *et al.* (2004), with simulated data, showed that in some cases the statistical tests used for model selection in the model-free protocol (Mandel *et al*., 1995), fail to select the appropriate model for interpreting backbone dynamics. These studies imply that by using an alternative test, the Bayesian information criteria (BIC) (Schwartz, 1978), more appropriate models are selected to describe motions. In the future it would be informative to process the ^{15}N relaxation data again, this time taking BIC testing into account at all three temperatures, to see if the effective correlation time (τ_e) of the internal motions had been underestimated, particularly for backbone amides in the α-helix and the A/B loop, which display relatively high order parameters with increasing temperature (Section 3.6.2).

4.2.2 Comparing Dynamics of β-Lg A, B and C Variants at Higher Temperatures

Comparing dynamics of the three variants at higher temperature (see Section 4.4.3), may shed light on results previously published, which suggested that at 320 K there is a difference in configuration of the Cys66-160 bond between the C variant and the major variants, A and B (Manderson *et al*., 1997). Even though these experiments were recorded at neutral pH and at lower protein concentrations, this finding should be investigated further and other potential dynamical differences resulting from amino acid substitutions, which may occur when the temperature is raised, need to be documented. Preceding NMR spectroscopy analyses, however, it would be informative to repeat the near-UV-CD experiments, but at pH 2.6, to note if this spectral difference among the variants also occurs at low pH.

4.2.3 Assessing Dynamics at More than One Static Magnetic Field Strength

Measuring ^{15}N relaxation rates at least at two different field strengths would provide a more thorough assessment of the slower motions (ms-μs), identified by the conformational exchange constant (R_{ex}), by offering a more precise assessment of rates at which conformational exchange occurs (Millet *et al*., 2000). Second; the 500 MHz spectrometer is capable of collecting spectra at higher temperatures (up to 373 K) than the 700 MHz spectrometer. For this reason, more interesting changes in dynamics could be spotted at higher temperatures, providing that β-Lg does not aggregate at the held temperature, over the length of time relaxation measurements are recorded. In addition,

further changes in dynamics could be detected at temperatures higher than 320 K that are a due to the internal and external amino acid polymorphisms within the A, B and C variants, providing a greater understanding of factors that differentiate these proteins during heat treatment. The 700 MHz spectrometer was used in these investigations as it offers higher resolution.

4.2.4 β-Lg's Putative Role as a Pheromone-Binding Protein

Investigations into β-Lg's putative biological role as a pheromone-binding protein have not been documented in any journals. Much of the research to date, in regards to its possible biological function, have mainly focussed on β-Lg hypothetically acting as a transporter protein, delivering nutrients from the mother's milk to the intestine of the suckling neonate (Section 1.11). However, β-Lg bears striking similarities to another lipocalin, the mouse urinary protein-1 (MUP-1) (Figure 1.9), which is secreted in the urine of male mice (Böcskei *et al*., 1992, Timm *et al.*, 2001). MUPs bind volatile pheromones, which elicits a female mouse's sexual response upon excretion. In many vertebrates, specific pheromones trigger innate behavioural responses such as attraction and recognition: for example, the suckling response in rabbits is controlled by the pheromone 2-methylbut-2-enal (Luo, 2004). Future investigations into β-Lg's potential for binding specific pheromones, to trigger a similar suckling/recognition response in bovine, would provide another avenue for NMR spectroscopy analyses that could potentially shed light on its true biological function. To compare dynamics of β-Lg, with and without potential pheromone candidates binding, experiments would need to be conducted at pH 6.5, where β-Lg is able to bind ligands. Therefore, the Ala34Cys mutant created in these studies would need to be investigated further to assess whether or not NMR spectroscopy dynamical experiments can be optimised to reduce peak broadening at neutral pH. If this is possible, these putative experiments could potentially provide a new perspective on how this protein behaves *in vivo*.

References

Ali, S. & J. Clark, (1988) Characterization of the gene encoding ovine β-lactoglobulin: Similarity to the genes for retinol binding protein and other secretory proteins. *Journal of Molecular Biology* **199**: 415-426.

Ariyaratne, K. A. N. S., R. Brown, A. Dasgupta, J. de Jonge, G. B. Jameson, T. S. Loo, C. Weinberg & G. E. Norris, (2002) Expression of bovine *β*-lactoglobulin as a fusion protein in *Escherichia coli*: a tool for investigating how structure affects function. *International Dairy Journal* **12**: 311-318.

Aschaffenberg, R. & J. Drewry, (1955) Occurrence of Different β-Lactoglobulins in Cow's Milk. *Nature* **176**: 218-219.

Batt, C. A., L. D. Rabson, D. W. Wong & J. E. Kinsella, (1990) Expression of recombinant bovine β-lactoglobulin in *Escherichia coli*. *Agricultural and Biological Chemistry* **54**: 169-178.

Belloque, J. & G. M. Smith, (1998) Thermal Denaturation of *β*-Lactoglobulin. A ^{1}H NMR Study. *Journal of Agricultural and Food Chemistry* **46**: 1805-1813.

Berjanskii, M. V. & D. S. Wishart, (2005) A Simple Method To Predict Protein Flexibility Using Secondary Chemical Shifts. *Journal of the American Chemical Society* **127**: 14970-14971.

Bessette, P. H., F. Aslund, G. Beckwith & G. Georgiou, (1999) Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proceedings of the National Academy of Sciences of the United States of America* **96**: 13703-13708.

Bewley, M. C., B. Y. Qin, G. B. Jameson, L. Sawyer & E. N. Baker, (1997) III.1 - Bovine β-lactoglobulin and its variants: a three-dimensional structural perspective. In: IDF Seminar. Palmerston North, New Zealand: International Dairy Federation, pp. 100-109.

Böcskei, Z., C. R. Groom, D. R. Flower, C. E. Wright, S. E. Phillips, A. Cavaggioni, J. B. Findlay & A. C. North, (1992) Pheromone binding to two rodent urinary proteins revealed by X-ray crystallography. *Nature* **360**: 186-188.

Bodenhausen, G. & D. J. Ruben, (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chemical Physics Letters* **69**: 185-189.

Bradford, M. M., (1976) Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding. *Analytical Biochemistry* **72**: 248-254.

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

Chaudhuri, B. N., G. J. Kleywegt, J. B. Björkman, L. D. Lehman-McKeeman, J. D. Oliver & T. A. Jones, (1999) The structures of α_{2u}-globulin and its complex with a hyaline droplet inducer. *Acta Crystallographica Section D: Biological Crystallography* **55**: 753-762.

Chen, J., C. L. Brooks & P. E. Wright, (2004) Model-free analysis of protein dynamics: assessment of accuracy and model-selection protocols based on molecular dynamics simulation. *Journal of Biomolecular NMR* **29**: 243-257.

REFERENCES

Clore, G. M., P. C. Driscoll, P. T. Wingfield & A. M. Gronenborn, (1990a) Analysis of the backbone dynamics of interleukin-10 using two-dimensional inverse detected heteronuclear ^{15}N-^{1}H NMR spectroscopy. . *Biochemistry* **29**: 7387-7401.

Clore, G. M., A. Szabo, A. Bax, L. E. Kay, P. C. Driscoll & A. M. Gronenborn, (1990b) Deviations from the Simple 2-Parameter Model-Free Approach to the Interpretation of Nitrogen-15 Nuclear Magnetic-Relaxation of Proteins. *Journal of the American Chemical Society* **112**: 4989-4991.

Cole, R. & P. Loria, (2003) FAST-Modelfree: A program for rapid automated analysis of solution NMR spin-relaxation data. *Journal of Biomolecular NMR* **26**: 203-212.

Cowan, S. W., M. E. Newcomer & T. A. Jones, (1990) Crystallographic refinement of human serum retinol binding protein at 2Å resolution. *Proteins* **8**: 44-61.

d'Auvergne, E. & P. R. Gooley, (2003) The use of model-selection in the model-free analysis of protein dynamics. *Journal of Biomolecular NMR* **25**: 25-39.

Delano, W. L., (2008) The PyMOL Molecular Graphics System. In. Palo Alto: *DeLano Scientific LLC*.

Denton, H., M. Smith, H. Husi, D. Uhrin, P. N. Barlow, C. A. Batt & L. Sawyer, (1998) Isotopically Labeled Bovine *β*-lactoglobulin for NMR studies Expressed in *Pichia pastoris*. *Protein Expression and Purification* **14**: 97-103.

Dillon, P. J. & C. R. Rosen, (1990) A rapid method for the construction of genes using the polymerase chain reaction. *Biotechniques* **9**: 298-300.

Edwards, P. J. B., L. K. Creamer & G. B. Jameson, (2008) Structure and stability of whey proteins. In: *Milk Proteins: From Expression to Food.* Boston: Academic Press, pp. 163-203.

Edwards, P. J. B., G. B. Jameson, G. E. Norris, T. S. Loo, D. Uhrin, P. N. Barlow & L. K. Creamer, (2004) Exploring the structure and dynamics of labeled β-lactoglobulin using high-field NMR spectroscopy. In: American Dairy Science Association (ADSA), American Society of Animal Science (ASAS), and Poultry Science Association, Joint Annual Meeting, St Louis, Missouri, USA.

Edwards, P. J. B., G. B. Jameson, K. P. Palmano & L. K. Creamer, (2002) Heat-resistant structural features of bovine *β*-lactoglobulin A revealed by NMR H/D exchange observations. *International Dairy Journal* **12**: 331-344.

Farrell, H. M., R. Jimenez-Flores, G. T. Bleck, E. M. Brown, J. E. Butler, L. K. Creamer, C. L. Hicks, C. M. Hollar, K. F. Ng-Kwai-Hang & H. E. Swaisgood, (2004) Nomenclature of the Proteins of Cows' Milk–Sixth Revision. *Journal of Dairy Science* **87**: 1641-1674.

Flower, D. R., (1996) The lipocalin protein family: structure and function. *Biochemical Journal* **318**: 1-14.

Flower, D. R., A. C. T. North & C. E. Sansom, (2000) The lipocalin protein family: structural and sequence overview. *Biochimica Et Biophysica Acta* **1482**: 9-24.

Fogolari, F., L. Ragona, L. Zetta, S. Romagnoli, K. G. De Kruif & H. Molinari, (1998) Monomeric bovine β-lactoglobulin adopts a β-barrel fold at pH 2. *FEBS Lett* **436**: 149-154.

Fugate, R. D. & P. S. Song, (1980) Spectroscopic characterization of β-lactoglobulin-retinol complex. *Biochimica et Biophysica Acta* **625**.

Fushman, D., (2003) Determination of Protein Dynamics using ^{15}N Relaxation Measurements. In: *BioNMR in Drug Research.* O. Zerbe (ed). Weinheim: VCH-Wiley, pp. 283-308.

Futterman, S. & J. Heller, (1972) The Enhancement of Fluorescence and the Decreased Susceptibility to Enzymatic Oxidation of Retinol Complexed with Bovine

Serum Albumin, β-Lactoglobulin, and the Retinol-binding Protein of Human Plasma *Journal of Biological Chemistry* **247**: 5168-5172.

Geier, G. E. & P. Modrich, (1979) Recognition Sequence of the *dam* Methylase of *Escherichia coli* K12 and Mode of Cleavage of *Dpn* 1 Endonuclease. *The Journal of Biological Chemistry* **254**: 1408-1413.

Grzesiek, S. & A. Bax, (1992a) Correlating Backbone Amide and Side Chain Resonances in Larger Proteins by Multiple Relayed Triple Resonance NMR. *Journal of the American Chemical Society* **114**: 6291-6293.

Grzesiek, S. & A. Bax, (1992b) An Efficient Experiment for Sequential Backbone Assignment of Medium-Sized Isotopically Enriched Proteins. *Journal of Magnetic Resonance* **99**: 201-207.

Grzesiek, S. & A. Bax, (1992c) Improved 3D Triple-Resonance NMR Techniques Applied to a 31 kDa Protein. *Journal of Magnetic Resonance* **96**: 432-440.

Hambling, S. G., A. S. McAlpine & L. Sawyer, (1992) β-lactoglobulin. In: *Advanced Dairy Chemistry - Volume 1.* London: Elsevier Science Publishers Ltd., pp. 141-190.

Hanahan, D., (1983) Studies on Transformation of *Escherichia Coli* with Plasmids. *Journal of Molecular Biology* **166**: 557-580.

Hattori, M., K. Hiramatsu, T. Kurata, M. Nishiura, K. Takahashi, A. Ametani & S. Kaminogawa, (2005) Unfolding and refolding of bovine β-lactoglobulin monitored by hydrogen exchange measurements. *Biochimica Et Biophysica Acta* **1752**: 154-165.

Hemley, R., B. E. Kohler & P. Siviski, (1979) Absorption spectra for the complexes formed from vitamin-A and beta-lactoglobulin. *Biophysical Journal* **28**: 448.

Hill, J. P., W. C. Thresher, M. J. Boland, L. K. Creamer, S. G. Anema, G. A. Manderson, D. E. Otter, G. Paterson, R. Lowe, R. C. Burr, R. L. Motion, A. Winkelman & B. Wickham, (1997) The Polymorphism of the Milk Protein β-Lactoglobulin. A Review. In: *Milk Composition, Production and Biotechnology.* R. A. S. Welch, D. J. W. Burns, S. R. Davis, A. I. Popay & C. G. Prosser (eds). Wallingford: CAB International, pp. 173-202.

Hoedemaeker, F. J., R. W. Visschers, A. C. Alting, K. G. de Kruif, M. E. Kuil & J. P. Abrahams, (2002) A novel pH-dependent dimerization motif in *β*-lactoglobulin from pig (*Sus scrofa*). *Acta Crystallographica Section D: Biological Crystallography* **58**.

Huber, R., M. Schneider, O. Epp, I. Mayr, A. Messerschmidt, J. Pflugrath & H. Kayser, (1987) Crystallization, crystal structure analysis and preliminary molecular model of the bilin binding protein from the insect *Pieris brassicae*. *Journal of Molecular Biology* **195**: 423-424.

Invernizzi, G., E. Annoni, A. Natalello, S. M. Doglia & M. Lotti, (2008) *In vivo* aggregation of bovine β-lactoglobulin is affected by Cys at position 121. *Protein Expression and Purification* **62**: 111-115.

Jameson, G. B., J. J. Adams & L. K. Creamer, (2002) Flexibility, functionality and hydrophobicity of bovine *β*-lactoglobulin. *International Dairy Journal* **12**: 319-329.

Jayat, D., J. C. Gaudin, J. M. Chobert, T. V. Burova, C. Holt, I. McNae, L. Sawyer & T. Haertle, (2004) A recombinant C121S mutant of bovine β-lactoglobulin is more susceptible to peptic digestion and to denaturation by reducing agents and heating. *Biochemistry* **43**.

Jenness, R., (1988) Composition of Milk. In: *Fundamentals of Dairy Chemistry.* New York: Van Nostrand Reinhold Company Inc, pp. 1-38.

REFERENCES

Joss, L. A. & G. B. Ralston, (1996) *β*-Lactoglobulin B: A Proposed Standard for the Study of Reversible Self-Association Reactions in the Analytical Ultracentrifuge? *Analytical Biochemistry* **236**: 20-26.

Kay, L. E., M. Ikura, R. Tschudin & A. Bax, (1990) Three-Dimensional Triple-Resonance NMR Spectroscopy of Isotopically Enriched Proteins. *Journal of Magnetic Resonance* **89**: 496-514.

Kay, L. E., P. Keifer & T. Saarinen, (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *Journal of the American Chemical Society* **114**: 10663-10665.

Kay, L. E., D. A. Torchia & A. Bax, (1989) Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application in staphlococcal nuclease. *Biochemistry* **28**: 8972-8979.

Kim, T. R., Y. Goto, N. Hirota, K. Kuwata, H. Denton, S. Y. Wu, L. Sawyer & C. A. Batt, (1997) High-level expression of bovine β-lactoglobulin in *Pichia pastoris* and characterization of its physical properties. *Protein Engineering* **10**: 1339-1345.

Kobayashi, T., M. Ikeguchi & S. Sugai, (2000) Molten globule structure of equine β-lactoglobulin probed by hydrogen exchange. *Journal of Molecular Biology* **299**: 757-770.

Kontopidis, G., C. Holt & L. Sawyer, (2002) The Ligand-binding Site of Bovine β-Lactoglobulin: Evidence for a Function? *Journal of Molecular Biology* **318**: 1043-1055.

Kontopidis, G., C. Holt & L. Sawyer, (2004) *Invited Review: β*-Lactoglobulin: Binding Properties, Structure, and Function. *Journal of Dairy Science* **87**: 785-796.

Kraulis, P. J., (1991) MOLSCRIPT: A Program to Produce Both Detailed and Schematic Plots of Protein Structures. *Journal of Applied Crystallography* **24**: 946-950.

Křížová, H., L. Žídek, M. J. Stone, M. V. Novotny & V. Sklenár, (2004) Temperature-dependent spectral density analysis applied to monitoring backbone dynamics of major urinary protein-I complexed with the pheromone 2-*sec*-butyl-4,5-dihydrothiazole. *Journal of Biomolecular NMR* **28**: 369-384.

Kroenke, C. D., M. Rance & A. G. Palmer, (1999) Variability of the ^{15}N Chemical Shift Anistropy in *Escherichia coli* Ribonuclease H in Solution. *Journal of the American Chemical Society* **121**: 10119-10125.

Kuwata, K., M. Hoshino, V. Forge, S. Era, C. A. Batt & Y. Goto, (1999) Solution structure and dynamics of bovine *β*-lactoglobulin A. *Protein Science* **8**: 2541-2545.

Laemmli, U. K., (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680-685.

Larson, B. L., (1979) Biosynthesis and Secretion of Milk Proteins: A Review. *Journal of Dairy Research* **46**: 161-174.

Levy, R., W. R., G. Chen, B. L. Iverson & G. Georgiou, (2001) Production of correctly folded Fab antibody fragment in the cytoplasm of *Escherichia coli trxB gor* mutants via the coexpression of molecular chaperones. *Protein Expression and Purification* **23**: 338-347.

Lipari, G. & A. Szabo, (1982a) Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. In: Journal of the American Chemical Society. pp. 4546-4559.

Lipari, G. & B. Szabo, (1982b) Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 2. Analysis of Experimental Results. *Journal of the American Chemical Society* **104**: 4559-4570.

Luo, M., (2004) Got Milk? A pheromonal message for newborn rabbits. *BioEssays* **26**: 6-9.

Mailliart, P. & B. Ribadeau-Dumas, (1988) Preparation of β-Lactoglobulin and β-Lactoglobulin-Free Proteins from Whey Retentate by NaCl Salting Out at Low pH. *Journal of Food Science* **53**: 743-745, 752.

Mandel, A. M., M. Akke & A. G. Palmer, (1995) Backbone Dynamics of *Escherichia-Coli* Ribonuclease HI - Correlations with Structure and Function in an Active Enzyme. *Journal of Molecular Biology* **246**: 144-163.

Manderson, G. A., L. K. Creamer & M. J. Hardman, (1999a) Effect of Heat Treatment on the Circular Dichroism Spectra of Bovine β-Lactoglobulin A, B, and C. *Journal of Agricultural and Food Chemistry* **47**: 4557-4567.

Manderson, G. A., M. J. Hardman & L. K. Creamer, (1997) Spectroscopic examination of the heat-induced changes in β-lactoglobulin A, B and C. *Milk Protein Polymorphism*: 204-211

Manderson, G. A., M. J. Hardman & L. K. Creamer, (1998) Effect of heat treatment on the conformation and aggregation of β-lactoglobulin A, B and C. *Journal of Agricultural and Food Chemistry* **46**: 4052-4061.

Manderson, G. A., M. J. Hardman & L. K. Creamer, (1999b) Effect of heat treatment on bovine β-lactoglobulin A, B and C explored using thiol availability and fluorescence. *Journal of Agricultural and Food Chemistry* **47**: 3617-3627.

Marion, D., P. C. Driscoll, L. E. Kay, P. T. Wingfield, A. Bax, A. M. Gronenborn & G. M. Clore, (1989a) Overcoming the Overlap Problem in the Assignment of ^{1}H NMR Spectra of Larger Proteins by Use of Three-Dimensional Heteronuclear ^{1}H-^{15}N Hartmann-Hahn Multiple Quantum Coherence and Nuclear Overhauser Multiple Quantum Coherence Spectroscopy: Application to Interleukin 1β. *Biochemistry* **28**: 6150-6156.

Marion, D., M. Ikura, R. Tschudin & A. Bax, (1989c) Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *Journal of Magnetic Resonance* **85**: 393-399.

Marion, D., L. E. Kay, S. W. Sparks, D. A. Torchia & A. Bax, (1989b) Three-Dimensional Heteronuclear NMR of ^{15}N-Labeled Proteins. *Journal of the American Chemical Society* **111**: 1515-1517.

Marion, D. & K. Wüthrich, (1983) Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^{1}H-^{1}H spin-spin coupling constants in proteins. *Biochemical and Biophysical Research Communications* **113**: 967-974.

Millet, O., J. P. Loria, C. D. Kroenke, M. Pons & A. G. Palmer, (2000) The static magnetic field dependence of chemical exchange line broadening defines the NMR chemical shift time scale. *Journal of American Chemical Society* **122**: 2867-2877.

Molinari, H., L. Ragona, L. Varani, G. Musco, R. Consonni, L. Zetta & H. L. Monaco, (1996) Partially folded structure of monomeric bovine β-lactoglobulin. *Febs Letters* **381**: 237-243.

Muhandiram, D. R. & L. E. Kay, (1994) Gradient-Enhanced Triple-Resonance Three-Dimensional NMR Experiments with Improved Sensitivity. *Journal of Magnetic Resonance Series B* **103**: 203-216.

Nakamura, Y., K. Wada, Y. Wada, H. Doi, S. Kanaya, T. Gojobori & T. Ikemura, (1995) Codon usage tabulated from the international DNA sequence databases. *Nucleic Acids Research* **24**: 214-215.

Nelson, M. & M. McClelland, (1992) Use of DNA methyltransferase/endonuclease enzyme combinations for megabase mapping of chromosomes. *Methods in Enzymology* **216**: 279-303.

REFERENCES

Oksanen, E., V. P. Jaakola, T. Tolonen, K. Valkonen, B. Akerstrom, N. Kalkkinen, V. Virtanen & A. Goldman, (2006) Reindeer β-lactoglobulin crystal structure with pseudo-body-centred noncrystallographic symmetry. *Acta Crystallographica Section D: Biological Crystallography* **D62**: 1369-1374.

Palmer, A. G., (1993) Dynamic properties of proteins from NMR spectroscopy. *Current Opinion in Biotechnology* **4**: 385-391.

Palmer, A. G., (1997) Probing molecular motion by NMR. *Current Opinion in Structural Biology* **7**: 732-737.

Palmer, A. G., M. Rance & P. E. Wright, (1991) Intramolecular Motions of a Zinc Finger DNA-Binding Domain from Xfin Characterized by Proton-Detected Natural Abundance ^{13}C Heteronuclear NMR Spectroscopy. *Journal of the American Chemical Society* **113**: 4371-4380.

Papaworth, C., J. C. Bauer, J. Braman & D. A. Wright, (1992) Site-directed mutagenesis in one day with >80% efficiency. *Strategies* **9**: 3-4.

Papiz, M. Z., L. Sawyer, E. E. Eliopoulos, A. C. North, J. B. Findlay, R. Sivaprasadarao, T. A. Jones, M. E. Newcomer & P. J. Kraulis, (1986) The structure of β-lactoglobulin and its similarity to plasma retinol-binding protein. *Nature* **324**: 383-385.

Pérez, M. D., C. Díaz de Villegas, L. Sánchez, P. Aranda, J. M. Ena & M. Calvo, (1989) Interaction of Fatty Acids with β-Lactoglobulin and Albumin from Ruminant milk. *The Journal of Biochemistry* **106**.

Pérez, M. D., P. Puyol, J. M. Ena & M. Calvo, (1993) Comparison of the ability to bind lipids of *β*-lactoglobulin and serum albumin of milk from ruminant and non-ruminant species. *Journal of Dairy Research* **60**: 55-63.

Pérez, M. D., L. Sánchez, P. Aranda, J. M. Ena, R. Oria & M. Calvo, (1992) Effect of b-lactoglobulin on the activity of pregastric lipase. Apossible role for this protein in ruminant milk. *Biochimica Et Biophysica Acta - Lipids and Lipid Metabolism* **1123**: 151-155.

Pervaiz, S. & K. Brew, (1985) Homology of β-lactoglobulin, Serum Retinol-Binding Protein, and Protein HC. *Science* **228**: 335-337.

Ponniah, K., T. S. Loo, P. J. B. Edwards, S. M. Pascal, G. B. Jameson & G. E. Norris, (2010) The production of soluble and correctly folded recombinant bovine β-lactoglobulin variants A and B in Escherichia coli for NMR studies. *Protein Expression and Purification* **70**: 283-289.

Qin, B. Y., M. C. Bewley, L. K. Creamer, E. N. Baker & G. B. Jameson, (1999) Functional implications of structural differences between variants A and B of bovine *β*-lactoglobulin. *Protein Science* **8**: 75-83.

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

Qin, B. Y., L. K. Creamer, E. N. Baker & G. B. Jameson, (1998b) 12-Bromododecanoic acid binds inside the calyx of bovine β-lactoglobulin. *FEBS Letters* **438**: 272-278.

Ragona, L., F. Fogolari, L. Zetta, M. D. Pérez, P. Puyol, K. G. De Kruif, F. Löhr, H. Rüterjans & H. Molinari, (2000) Bovine β-Lactoglobulin: Interaction studies with palmitic acid. *Protein Science* **9**: 1347-1358.

Ragona, L., F. Pusterla, L. Zetta, H. L. Monaco & H. Molinari, (1997) Identification of a conserved hydrophobic cluster in partially folded bovine β-lactoglobulin at pH 2. *Folding and Design* **2**: 281-290.

Rocha, T. L., G. Paterson, K. Crimmins, A. Boyd, L. Sawyer & Fothergill-Gilmore, (1996) Expression and secretion of recombinant ovine β-lactoglobulin in

Saccharomyces cerevisiae and *Kluyveromyces lactis*. *Biochemistry Journal* **313**: 927-932.

Rule, G. S. & K. T. Hitchens, (2006) *Fundamentals of Protein NMR Spectroscopy*. Springer, The Netherlands.

Said, H. M., D. E. Ong & J. L. Shingleton, (1989) Intestinal uptake of retinol: enhancement by bovine milk *β*-lactoglobulin. *American Journal of Clinical Nutrition* **40**: 690-694.

Sakurai, K. & Y. Goto, (2002) Manipulating Monomer-Dimer Equilibrium of Bovine *β*-Lactoglobulin by Amino Acid Substitution. *Journal of Biological Chemistry* **277**: 25735-25740.

Sakurai, K. & Y. Goto, (2006) Dynamics and Mechanism of the Tanford Transition of Bovine β-Lactoglobulin Studied using Heteronuclear NMR Spectroscopy. *Journal of Molecular Biology* **356**: 483-496.

Sakurai, K. & Y. J. Goto, (2007) Principal component analysis of the pH-dependent conformational transitions of bovine *β*-lactoglobulin monitored by heteronuclear NMR. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 15346-15351.

Sakurai, K., T. Konuma, M. Yagi & Y. Goto, (2009) Structural dynamics and folding of β-lactoglobulin probed by heteronuclear NMR. *Biochimica Et Biophysica Acta* **1790**: 527-537.

Sakurai, K., M. Oobatake & Y. Goto, (2001) Salt-dependent monomer–dimer equilibrium of bovine β-lactoglobulin at pH 3. *Protein Science* **10**: 2325-2335.

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

Schwartz, G., (1978) Estimating the dimension of a model. *The annals of statistics* **6**: 461-464.

Scopes, R. K., (1974) Measurement of Protein by Spectrophotometry at 205 nm. *Analytical Biochemistry* **59**: 277-282.

Tanford, C., L. G. Bunville & Y. Nozaki, (1959) The reversible transformation of b-lactoglobulin at pH 7.5. *Journal of the American Chemical Society* **81**.

Taulier, N. & T. V. Chalikian, (2001) Characterization of pH-induced transitions of β-lactoglobulin: ultrasonic, densimetric, and spectroscopic studies. *Journal of Molecular Biology* **314**: 873-889.

Tegoni, M., R. Ramoni, E. Bignetti, S. Spinelli & C. Cambillau, (1996) Domain swapping creates a third putative combining site in bovine odorant binding protein dimer. *Nature Structural Biology* **3**: 863-867.

Timm, D. E., L. J. Baker, H. Mueller, L. Zidek & M. V. Novotny, (2001) Structural basis of pheromone binding to mouse major urinary protein (MUP-I). *Protein Science* **10**: 997-1004.

Tjandra, N., S. E. Feller, R. W. Pastor & A. Bax, (1995) Rotational Diffusion Anistropy of Human Ubiquitin from ^{15}N NMR Relaxation. *Journal of the American Chemical Society* **117**: 12562-12566.

Totsuka, M., Y. Katakura, M. Shimizu, I. Kumagai, K. Miura & S. Kaminogawa, (1990) Expression and Secretion of Bovine β-Lactoglobulin in *Saccharomyces cerevisiae*. *Agricultural and Biological Chemistry* **54**: 3111-3116.

Ugolini, R., L. Ragona, E. Silletti, F. Fogolari, R. W. Visschers, A. C. Alting & H. Molinari, (2001) Dimerisation, stability and electrostatic properties of porcine β-lactoglobulin. *European Journal of Biochemistry* **268**: 4477-4488.

Uhrínová, S., M. H. Smith, G. B. Jameson, D. Uhrín, L. Sawyer & P. N. Barlow, (2000) Structural Changes Accompanying pH-Induced Dissociation of the *β*-Lactoglobulin Dimer. *Biochemistry* **39**: 3565-3574.

REFERENCES

Uhrínová, S., D. Uhrín, H. Denton, M. Smith, L. Sawyer & P. N. Barlow, (1998) Complete assignment of ^{1}H, ^{13}C and ^{15}N chemical shifts for bovine β-lactoglobulin: Secondary structure and topology of the native state is retained in a partially unfolded form. *Journal of Biomolecular NMR* **12**: 89-107.

Vranken, W. F., W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, P. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides & E. D. Laue, (2005) The CCPN Data model for NMR Spectroscopy: Development of a Software Pipeline. *Proteins: Structure, Function, and Bioinformatics* **59**: 687-696.

Wu, S. Y., M. D. Pérez, P. Puyol & L. Sawyer, (1999) β-Lactoglobulin Binds Palmitate Within its Central Cavity. *Journal of Biological Chemistry* **274**: 170-174.

Zapun, A., D. Missiakas, S. Raina & T. E. Creighton, (1995) Structural and Functional Characterization of DsbC, a Protein Involved in Disulfide Bond Formation in *Escherichia coli*. *Biochemistry* **34**: 5075-5089.

Zhang, F. L. & R. Brűschweiler, (2002) Contact Model for the Prediction of NMR N−H Order Parameters in Globular Proteins. *Journal of the American Chemical Society* **124**: 12654-12655.

A

Molecular Biology

A.1 General Chemicals Used

Table A.1 outlays the general chemicals used throughout this study, the origins and suppliers.

Table A.1 General chemicals used.

Chemical	Formula and/or Abbreviation	Supplier	Origin
Agar	-	Oxoid Ltd	Hampshire, England
Agarose	Type ST/L	Pure Science Ltd	Wellington, NZ
Potassium dihydrogen phosphate	KH_2PO_4	BDH	Poole, England
Disodium hydrogen phosphate	Na_2HPO_4	Ajax Chemicals	Auckland, NZ
Sodium dihydrogen phosphate	NaH_2PO_4	Ajax Chemicals	Auckland, NZ
Dipotassium hydrogen phosphate	K_2HPO_4	Ajax Chemicals	Auckland, NZ
Tetramethylethylenediamine	TEMED/ $(CH_3)_2NCH_2CH_2N(CH_3)_2$	BDH	Poole, England
Bis-tris propane	Bis-Tris buffer/$C_8H_{19}NO_5$	Lancaster Synthesis Ltd.	Morecombe, England
3-(N-Morpholino)-propanesulfonic acid	MOPs buffer /$C_7H_{15}NO_4S$	Sigma	Auckland, NZ
Acetic acid	CH_3COOH	BDH	Poole, England
Ammonium chloride	NH_4Cl	BDH	Poole, England
Ammonium persulfate	APS/$(NH_4)_2S_2O_8$	Ajax Chemicals	Auckland, NZ
β-mercaptothanol	BME /C_2H_6OS	Sigma	Auckland, NZ
Bromophenol blue	$C_{19}H_{10}Br_4O_5S$	Affymetrix	Cleveland, USA
Calcium chloride	$CaCl_2$		
D-glucose	Drose/$C_6H_{12}O_6$	Ajax Chemicals	Auckland, NZ
Ethanol	EtOH	Merck	Darmstadt, GER
Ethidium bromide	$C_{21}H_2OBrN_3$ / EtBr		
Ethylene diamine tetra-acetic acid	EDTA	Sigma	Auckland, NZ
Glycerol	$C_3H_5(OH)_3$	BDH	Poole, England
Glycine	$C_2H_5NO_2$	Sigma	Auckland, NZ
Hydrochloric acid	HCl	-	-
Isopropyl-β-D-Thiogalactopyranoside	IPTG	Sigma	Auckland, NZ

Continued on next page

Table A.1 – continued from previous page

Chemical	Formula and/or Abbreviation	Supplier	Origin
Luria Broth powder	LB	Invitrogen	Scotland
Magnesium sulfate	$MgSO_4$	M&B	Victoria, Australia
Orthophosphoric acid	H_3PO_4	Ajax Chemicals	Auckland, NZ
Potassium acetate	CH_3COOK	BDH	Poole, England
Potassium hydroxide	KOH	ProLabo	Paris, France
Rubidium chloride	RbCl	Sigma	Auckland, NZ
Sodium azide	NaN_3	BDH	Poole, England
Sodium chloride	NaCl	Pure Science	Wellington, NZ
Sodium dodecyl sulfate	SDS/ $NaC_{12}H_{25}SO_4$	BDH	Poole, England
Sodium hydroxide	NaOH	Ajax Chemicals	Auckland, NZ
Thiamine	Vitamin B_1	Sigma	Auckland, NZ
Tris (hydroxymethyl) aminomethane	$C_4H_{11}NO_3$ / Tris buffer	Ajax Chemicals	Auckland, NZ

A.2 The Genetic Code

The following translation table was used to generate rationalised site-directed mutations. Table A.2 is sourced from http://ewydanie.nczas.com.

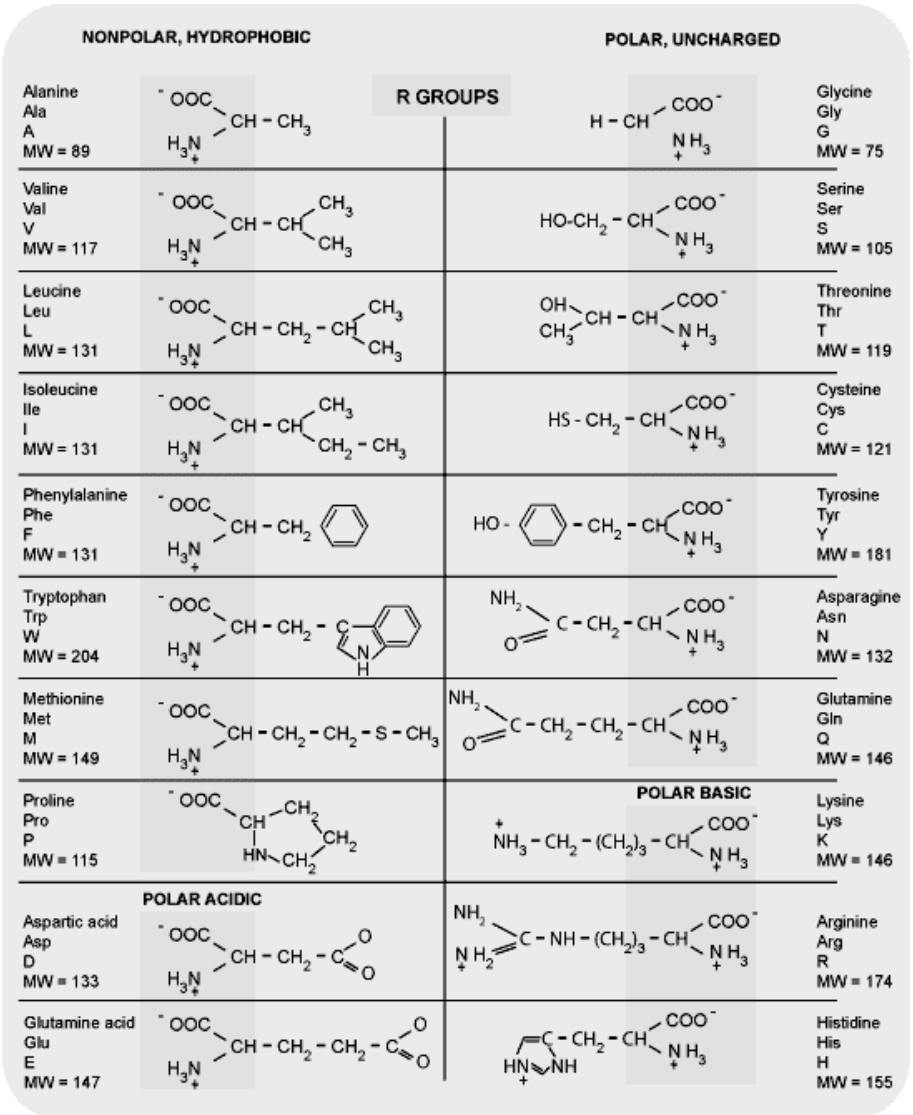
Table A.2 Translation Table

First Position	Second Position: U	Second Position: C	Second Position: A	Second Position: G	Third Position
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	C
	UUA Leu	UCA Ser	UAA *Stop*	UGA *Stop*	A
	UUG Leu	UCG Ser	UAG *Stop*	UGG Trp	G
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U
	CUC Leu	CCC Pro	CAC His	CGC Arg	C
	CUA Leu	CCA Pro	CAA Gln	CGA Arg	A
	CUG Leu	CCG Pro	CAG Gln	CGG Arg	G
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser	U
	AUC Ile	ACC Thr	AAC Asn	AGC Ser	C
	AUA Ile	ACA Thr	AAA Lys	AGA Arg	A
	AUG Met	ACG Thr	AAG Lys	AGG Arg	G
G	GUU Val	GCU Ala	GAU Asp	GGU Gly	U
	GUC Val	GCC Ala	GAC Asp	GGC Gly	C
	GUA Val	GCA Ala	GAA Glu	GGA Gly	A
	GUG Val	GCG Ala	GAG Glu	GGG Gly	G

A.3 Structures & Abbreviations of Standard Amino Acids

Table A.3 is a list of standard amino acids with their appropriate structure, one and three letter codes. Table A.3 is sourced from: http://biotech.matcmadison.edu/resources/proteins/labManual/images/amino_000.gif.

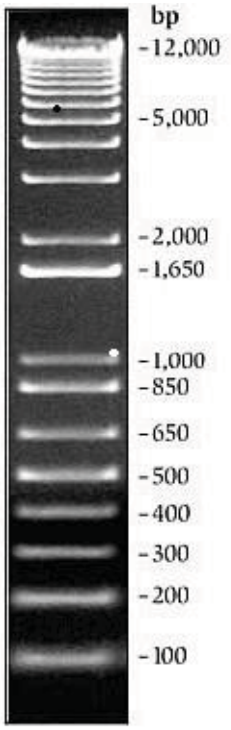
Table A.3 Structures and symbols of standard amino acids.



A.4 DNA Ladder and Protein Molecular Weight Marker

Below are the markers that were used to estimate the length of linear DNA and protein molecular weights.

1 Kb plus DNA ladder

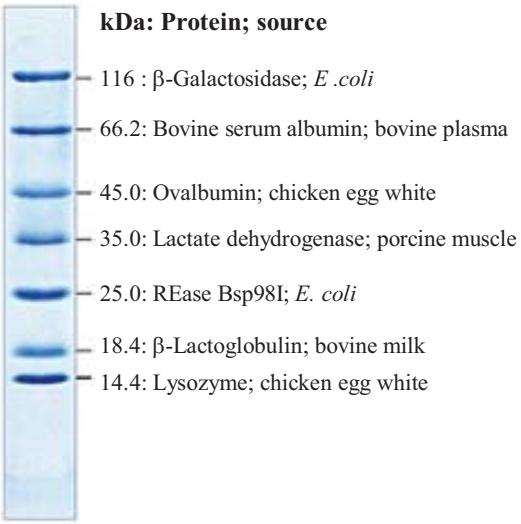


DNA 0.9 μg/lane

0.9 % agarose
stained with
ethidium bromide

Invitrogen

Protein Molecular Weight Marker



Protein 0.1-0.2 mg/mL

8-16 % Tris-glycine
stained with Page Blue
Protein Staining Solution

Fermentas Life Sciences

A.5 Synthetic β-Lg A Sequence

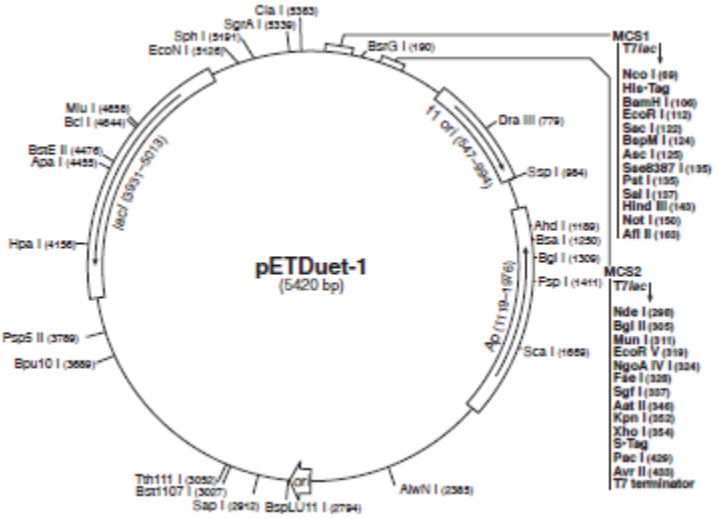
Below is the synthetic BLG A DNA sequence and the respective translated primary sequence (Ariyaratne *et al*., 2002).

```
     L   I   V   T   Q   T   M   K   G   L   D   I   Q   K   V   A   G   T   W   Y
1   ctg att gtg acc cag acc atg aaa ggt ctg gat att cag aaa gtg gcg ggt acg tgg tat  60

     S   L   A   M   A   A   S   D   I   S   L   L   D   A   Q   S   A   P   L   R
61  agc ctg gcc atg gct gcc agc gat att agt ctg ctg gat gcg cag tcg gcg ccg ctg cgt 120

     V   Y   V   E   E   L   K   P   T   P   E   G   D   L   E   I   L   L   Q   K
121 gtg tac gtg gaa gag ctc aag ccg acc ccg gaa ggc gat ctc gag att ctg ctg cag aaa 180

     W   E   N   D   E   C   A   Q   K   K   I   I   A   E   K   T   K   I   P   A
181 tgg gaa aac gat gaa tgt gcg cag aag aag att atc gct gag aag acc aag att ccg gct 240

     V   F   K   I   D   A   L   N   E   N   K   V   L   V   L   D   T   D   Y   K
241 gtg ttt aag atc gat gcg ctg aac gag aac aaa gtg ctg gtt cta gat acc gat tat aag 300

     K   Y   L   L   F   C   M   E   N   S   A   E   P   E   Q   S   L   V   C   Q
301 aaa tat ctg ctg ttc tgc atg gag aat tct gcg gag ccg gag cag tct ctg gtg tgc cag 360

     C   L   V   R   T   P   E   V   D   D   E   A   L   E   K   F   D   K   A   L
361 tgt ctg gtt cga acc ccg gaa gtc gac gat gaa gcg ctg gag aaa ttt gac aaa gcc ctg 420

     K   A   L   P   M   H   I   R   L   S   F   N   P   T   Q   L   E   E   Q   C
421 aag gcg ctg ccg atg cat att cgt ctg tct ttt aac ccg acc cag ctg gaa gag cag tgc 480

     H   I   *   *
481 cat att taa taa 492
```

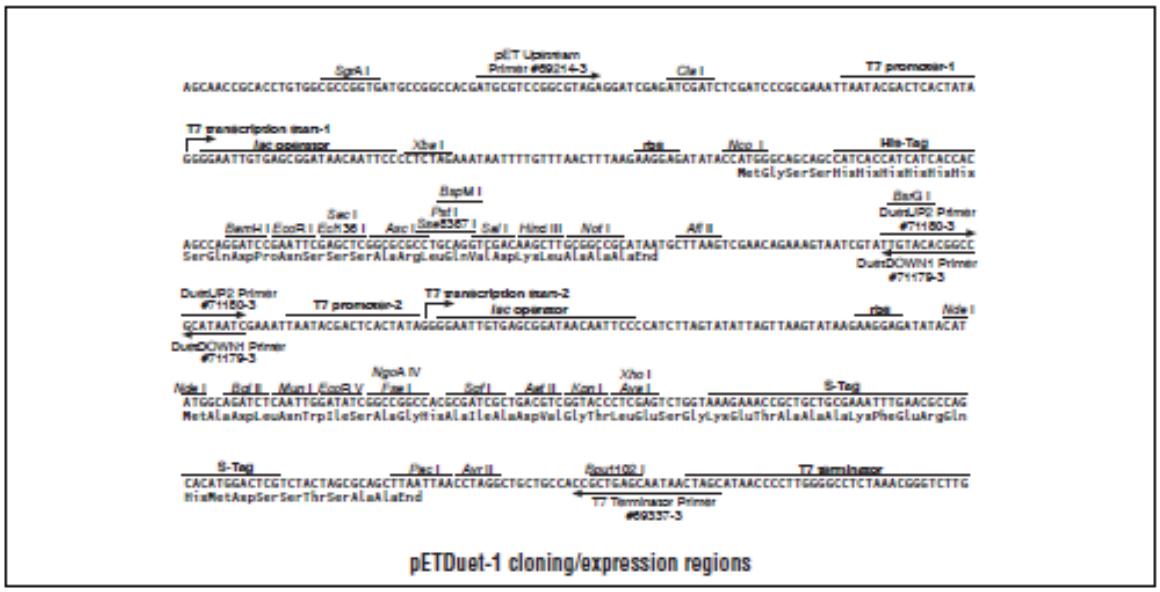
A.6 pETDuet-1 Vector Map

pETDuet-1 Vector Map (Novagen)

(A)



(B)



Maps of the pETDuet-1 Expression Vector from Novagen.

Figure (A) maps the whole plasmid of the pETDuet-1 expression vector used in the studies. Figure (B) shows a map of the pETDuet-1 cloning and expression regions.

B

Chemical Shift Tables

B.1 Chemical Shifts for β-Lg C at Three Temperatures

Table B.1 lists the NMR spectroscopy ^{1}H and ^{15}N backbone chemical shifts for β-Lg-C assigned at 305 K, 312 K and 320 K.

Table B.1 β-Lg C chemical shifts assigned at 305 K, 313 K and 320 K

	β-Lg C 305 K		β-Lg C 313 K		β-Lg C 320 K	
Residue	**H^N**	**N**	**H^N**	**N**	**H^N**	**N**
Met0	-	-	-	-	-	-
Leu1	-	-	-	-	-	-
Ile2	-	-	-	-	-	-
Val3	8.10	125.22	8.02	124.82	7.94	124.45
Thr4	7.99	116.55	7.92	116.31	7.85	116.13
Gln5	8.29	123.71	8.21	123.50	-	-
Thr6	-	-	-	-	-	-
Met7	7.72	120.48	7.66	120.44	7.61	120.39
Lys8	8.63	127.68	8.53	127.18	8.44	126.67
Gly9	8.57	109.62	8.46	109.31	8.36	109.05
Leu10	7.55	120.15	7.49	120.04	7.45	119.99
Asp11	8.60	126.57	8.55	126.53	8.49	126.45
Ile12	8.89	123.12	8.80	123.04	8.72	122.98
Gln13	7.92	117.23	7.90	117.27	7.87	117.28
Lys14	7.18	113.61	7.16	113.65	7.13	113.69
Val15	7.39	110.33	7.35	110.36	7.32	110.41
Ala16	6.79	120.96	6.77	120.96	6.74	120.95
Gly17	9.13	110.11	9.10	109.94	9.06	109.81
Thr18	8.39	118.40	8.31	118.17	8.23	117.96
Trp19	7.47	124.07	7.45	123.98	7.42	123.88
Tyr20	9.14	112.91	9.10	112.91	9.05	112.94
Ser21	9.94	120.28	9.85	120.30	9.76	120.32
Leu22	9.04	129.87	8.97	129.48	8.90	129.08
Ala23	7.65	114.78	7.62	114.71	7.59	114.66
Met24	9.00	115.30	8.96	115.33	-	-
Ala25	9.21	123.31	9.18	123.23	9.15	123.15
Ala26	8.38	120.02	8.36	120.00	8.33	119.96
Ser27	8.53	116.13	8.50	116.07	8.47	116.04
Asp28	-	-	-	-	-	-
Ile29	-	-	-	-	-	-
Ser30	7.93	114.35	7.90	114.34	7.86	114.34
Leu31	7.64	117.17	7.59	117.07	7.55	117.00
Leu32	7.08	111.50	7.06	111.58	7.04	111.64

Continued on next page

Table B.1 – continued from previous page

Residue	β-Lg C 305 K		β-Lg C 313 K		β-Lg C 320 K	
	H^N	N	H^N	N	H^N	N
Asp33	-	-	-	-	-	-
Ala34	7.63	120.70	7.59	120.66	7.56	120.58
Gln35	-	-	-	-	-	-
Ser36	7.70	108.01	7.65	107.93	7.60	107.92
Ala37	7.48	126.17	7.44	126.09	7.40	126.00
Pro38	-	-	-	-	-	-
Leu39	7.84	112.05	7.80	111.96	7.76	111.91
Arg40	7.57	123.63	7.53	123.55	7.50	123.44
Val41	6.93	116.13	6.91	116.19	6.89	116.24
Tyr42	-	-	-	-	-	-
Val43	8.98	127.95	8.91	127.68	8.85	127.46
Glu44	-	-	-	-	-	-
Glu45	7.60	115.57	7.57	115.48	7.54	115.41
Leu46	8.66	122.13	8.62	122.19	8.58	122.23
Lys47	9.14	122.59	9.10	122.56	9.06	122.50
Pro48	-	-	-	-	-	-
Thr49	8.56	114.27	8.52	114.14	8.47	113.99
Pro50	-	-	-	-	-	-
Glu51	7.43	112.22	7.39	112.16	-	-
Gly52	8.26	106.88	8.23	106.88	8.19	106.93
Asp53	7.16	115.01	7.11	114.91	7.06	114.86
Leu54	8.78	119.44	8.72	119.44	8.67	119.42
Glu55	8.60	126.05	8.64	125.93	8.65	125.79
Ile56	9.26	126.56	9.22	126.49	9.17	126.38
Leu57	8.54	128.25	8.51	128.26	8.48	128.24
Leu58	8.98	122.41	8.95	122.45	8.91	122.46
His59	9.23	115.88	9.22	115.86	9.19	115.83
Lys60	8.95	121.02	-	-	8.90	120.89
Trp61	-	-	-	-	-	-
Glu62	8.91	126.11	8.86	125.95	8.81	125.78
Asn63	9.09	123.07	8.99	122.88	-	-
Gly64	7.34	126.28	7.06	125.88	-	-
Glu65	7.07	116.99	7.08	116.99	7.08	116.99
Cys66	-	-	-	-	-	-
Ala67	8.82	108.49	8.78	108.13	8.73	107.75
Gln68	-	-	-	-	-	-
Lys69	-	-	-	-	-	-
Lys70	8.49	123.59	8.42	123.51	8.36	123.41

Continued on next page

Table B.1 – continued from previous page

Residue	β-Lg C 305 K		β-Lg C 313 K		β-Lg C 320 K	
	H^N	N	H^N	N	H^N	N
Ile71	8.94	124.74	8.90	124.50	8.85	124.29
Ile72	8.50	126.18	8.43	126.06	8.37	125.90
Ala73	9.23	109.25	9.20	109.13	9.17	108.96
Glu74	9.42	124.54	9.37	124.46	9.33	124.40
Lys75	8.38	123.08	8.31	123.08	8.26	123.06
Thr76	8.42	112.64	8.36	112.55	8.30	112.50
Lys77	8.24	116.59	8.18	116.58	8.12	116.54
Ile78	7.96	122.34	7.95	122.27	7.93	122.18
Pro79	-	-	-	-	-	-
Ala80	7.83	115.13	7.79	115.13	7.75	115.15
Val81	7.30	117.49	7.27	117.45	7.23	117.38
Phe82	9.53	126.35	9.49	126.32	9.45	126.25
Lys83	9.45	123.21	9.40	123.12	9.36	123.04
Ile84	7.70	115.87	-	-	-	-
Asp85	8.80	119.11	8.72	119.06	8.65	118.99
Ala86	7.22	121.99	7.19	121.92	7.17	121.88
Leu87	9.00	114.74	-	-	-	-
Asn88	8.68	111.67	8.62	111.58	8.56	111.50
Glu89	8.23	117.23	8.18	117.17	8.13	117.13
Asn90	9.06	115.74	9.01	115.59	-	-
Lys91	8.56	121.34	-	-	-	-
Val92	9.00	121.26	-	-	-	-
Leu93	9.40	125.99	9.37	125.97	9.33	125.94
Val94	9.33	124.09	9.27	124.07	9.21	124.02
Leu95	8.42	128.13	8.37	128.05	8.33	127.95
Asp96	7.17	109.55	7.14	109.52	7.11	109.54
Thr97	7.78	118.76	-	-	-	-
Asp98	6.74	122.82	6.73	122.77	6.72	122.78
Tyr99	9.02	117.85	8.99	117.73	8.96	117.56
Lys100	8.98	119.45	8.93	119.43	8.89	119.43
Lys101	8.95	118.10	8.91	118.00	8.91	118.02
Tyr102	8.76	115.89	8.70	115.75	8.66	115.68
Leu103	9.13	125.13	9.10	125.13	9.07	125.11
Leu104	9.32	124.94	9.27	125.00	9.22	125.01
Phe105	8.89	120.96	8.86	120.95	8.83	120.91
Cys106	9.66	116.88	9.61	116.90	9.57	116.94
Met107	9.83	121.01	9.80	121.00	9.77	120.98
Glu108	8.41	114.16	8.39	114.15	8.36	114.16

Continued on next page

Table B.1 – continued from previous page

	β-Lg C 305 K		β-Lg C 313 K		β-Lg C 320 K	
Residue	**H^N**	**N**	**H^N**	**N**	**H^N**	**N**
Asn109	9.36	120.28	9.29	120.25	9.23	120.22
Ser110	9.80	122.28	9.75	122.18	9.69	122.08
Ala111	8.27	123.41	-	-	-	-
Glu112	7.72	113.11	7.68	113.11	7.64	113.12
Pro113	-	-	-	-	-	-
Glu114	8.33	114.23	8.28	114.07	8.23	113.97
Gln115	7.76	115.97	-	-	-	-
Ser116	7.55	109.58	7.52	109.51	7.49	109.51
Leu117	7.22	126.64	7.19	126.56	7.15	126.47
Ala118	8.76	128.45	8.72	128.43	8.69	128.33
Cys119	9.49	118.78	9.45	118.71	9.40	118.69
Gln120	9.22	118.60	9.18	118.59	9.14	118.59
Cys121	7.82	120.00	-	-	-	-
Leu122	9.71	128.32	9.65	128.28	9.61	128.26
Val123	1.40	111.32	9.41	111.33	9.38	111.38
Arg124	-	-	-	-	-	-
Thr125	7.47	108.16	7.43	108.18	7.40	108.21
Pro126	-	-	-	-	-	-
Glu127	8.29	119.82	8.24	119.74	8.18	119.71
Val128	-	-	-	-	-	-
Asp129	6.58	123.52	6.60	123.39	6.59	123.31
Asp130	-	-	-	-	-	-
Glu131	8.35	121.14	8.28	120.93	8.22	120.75
Ala132	7.53	121.26	7.53	121.22	7.53	121.16
Leu133	7.75	116.62	7.72	116.59	7.69	116.55
Glu134	7.98	119.61	-	-	-	-
Lys135	-	-	-	-	-	-
Phe136	8.54	121.62	-	-	-	-
Asp137	8.74	117.01	8.69	116.93	8.64	116.88
Lys138	-	-	-	-	-	-
Ala139	7.80	121.79	7.77	121.71	7.75	121.62
Leu140	7.45	114.26	7.41	114.16	7.36	114.08
Lys141	7.13	119.82	7.10	119.72	7.07	119.64
Ala142	7.87	118.46	7.81	118.43	7.76	118.42
Leu143	7.73	119.66	7.69	119.54	7.65	119.45
Pro144	-	-	-	-	-	-
Met145	7.68	118.21	7.64	118.23	7.60	118.25
His146	8.73	118.13	8.67	118.15	8.61	118.17

Continued on next page

Table B.1 – continued from previous page

	β-Lg C 305 K		β-Lg C 313 K		β-Lg C 320 K	
Residue	**H^N**	**N**	**H^N**	**N**	**H^N**	**N**
Ile147	7.71	117.46	7.68	117.35	7.66	117.27
Arg148	8.33	126.59	8.26	126.49	8.20	126.37
Leu149	-	-	-	-	-	-
Ser150	8.17	116.31	8.11	116.15	8.05	116.05
Phe151	8.05	120.41	7.98	119.94	8.05	120.32
Asn152	8.65	119.72	8.60	119.66	8.55	119.60
Pro153	-	-	-	-	-	-
Thr154	-	-	-	-	-	-
Gln155	-	-	-	-	-	-
Leu156	-	-	-	-	-	-
Glu157	-	-	-	-	-	-
Glu158	7.37	119.19	7.34	118.89	7.32	118.62
Gln159	-	-	-	-	-	-
Cys160	-	-	-	-	-	-
His161	-	-	-	-	-	-
Ile162	-	-	-	-	-	-

B.2 Chemical Shifts for β-Lg Variants A, B and C

Table B.2 lists the NMR spectroscopy ^{1}H and ^{15}N backbone chemical shifts for β-Lg A, B and C, assigned at 305 K.

Table B.2 β-Lg chemical shifts for variants A, B and C

	β-Lg A 305 K		β-Lg B 305 K		β-Lg C 305 K	
Residue	**H^N**	**N**	**H^N**	**N**	**H^N**	**N**
Met0	-	-	-	-	-	-
Leu1	-	-	-	-	-	-
Ile2	-	-	-	-	-	-
Val3	8.09	125.29	-	-	8.10	125.22
Thr4	8.00	116.62	7.98	116.68	7.99	116.55
Gln5	8.31	123.85	-	-	8.29	123.71
Thr6	7.67	111.32	7.65	111.22	-	-
Met7	7.72	120.50	7.69	120.50	7.72	120.48
Lys8	8.63	127.61	8.61	127.62	8.63	127.68
Gly9	8.54	109.49	8.52	109.48	8.57	109.62
Leu10	7.53	120.16	7.51	120.15	7.55	120.15
Asp11	8.60	126.83	8.58	126.81	8.60	126.57
Ile12	8.85	122.92	8.83	122.87	8.89	123.12
Gln13	7.95	117.19	7.93	117.17	7.92	117.23
Lys14	7.20	113.76	7.17	113.75	7.18	113.61
Val15	7.37	110.27	7.35	110.26	7.39	110.33
Ala16	6.79	120.96	6.77	120.94	6.79	120.96
Gly17	9.17	110.21	9.14	110.16	9.13	110.11
Thr18	8.38	118.45	-	-	8.39	118.40
Trp19	7.43	124.07	7.41	124.03	7.47	124.07
Tyr20	9.09	112.17	9.06	112.12	9.14	112.91
Ser21	9.89	119.52	9.85	119.61	9.94	120.28
Leu22	9.19	131.67	9.17	107.26	9.04	129.87
Ala23	7.70	114.92	7.66	115.11	7.65	114.78
Met24	9.05	114.80	9.01	115.27	9.00	115.30
Ala25	9.29	123.47	9.21	123.30	9.21	123.31
Ala26	8.31	119.72	8.38	119.92	8.38	120.02
Ser27	8.34	116.26	8.57	116.01	8.53	116.13
Asp28	-	-	7.54	119.54	-	-
Ile29	8.79	124.33	8.83	124.86	-	-
Ser30	7.95	114.30	7.96	114.32	7.93	114.35
Leu31	7.58	117.14	7.60	117.16	7.64	117.17
Leu32	7.02	111.01	7.01	111.38	7.08	111.50

Continued on next page

Table B.2 – continued from previous page

	β-Lg A 305 K		β-Lg B 305 K		β-Lg C 305 K	
Residue	**H^N**	**N**	**H^N**	**N**	**H^N**	**N**
Asp33	7.74	118.68	7.72	118.67	-	-
Ala34	7.61	120.74	7.55	120.66	7.63	120.70
Gln35	-	-	8.98	121.25	9.00	121.26
Ser36	7.70	107.60	7.63	107.63	7.70	108.01
Ala37	7.51	126.31	7.47	126.24	7.48	126.17
Pro38	-	-	-	-	-	-
Leu39	7.74	114.18	7.80	112.05	7.84	112.05
Arg40	7.39	122.34	7.51	123.58	7.57	123.63
Val41	6.94	117.01	6.86	116.35	6.93	116.13
Tyr42	8.58	120.63	8.56	121.07	8.56	121.07
Val43	8.62	125.22	8.56	125.48	8.98	127.95
Glu44	8.95	121.96	8.92	121.91	-	-
Glu45	7.59	116.34	7.58	116.33	7.60	115.57
Leu46	8.64	122.07	8.62	122.07	8.66	122.13
Lys47	9.14	122.49	9.11	122.46	9.14	122.59
Pro48	-	-	-	-	-	-
Thr49	8.58	114.36	8.56	114.34	8.56	114.27
Pro50	-	-	-	-	-	-
Glu51	7.46	112.31	7.45	112.25	7.43	112.22
Gly52	8.27	107.09	8.24	107.07	8.26	106.88
Asp53	7.16	115.28	7.13	115.24	7.16	115.01
Leu54	8.79	119.44	8.77	119.42	8.78	119.44
Glu55	8.60	126.13	8.59	126.05	8.60	126.05
Ile56	9.25	126.53	9.23	126.50	9.26	126.56
Leu57	8.46	128.66	8.45	128.69	8.54	128.25
Leu58	9.04	123.38	9.01	123.31	8.98	122.41
Gln59His	9.22	117.77	9.16	117.76	9.23	115.88
Lys60	8.93	121.52	8.91	121.44	8.95	121.02
Trp61	9.73	129.91	9.71	129.92	9.70	128.58
Glu62	8.73	127.31	8.81	127.06	8.91	126.11
Asn63	8.95	121.79	8.99	122.45	9.09	123.07
Asp64Gly	8.13	111.01	7.05	125.88	7.34	126.28
Glu65	6.74	114.56	6.81	116.57	7.07	116.99
Cys66	8.65	120.25	8.65	120.82	-	-
Ala67	8.96	133.57	8.93	109.48	8.82	108.49
Gln68	-	-	8.34	118.59	-	-
Lys69	9.14	123.72	-	-	-	-
Lys70	8.53	123.77	8.50	123.72	8.49	123.59

Continued on next page

Table B.2 – continued from previous page

	β-Lg A 305 K		β-Lg B 305 K		β-Lg C 305 K	
Residue	**H^N**	**N**	**H^N**	**N**	**H^N**	**N**
Ile71	8.94	124.27	8.93	124.22	8.94	124.74
Ile72	8.49	125.56	8.46	125.58	8.50	126.18
Ala73	9.26	133.39	9.22	109.22	9.23	109.25
Glu74	9.42	124.55	9.39	124.59	9.42	124.54
Lys75	8.48	123.17	8.46	123.19	8.38	123.08
Thr76	8.42	112.60	8.39	112.62	8.42	112.64
Lys77	8.27	116.50	8.24	116.48	8.24	116.59
Ile78	7.96	122.33	7.93	122.29	7.96	122.34
Pro79	-	-	-	-	-	-
Ala80	7.86	115.22	7.85	115.19	7.83	115.13
Val81	7.29	117.44	7.26	117.47	7.30	117.49
Phe82	9.53	126.38	9.51	126.39	9.53	126.35
Lys83	9.46	123.27	9.44	123.28	9.45	123.21
Ile84	7.70	116.05	7.66	115.89	7.70	115.87
Asp85	8.79	119.21	8.77	119.23	8.80	119.11
Ala86	7.20	121.98	7.16	121.90	7.22	121.99
Leu87	9.01	114.54	8.99	114.43	9.00	114.74
Asn88	8.68	111.33	8.66	111.61	8.68	111.67
Glu89	8.24	117.37	8.22	117.15	8.23	117.23
Asn90	9.06	115.67	9.03	115.55	9.06	115.74
Lys91	-	-	-	-	8.56	121.34
Val92	-	-	-	-	9.00	121.26
Leu93	9.40	126.04	9.37	125.93	9.40	125.99
Val94	9.32	124.01	9.31	124.05	9.33	124.09
Leu95	8.43	128.05	8.41	128.06	8.42	128.13
Asp96	7.17	109.89	7.14	109.84	7.17	109.55
Thr97	7.78	119.02	7.75	119.00	7.78	118.76
Asp98	6.76	123.07	6.73	123.06	6.74	122.82
Tyr99	9.01	117.78	8.98	117.77	9.02	117.85
Lys100	8.98	119.58	8.96	119.53	8.98	119.45
Lys101	9.15	118.77	9.11	118.76	8.95	118.10
Tyr102	8.90	116.32	8.88	116.30	8.76	115.89
Leu103	9.14	125.04	9.12	125.03	9.13	125.13
Leu104	9.36	124.98	9.32	124.83	9.32	124.94
Phe105	8.87	121.23	8.85	121.04	8.89	120.96
Cys106	9.63	117.18	9.62	116.94	9.66	116.88
Met107	9.89	120.67	9.79	120.84	9.83	121.01
Glu108	8.38	114.01	8.40	113.98	8.41	114.16

Continued on next page

Table B.2 – continued from previous page

Residue	β-Lg A 305 K		β-Lg B 305 K		β-Lg C 305 K	
	H^N	N	H^N	N	H^N	N
Asn109	9.37	120.30	9.29	120.28	9.36	120.28
Ser110	9.79	122.17	9.81	122.26	9.80	122.28
Ala111	8.30	123.44	8.26	123.48	8.27	123.41
Glu112	7.72	113.24	7.69	113.06	7.72	113.11
Pro113	-	-	-	-	-	-
Glu114	8.35	114.34	-	-	8.33	114.23
Gln115	7.80	116.05	7.69	113.06	7.76	115.97
Ser116	7.57	109.57	7.52	109.68	7.55	109.58
Leu117	7.25	126.44	7.21	126.80	7.22	126.64
Ala118Val	8.78	126.08	8.72	128.31	8.76	128.45
Cys119	9.38	121.86	9.47	118.73	9.49	118.78
Gln120	9.33	118.65	9.21	118.53	9.22	118.60
Cys121	7.84	120.06	7.82	120.09	7.82	120.00
Leu122	9.73	128.55	9.69	128.48	9.71	128.32
Val123	9.43	111.41	9.40	111.42	1.40	111.32
Arg124	-	-	-	-	-	-
Thr125	7.49	108.42	7.46	108.35	7.47	108.16
Pro126	-	-	-	-	-	-
Glu127	8.27	120.02	8.25	120.06	8.29	119.82
Val128	8.30	118.87	8.30	119.03	-	-
Asp129	6.52	124.04	6.47	124.08	6.58	123.52
Asp130	9.07	123.85	9.06	123.94	-	-
Glu131	8.35	121.15	8.31	121.05	8.35	121.14
Ala132	7.59	121.43	7.54	121.40	7.53	121.26
Leu133	7.75	116.47	-	-	7.75	116.62
Glu134	-	-	7.99	119.82	7.98	119.61
Lys135	-	-	7.89	119.78	-	-
Phe136	-	-	8.50	121.67	8.54	121.62
Asp137	8.73	117.26	8.68	117.20	8.74	117.01
Lys138	-	-	-	-	-	-
Ala139	7.80	121.83	7.79	121.79	7.80	121.79
Leu140	7.47	114.34	7.44	114.31	7.45	114.26
Lys141	7.14	119.78	7.12	119.80	7.13	119.82
Ala142	7.83	118.40	7.82	118.39	7.87	118.46
Leu143	7.74	119.70	7.72	119.68	7.73	119.66
Pro144	-	-	-	-	-	-
Met145	7.67	118.09	7.64	118.18	7.68	118.21
His146	8.78	118.06	8.70	118.01	8.73	118.13

Continued on next page

Table B.2 – continued from previous page

	β-Lg A 305 K		β-Lg B 305 K		β-Lg C 305 K	
Residue	**H^N**	**N**	**H^N**	**N**	**H^N**	**N**
Ile147	7.70	117.40	7.67	117.37	7.71	117.46
Arg148	8.35	126.74	8.31	126.68	8.33	126.59
Leu149	9.13	124.12	-	-	9.13	123.73
Ser150	8.20	116.60	8.17	116.59	8.17	116.31
Phe151	8.01	119.84	7.97	120.02	8.05	120.41
Asn152	8.74	119.37	8.71	119.33	8.65	119.72
Pro153	-	-	-	-	-	-
Thr154	7.74	111.59	7.71	111.56	-	-
Gln155	7.77	119.26	7.74	119.24	-	-
Leu156	7.52	115.74	7.47	115.70	-	-
Glu157	-	-	-	-	-	-
Glu158	7.17	118.80	7.14	118.77	7.37	119.19
Gln159	-	-	-	-	-	-
Cys160	-	-	-	-	-	-
His161	-	-	-	-	-	-
Ile162	-	-	-	-	-	-

C

Relaxation Parameters for β-Lg

C.1 ^{15}N Relaxation Parameters for β-Lg C at 305 K

Table C.1 lists relaxation parameters and errors for β-Lg C sampled at 305 K.

Table C.1 β-Lg C relaxation parameters sampled at 305K

Residue	NOE	δNOE	R_1 (s^{-1})	δR_1 (s^{-1})	R_2 (s^{-1})	δR_2 (s^{-1})	R_2/R_1	$\delta R_2/R_1$
Met0								
Leu1								
Ile2								
Val3	0.41	0.02	1.26	0.03	9.51	0.23	7.56	0.28
Thr4	0.42	0.02	1.13	0.02	9.32	0.13	8.28	0.19
Gln5	0.43	0.01	1.11	0.02	8.15	0.19	7.33	0.24
Thr6								
Met7	0.69	0.02	1.15	0.02	10.47	0.12	9.11	0.22
Leu8	0.53	0.02	1.13	0.03	10.93	0.12	9.64	0.30
Gly9	0.63	0.02	1.39	0.02	8.58	0.17	6.16	0.17
Lys10	0.55	0.01	1.27	0.01	9.26	0.52	7.27	0.42
Asp11	0.73	0.02	1.10	0.05	11.07	0.28	10.03	0.59
Ile12	0.71	0.02	1.04	0.03	11.00	0.20	10.61	0.39
Gln13	0.76	0.02	1.16	0.02	12.82	0.18	11.06	0.31
Lys14	0.74	0.02	1.13	0.03	10.71	0.37	9.51	0.42
Val15	0.80	0.02	1.09	0.02	11.76	0.32	10.75	0.35
Ala16	0.80	0.01	1.11	0.03	12.67	0.14	11.37	0.38
Gly17	0.80	0.03	1.15	0.02	11.07	0.31	9.65	0.33
Thr18	0.76	0.02	1.08	0.03	11.74	0.26	10.90	0.40
Trp19	0.77	0.04	1.11	0.04	12.64	0.74	11.34	0.79
Tyr20	0.76	0.03	1.16	0.06	10.75	0.27	9.25	0.64
Ser21	0.76	0.06	1.17	0.06	12.14	0.44	10.39	0.74
Leu22	0.73	0.06	1.28	0.08	22.85	1.53	17.91	1.86
Ala23	0.78	0.03	1.15	0.05	12.16	0.25	10.62	0.61
Met24	0.77	0.02	1.17	0.07	11.54	0.31	9.86	0.77
Ala25	0.76	0.03	1.11	0.06	11.51	0.16	10.36	0.63
Ala26	0.74	0.02	1.22	0.07	11.59	0.14	9.52	0.69
Ser27	0.77	0.03	1.24	0.09	11.44	0.33	9.20	0.83
Asp28								
Ile29								
Ser30	0.75	0.02	1.31	0.07	11.51	0.33	8.77	0.69
Leu31	0.78	0.02	1.33	0.07	12.35	0.27	9.29	0.69
Leu32	0.76	0.04	1.30	0.10	12.50	0.20	9.61	0.97
Asp33								

Continued on next page

Table C.1 – continued from previous page

Residue	NOE	δNOE	R_1 (s^{-1})	δR_1 (s^{-1})	R_2 (s^{-1})	δR_2 (s^{-1})	R_2/R_1	$\delta R_2/R_1$
Ala34	0.61	0.02	1.10	0.05	7.59	0.12	6.91	0.37
Gln35								
Ser36	0.69	0.02	1.23	0.07	11.78	0.42	9.55	0.78
Ala37	0.74	0.02	1.31	0.06	11.62	0.32	8.88	0.58
Pro38								
Leu39	0.78	0.06	1.17	0.06	21.32	2.61	18.19	2.45
Arg40	0.77	0.02	1.27	0.07	10.73	0.79	8.42	0.85
Val41	0.70	0.03	1.21	0.09	12.62	0.40	10.40	0.97
Tyr42								
Val43	0.76	0.07	1.20	0.08	15.19	0.86	12.63	1.20
Glu44								
Glu45	0.76	0.02	1.12	0.03	11.25	0.30	10.02	0.43
Leu46	0.74	0.03	1.12	0.03	11.93	0.33	10.64	0.45
Lys47	0.74	0.03	1.09	0.03	11.26	0.17	10.29	0.37
Pro48								
Thr49	0.74	0.03	1.10	0.03	12.03	0.41	10.98	0.52
Pro50								
Glu51	0.60	0.01	1.02	0.04	8.96	0.45	8.80	0.54
Gly52	0.72	0.02	1.13	0.04	12.22	0.16	10.78	0.47
Asp53	0.70	0.02	1.13	0.04	10.96	0.28	9.68	0.48
Leu54								
Glu55	0.70	0.03	1.14	0.03	11.16	0.31	9.78	0.42
Ile56	0.77	0.03	1.15	0.04	11.83	0.18	10.25	0.46
Leu57	0.75	0.03	1.16	0.03	13.78	0.43	11.86	0.51
Leu58	0.74	0.02	1.17	0.04	11.94	0.42	10.18	0.53
His59	0.76	0.08	1.15	0.04	22.97	2.28	20.04	2.17
Lys60								
Trp61								
Glu62	0.68	0.07	1.16	0.06	21.95	3.73	18.99	3.40
Asn63	0.58	0.02	1.05	0.02	11.68	0.47	11.16	0.50
Gly64								
Glu65	0.67	0.07	1.10	0.02	14.05	1.67	12.72	1.53
Cys66								
Ala67	0.71	0.07	1.07	0.04	22.35	1.86	20.88	1.97
Gln68								
Lys69								
Lys70	0.72	0.04	1.07	0.02	13.03	0.34	12.22	0.42
Ile71	0.73	0.03	1.11	0.03	15.28	0.40	13.81	0.52

Continued on next page

Table C.1 – continued from previous page

Residue	NOE	δNOE	R_1 (s^{-1})	δR_1 (s^{-1})	R_2 (s^{-1})	δR_2 (s^{-1})	R_2/R_1	δR_2/R_1
Ile72	0.75	0.04	1.08	0.03	14.72	0.31	13.60	0.45
Ala73	0.68	0.05	1.13	0.04	17.07	1.05	15.13	1.09
Glu74	0.77	0.04	1.09	0.03	11.93	0.36	10.98	0.46
Lys75	0.74	0.02	1.12	0.03	11.50	0.10	10.26	0.35
Thr76	0.77	0.02	1.14	0.04	10.90	0.22	9.57	0.46
Lys77	0.70	0.02	1.09	0.04	10.03	0.23	9.22	0.43
Ile78	0.74	0.02	1.12	0.04	11.54	0.28	10.28	0.49
Pro79								
Ala80	0.75	0.03	1.01	0.02	11.01	0.20	10.89	0.30
Val81	0.78	0.02	1.07	0.04	11.97	0.23	11.16	0.46
Phe82	0.79	0.03	1.08	0.04	12.08	0.35	11.17	0.53
Lys83	0.76	0.03	1.09	0.03	11.89	0.50	10.92	0.59
Ile84								
Asp85								
Ala86	0.74	0.04	1.04	0.02	11.96	0.47	11.44	0.51
Leu87								
Asn88	0.72	0.05	1.14	0.02	22.24	1.34	19.51	1.23
Glu89	0.75	0.06	1.13	0.05	13.18	0.60	11.71	0.75
Asn90	0.79	0.06	1.13	0.05	14.82	0.82	13.13	0.98
Lys91								
Val92								
Leu93	0.74	0.02	1.08	0.04	11.30	0.22	10.49	0.44
Val94	0.79	0.02	1.03	0.03	11.11	0.14	10.82	0.33
Leu95	0.74	0.04	1.16	0.03	11.68	0.29	10.09	0.42
Asp96	0.71	0.02	1.08	0.03	10.88	0.25	10.10	0.40
Thr97								
Asp98	0.81	0.03	1.15	0.04	12.00	0.28	10.45	0.49
Tyr99	0.79	0.04	1.20	0.05	13.54	0.63	11.33	0.76
Lys100	0.72	0.02	1.05	0.02	13.95	0.23	13.28	0.38
Lys101	0.76	0.03	1.11	0.03	12.38	0.30	11.12	0.46
Tyr102	0.80	0.02	1.15	0.04	12.20	0.19	10.59	0.46
Leu103	0.79	0.02	1.13	0.04	12.04	0.26	10.65	0.47
Leu104	0.81	0.03	1.14	0.04	11.76	0.31	10.30	0.48
Phe105								
Cys106	0.80	0.03	1.07	0.03	11.79	0.24	11.00	0.41
Met107	0.77	0.04	1.07	0.04	11.42	0.50	10.69	0.62
Glu108								
Asn109	0.79	0.1	1.08	0.06	18.78	1.89	17.35	2.05

Continued on next page

Table C.1 – continued from previous page

Residue	NOE	δNOE	R_1 (s^{-1})	δR_1 (s^{-1})	R_2 (s^{-1})	δR_2 (s^{-1})	R_2/R_1	$\delta R_2/R_1$
Ser110	0.75	0.02	1.11	0.05	11.51	0.30	10.37	0.61
Ala111								
Glu112	0.71	0.02	1.04	0.03	11.51	0.15	11.11	0.37
Pro113								
Glu114								
Gln115								
Ser116	0.73	0.02	1.22	0.05	11.66	0.34	9.54	0.59
Leu117	0.78	0.03	1.25	0.07	11.56	0.23	9.25	0.68
Ala118	0.79	0.03	1.19	0.07	12.73	0.33	10.66	0.76
Cys119	0.77	0.04	1.10	0.06	12.29	0.32	11.22	0.77
Gln120	0.79	0.06	1.13	0.05	13.12	0.36	11.63	0.63
Cys121								
Leu122								
Val123	0.79	0.04	1.18	0.04	12.27	0.31	10.40	0.51
Arg124								
Thr125	0.74	0.02	1.11	0.04	11.34	0.23	10.26	0.48
Pro126								
Glu127	0.64	0.03	1.06	0.02	16.81	0.41	15.93	0.54
Val128								
Asp129	0.65	0.08	1.14	0.05	14.80	1.61	12.97	1.55
Asp130								
Glu131	0.73	0.02	1.12	0.02	11.98	0.11	10.72	0.28
Ala132	0.78	0.03	1.17	0.04	12.28	0.43	10.49	0.52
Leu133	0.75	0.03	1.15	0.04	12.40	0.55	10.76	0.65
Glu134								
Lys135								
Phe136								
Asp137	0.76	0.03	1.24	0.06	11.26	0.24	9.09	0.56
Lys138								
Ala139	0.75	0.02	1.19	0.04	11.01	0.17	9.27	0.41
Leu140	0.70	0.02	1.13	0.03	10.70	0.33	9.49	0.40
Lys141	0.72	0.02	1.18	0.04	11.57	0.26	9.77	0.43
Ala142	0.60	0.02	1.10	0.04	9.73	0.19	8.83	0.39
Leu143	0.67	0.01	1.07	0.03	11.00	0.14	10.27	0.31
Pro144								
Met145	0.71	0.02	1.06	0.04	9.98	0.30	9.41	0.45
His146	0.71	0.03	1.15	0.07	11.10	0.43	9.68	0.81
Ile147	0.78	0.02	1.19	0.06	14.12	0.41	11.85	0.75

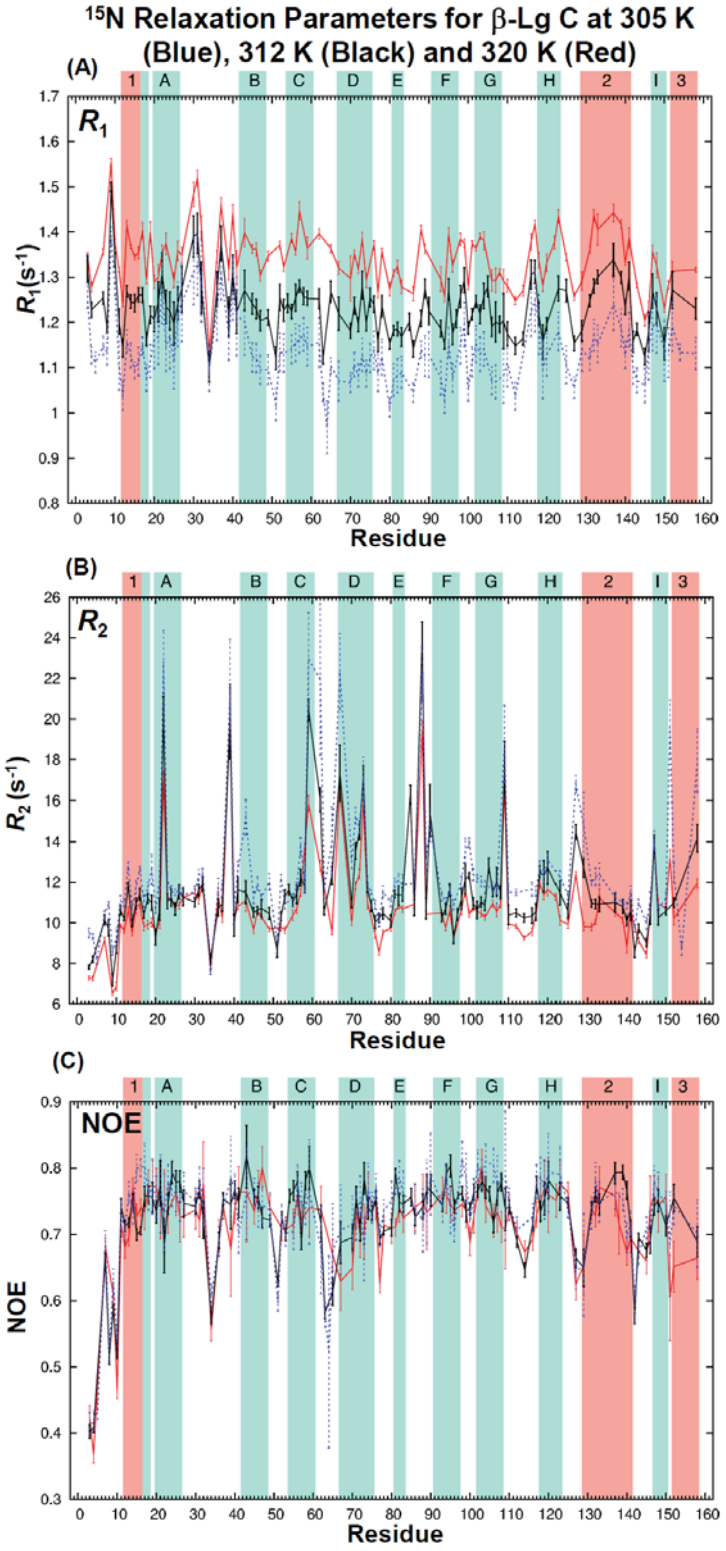
Continued on next page

Table C.1 – continued from previous page

Residue	NOE	δNOE	R_1 (s^{-1})	δR_1 (s^{-1})	R_2 (s^{-1})	δR_2 (s^{-1})	R_2/R_1	δR_2/R_1
Arg148	0.77	0.03	1.20	0.06	10.68	0.18	8.93	0.60
Leu149								
Ser150	0.74	0.02	1.14	0.07	11.22	0.30	9.86	0.73
Phe151	0.69	0.08	1.13	0.06	19.62	1.32	17.32	1.51
Asn152	0.74	0.03	1.19	0.06	12.46	0.46	10.46	0.70
Pro153								
Thr154								
Gln155								
Leu156								
Glu157								
Glu158	0.70	0.05	1.13	0.04	17.96	1.55	15.88	1.48
Gln159								
Cys160								
His161								
Ile162								

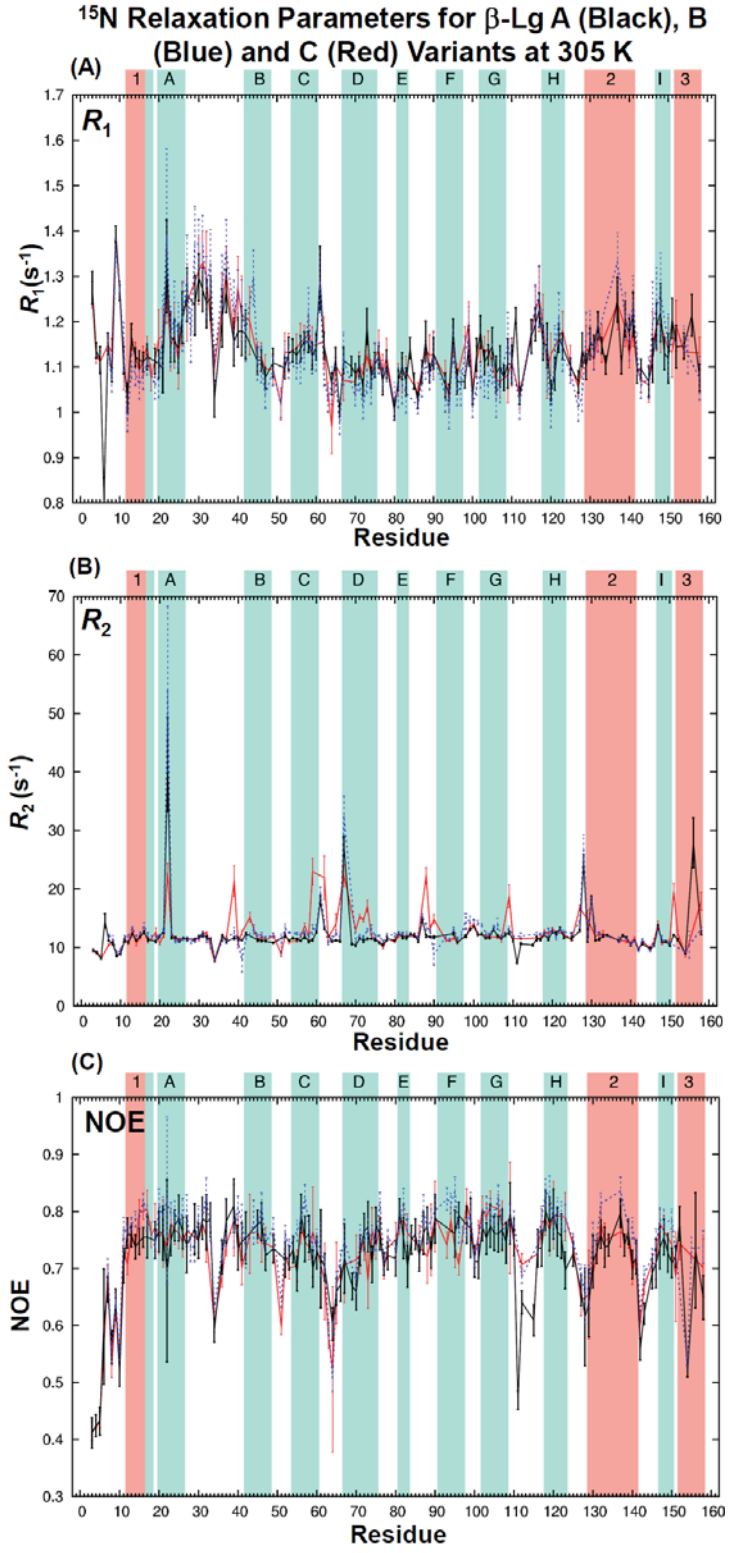
C.2 ^{15}N Relaxation Parameters for β-Lg C at 305 K, 313K and 320 K

Figure C.2 is an overlay traces for ^{15}N (A) R_1, (B) R_2 and (C) NOE enhancement values and errors for β-Lg C sampled at 305 K (blue), 313 K (black) and 320K (red).



C.3 ^{15}N Relaxation Parameters for β-Lg A, B and C Variants at 305 K

Figure C.3 is an overlay traces for ^{15}N (A) R_1, (B) R_2 and (C) NOE enhancement values and errors for β-Lg Variants A (black), B (blue) and C (red).



D

Model-Free Parameters

D.1 β-Lg C Model-Free at 305 K

Table D.1 lists the β-Lg C model-free values and respective errors derived at 305 K.

Table D.4 β-Lg C model-free parameters at 305 K.

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Met0											
Leu1											
Ile2											
Val3	5	0.65	0.02	0.9	0.02	0.72	0.02	656.00	45.90		
Thr4	5	0.64	0.01	0.83	0.01	0.77	0.01	531.00	31.80		
Gln5	5	0.55	0.01	0.78	0.01	0.71	0.02	718.00	43.00		
Thr6											
Met7	5	0.73	0.01	0.83	0.01	0.88	0.01	827.00	104.00		
Lys8	2	0.78	0.01			0.78	0.01	69.40	6.07		
Gly9	5	0.55	0.01	0.84	0.01	0.66	0.01	1360.00	98.00		
Leu10	5	0.62	0.04	0.86	0.02	0.73	0.03	910.00	78.10		
Asp11	2	0.79	0.02			0.79	0.02	24.80	5.74		
Ile12	2	0.78	0.01			0.78	0.01	26.20	4.12		
Gln13	4	0.86	0.02			0.86	0.02	28.10	6.95	0.72	0.31
Lys14	5	0.75	0.03	0.83	0.02	0.91	0.02	865.00	235.00		
Val15	1	0.83	0.01			0.83	0.01				
Ala16	1	0.90	0.01			0.90	0.01				
Gly17	1	0.85	0.01			0.85	0.01				
Thr18	2	0.82	0.02			0.82	0.02	20.70	6.01		
Trp19	1	0.86	0.03			0.86	0.03				
Tyr20	1	0.78	0.02			0.78	0.02				
Ser21	1	0.87	0.03			0.87	0.03				
Leu22	3	0.97	0.04			0.97	0.04			9.29	1.57
Ala23	2	0.87	0.02			0.87	0.02	22.70	9.46		
Met24	2	0.83	0.02			0.83	0.02	21.00	6.97		
Ala25	2	0.82	0.01			0.82	0.01	20.80	7.21		
Ala26	2	0.83	0.01			0.83	0.01	27.50	7.13		
Ser27	1	0.83	0.02			0.83	0.02				
Asp28											
Ile29											
Ser30	5	0.80	0.03	0.91	0.03	0.88	0.03	1310.00	731.00		
Leu31	1	0.90	0.02			0.90	0.02				
Leu32	1	0.90	0.01			0.90	0.01				
Asp33											

Continued on next page

Table D.1 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Ala34	5	0.50	0.01	0.71	0.02	0.71	0.03	1140.00	94.40		
Gln35											
Ser36	2	0.85	0.02			0.85	0.02	52.90	140.00		
Ala37	5	0.81	0.03	0.92	0.02	0.88	0.03	1190.00	272.00		
Pro38											
Leu39	3	0.89	0.04			0.89	0.04			8.87	2.70
Arg40	1	0.87	0.04			0.87	0.04				
Val41	2	0.90	0.03			0.90	0.03	75.10	263.00		
Tyr42											
Val43	1	0.99	0.03			0.99	0.03				
Glu44											
Glu45	2	0.82	0.02			0.82	0.02	20.10	5.39		
Leu46	2	0.84	0.02			0.84	0.02	28.40	10.50		
Lys47	2	0.81	0.01			0.81	0.01	23.10	6.02		
Pro48											
Thr49	2	0.83	0.02			0.83	0.02	28.10	8.33		
Pro50											
Glu51	5	0.62	0.03	0.74	0.02	0.84	0.03	674.00	129.00		
Gly52	2	0.87	0.01			0.87	0.01	43.40	9.42		
Asp53	2	0.80	0.02			0.8	0.02	31.70	5.67		
Leu54	2	0.81	0.02			0.81	0.02	36.00	7.83		
Glu55											
Ile56	2	0.85	0.01			0.85	0.01	22.20	9.66		
Leu57	1	0.92	0.02			0.92	0.02				
Leu58	2	0.86	0.02			0.86	0.02	37.10	10.60		
His59	3	0.87	0.03			0.87	0.03			10.79	2.35
Lys60											
Trp61											
Glu62	1	0.90	0.04			0.90	0.04				
Asn63	4	0.74	0.01			0.74	0.01	45.40	4.60	1.21	0.49
Gly64											
Glu65	1	0.84	0.01			0.84	0.01				
Cys66											
Ala67	4	0.79	0.04			0.79	0.04	28.30	13.40	11.32	1.91
Gln68											
Lys69											
Lys70	4	0.79	0.02			0.79	0.02	26.20	8.38	2.02	0.42
Ile71	4	0.82	0.02			0.82	0.02	29.40	8.90	3.82	0.49

Continued on next page

Table D.1 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Ile72	4	0.80	0.02			0.80	0.02	21.00	7.70	3.46	0.42
Ala73	4	0.82	0.03			0.82	0.03	44.30	15.20	5.52	1.13
Glu74	1	0.84	0.02			0.84	0.02				
Lys75	2	0.82	0.01			0.82	0.01	26.90	6.42		
Thr76	1	0.80	0.01			0.80	0.01				
Lys77	5	0.70	0.02	0.79	0.02	0.89	0.02	816.00	168.00		
Ile78	2	0.83	0.02			0.83	0.02	28.40	5.66		
Pro79											
Ala80	1	0.78	0.01			0.78	0.01				
Val81	1	0.84	0.02			0.84	0.02				
Phe82	1	0.85	0.02			0.85	0.02				
Lys83	2	0.82	0.02			0.82	0.02	20.20	8.58		
Ile84											
Asp85											
Ala86	1	0.81	0.01			0.81	0.01				
Leu87											
Asn88	4	0.84	0.02			0.84	0.02	36.40	12.30	10.45	1.34
Glu89	1	0.89	0.03			0.89	0.03				
Asn90	3	0.86	0.04			0.86	0.04			2.83	1.00
Lys91											
Val92											
Leu93	2	0.80	0.01			0.80	0.01	23.20	5.74		
Val94	2	0.79	0.01			0.79	0.01	11.20	4.41		
Leu95	2	0.84	0.02			0.84	0.02	33.60	11.70		
Asp96	2	0.78	0.01			0.78	0.01	27.70	4.39		
Thr97											
Asp98	1	0.86	0.02			0.86	0.02				
Tyr99	1	0.93	0.03			0.93	0.03				
Lys100	4	0.77	0.02			0.77	0.02	23.70	3.79	3.09	0.31
Lys101	1	0.87	0.02			0.87	0.02				
Tyr102	1	0.87	0.01			0.87	0.01				
Leu103	1	0.86	0.02			0.86	0.02				
Leu104	1	0.85	0.02			0.85	0.02				
Phe105											
Cys106	1	0.83	0.01			0.83	0.01				
Met107	1	0.81	0.02			0.81	0.02				
Glu108											
Asn109	3	0.82	0.05			0.82	0.05			7.28	1.88

Continued on next page

Table D.1 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Ser110	2	0.82	0.02			0.82	0.02	25.00	7.00		
Ala111											
Glu112	4	0.76	0.02			0.76	0.02	24.40	4.63	0.83	0.35
Pro113											
Glu114											
Gln115											
Ser116	2	0.85	0.02			0.85	0.02	38.40	46.90		
Leu117	1	0.84	0.02			0.84	0.02				
Ala118	1	0.91	0.02			0.91	0.02				
Cys119	1	0.87	0.02			0.87	0.02				
Gln120	1	0.91	0.02			0.91	0.02				
Cys121											
Leu122											
Val123	1	0.88	0.02			0.88	0.02				
Arg124											
Thr125	2	0.81	0.01			0.81	0.01	24.70	4.34		
Pro126											
Glu127	4	0.76	0.02			0.76	0.02	36.90	5.78	6.09	0.51
Val128											
Asp129	1	0.89	0.04			0.89	0.04				
Asp130											
Glu131	2	0.85	0.01			0.85	0.01	36.00	5.75		
Ala132	1	0.88	0.02			0.88	0.02				
Leu133	2	0.87	0.03			0.87	0.03	33.70	47.40		
Glu134											
Lys135											
Phe136											
Asp137	1	0.82	0.02			0.82	0.02				
Lys138											
Ala139	5	0.77	0.02	0.85	0.02	0.91	0.02	1090.00	921.00		
Leu140	5	0.75	0.03	0.83	0.02	0.90	0.02	689.00	167.00		
Lys141	2	0.84	0.02			0.84	0.02	37.80	6.88		
Ala142	5	0.68	0.02	0.80	0.02	0.84	0.02	673.00	103.00		
Leu143	2	0.78	0.01			0.78	0.01	36.00	3.75		
Pro144											
Met145	5	0.70	0.02	0.78	0.02	0.9	0.02	781.00	197.00		
His146	2	0.80	0.03			0.8	0.03	31.50	45.80		
Ile147	1	0.98	0.02			0.98	0.02				

Continued on next page

Table D.1 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Arg148	1	0.77	0.01			0.77	0.01				
Leu149											
Ser150	2	0.81	0.02			0.81	0.02	23.80	6.42		
Phe151	4	0.83	0.04			0.83	0.04	43.50	29.90	8.00	1.53
Asn152	2	0.89	0.03			0.89	0.03	44.20	85.20		
Pro153											
Thr154											
Gln155											
Leu156											
Glu157											
Glu158	4	0.83	0.03			0.83	0.03	39.20	14.40	6.32	1.55
Gln159											
Cys160											
His161											
Ile162											

D.2 β-Lg C Model-Free at 313 K

Table D.2 lists the β-Lg C model-free values and respective errors derived at 313 K.

Table D.2 β-Lg C model-free parameters at 313 K.

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Met0											
Leu1											
Ile2											
Val3	5	0.55	0.01	0.86	0.01	0.64	0.02	786.00	28.40		
Thr4	5	0.59	0.01	0.84	0.01	0.71	0.01	663.00	30.60		
Gln5											
Thr6											
Met7	5	0.75	0.02	0.87	0.01	0.87	0.02	743.00	96.60		
Leu8	2	0.78	0.01			0.78	0.01	78.10	5.11		
Gly9	5	0.46	0.03	0.83	0.02	0.56	0.02	1380.00	46.70		
Lys10	5	0.62	0.02	0.87	0.01	0.71	0.01	867.00	34.40		
Asp11	2	0.82	0.01			0.82	0.01	27.40	3.77		
Ile12	2	0.78	0.01			0.78	0.01	27.20	2.09		
Gln13	2	0.89	0.01			0.89	0.01	60.60	5.76		
Lys14	5	0.74	0.02	0.83	0.01	0.89	0.01	1250.00	124.00		
Val15	2	0.85	0.01			0.85	0.01	49.90	5.23		
Ala16	2	0.87	0.01			0.87	0.01	57.00	5.68		
Gly17	5	0.78	0.02	0.86	0.01	0.90	0.01	1140.00	181.00		
Thr18	2	0.83	0.01			0.83	0.01	22.50	3.97		
Trp19	2	0.85	0.01			0.85	0.01	29.30	7.57		
Tyr20	5	0.68	0.03	0.79	0.02	0.86	0.03	1260.00	234.00		
Ser21	1	0.86	0.03			0.86	0.03				
Leu22	4	0.90	0.03			0.90	0.03	96.40	96.20	8.36	0.91
Ala23	2	0.85	0.02			0.85	0.02	32.80	7.76		
Met24	2	0.86	0.02			0.86	0.02	17.80	7.17		
Ala25	2	0.82	0.02			0.82	0.02	17.20	4.99		
Ala26	2	0.88	0.02			0.88	0.02	25.90	7.77		
Ser27	2	0.87	0.02			0.87	0.02	38.90	69.20		
Asp28											
Ile29											
Ser30	5	0.81	0.02	0.93	0.02	0.88	0.02	1210.00	208.00		
Leu31	5	0.87	0.03	0.95	0.02	0.91	0.02	1190.00	296.00		
Leu32	2	0.90	0.02			0.90	0.02	66.70	201.00		
Asp33											

Continued on next page

Table D.2 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Ala34	5	0.60	0.02	0.75	0.02	0.80	0.03	740.00	97.20		
Gln35											
Ser36	2	0.85	0.02			0.85	0.02	54.80	22.90		
Ala37	5	0.81	0.03	0.92	0.02	0.88	0.02	1290.00	232.00		
Pro38											
Leu39	4	0.85	0.03			0.85	0.03	32.00	14.50	8.71	1.91
Arg40	5	0.72	0.05	0.84	0.03	0.86	0.04	1700.00	374.00		
Val41	2	0.86	0.03			0.86	0.03	35.80	48.60		
Tyr42											
Val43	1	0.90	0.02			0.90	0.02				
Glu44											
Glu45	5	0.79	0.02	0.86	0.01	0.92	0.02	986.00	227.00		
Leu46	5	0.80	0.02	0.86	0.01	0.93	0.01	845.00	199.00		
Lys47	2	0.83	0.01			0.83	0.01	35.00	5.52		
Pro48											
Thr49	5	0.79	0.02	0.85	0.01	0.92	0.02	665.00	144.00		
Pro50											
Glu51	5	0.64	0.03	0.77	0.02	0.83	0.03	777.00	121.00		
Gly52	5	0.77	0.02	0.86	0.01	0.89	0.02	937.00	154.00		
Asp53	2	0.85	0.01			0.85	0.01	49.70	6.40		
Leu54	2	0.88	0.01			0.88	0.01	35.10	5.51		
Glu55	2	0.85	0.01			0.85	0.01	26.80	5.24		
Ile56	2	0.88	0.02			0.88	0.02	26.90	8.01		
Leu57	2	0.89	0.01			0.89	0.01	74.80	9.09		
Leu58	2	0.89	0.01			0.89	0.01	28.70	6.63		
His59	3	0.90	0.01			0.90	0.01			8.69	0.52
Lys60											
Trp61											
Glu62	4	0.87	0.02			0.87	0.02	50.70	12.40	4.58	0.58
Asn63	4	0.75	0.01			0.75	0.01	49.90	3.07	0.63	0.23
Gly64											
Glu65	2	0.87	0.01			0.87	0.01	104.00	23.20		
Cys66											
Ala67	4	0.84	0.03			0.84	0.03	49.70	14.60	6.27	1.39
Gln68											
Lys69											
Lys70	2	0.82	0.01			0.82	0.01	39.90	5.75		
Ile71	4	0.87	0.01			0.87	0.01	33.20	6.74	2.13	0.45

Continued on next page

Table D.2 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Ile72	4	0.83	0.01			0.83	0.01	39.10	5.29	3.56	0.39
Ala73											
Glu74	4	0.84	0.01			0.84	0.01	30.90	5.40	0.71	0.28
Lys75	5	0.82	0.02	0.88	0.01	0.92	0.02	712.00	164.00		
Thr76	5	0.75	0.03	0.83	0.02	0.89	0.02	1220.00	213.00		
Lys77	2	0.79	0.01			0.79	0.01	33.70	2.69		
Ile78	5	0.78	0.02	0.87	0.01	0.90	0.02	706.00	141.00		
Pro79											
Ala80	2	0.79	0.01			0.79	0.01	30.90	3.10		
Val81	2	0.84	0.01			0.84	0.01	15.20	2.05		
Phe82	2	0.83	0.01			0.83	0.01	27.10	4.10		
Lys83	2	0.83	0.01			0.83	0.01	26.90	4.27		
Ile84											
Asp85	4	0.84	0.01			0.84	0.01	26.90	4.56	5.23	0.48
Ala86	2	0.80	0.01			0.80	0.01	27.00	3.34		
Leu87											
Asn88	4	0.84	0.01			0.84	0.01	32.30	6.18	12.79	0.9
Glu89	1	0.90	0.01			0.90	0.01				
Asn90	1	0.89	0.02			0.89	0.02				
Lys91											
Val92											
Leu93	5	0.77	0.02	0.83	0.01	0.93	0.02	768.00	172.00		
Val94	2	0.81	0.01			0.81	0.01	12.70	3.00		
Leu95	2	0.90	0.01			0.90	0.01	20.30	9.27		
Asp96	5	0.68	0.02	0.78	0.02	0.88	0.02	1230.00	166.00		
Thr97	5	0.78	0.03	0.84	0.02	0.93	0.02	885.00	209.00		
Asp98	2	0.88	0.02			0.88	0.02	34.00	6.89		
Tyr99	2	0.90	0.02			0.9	0.02	62.20	24.70		
Lys100	4	0.83	0.01			0.83	0.01	29.70	3.05	1.39	0.19
Lys101	2	0.85	0.01			0.85	0.01	26.90	4.38		
Tyr102	5	0.80	0.04	0.86	0.02	0.93	0.02	1010.00	282.00		
Leu103	2	0.85	0.01			0.85	0.01	21.40	4.54		
Leu104											
Phe105	4	0.90	0.02			0.90	0.02	42.40	11.10	1.00	0.43
Cys106	4	0.83	0.02			0.83	0.02	36.20	4.84	0.59	0.27
Met107	2	0.86	0.03			0.86	0.03	28.80	32.30		
Glu108	2	0.85	0.01			0.85	0.01	20.60	4.65		
Asn109	4	0.84	0.04			0.84	0.04	32.30	21.80	6.57	1.46

Continued on next page

Table D.2 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Ser110	2	0.81	0.02			0.81	0.02	26.10	3.87		
Ala111											
Glu112	2	0.79	0.01			0.79	0.01	35.30	2.77		
Pro113											
Glu114	2	0.79	0.01			0.79	0.01	45.30	3.45		
Gln115											
Ser116	5	0.77	0.03	0.88	0.02	0.87	0.02	1030.00	174.00		
Leu117	5	0.78	0.04	0.87	0.02	0.89	0.03	1480.00	738.00		
Ala118	4	0.85	0.02			0.85	0.02	36.40	7.22	1.32	0.39
Cys119	1	0.85	0.03			0.85	0.03				
Gln120	1	0.86	0.01			0.86	0.01				
Cys121											
Leu122											
Val123	2	0.89	0.02			0.89	0.02	41.40	13.00		
Arg124											
Thr125	5	0.79	0.03	0.87	0.02	0.91	0.02	1010.00	216.00		
Pro126											
Glu127	4	0.79	0.01			0.79	0.01	40.90	4.15	4.06	0.39
Val128											
Asp129	4	0.81	0.02			0.81	0.02	49.50	8.06	2.07	0.50
Asp130											
Glu131	5	0.83	0.01	0.88	0.01	0.93	0.01	642.00	122.00		
Ala132	5	0.82	0.02	0.89	0.01	0.92	0.01	963.00	194.00		
Leu133	5	0.82	0.03	0.90	0.02	0.91	0.02	948.00	487.00		
Glu134											
Lys135											
Phe136											
Asp137	1	0.92	0.02			0.92	0.02				
Lys138											
Ala139	5	0.80	0.03	0.88	0.02	0.92	0.02	1580.00	1130.00		
Leu140	5	0.75	0.02	0.83	0.01	0.90	0.02	1280.00	530.00		
Lys141	5	0.79	0.04	0.89	0.02	0.88	0.03	889.00	200.00		
Ala142	5	0.63	0.02	0.79	0.01	0.80	0.02	764.00	82.90		
Leu143	5	0.73	0.02	0.82	0.01	0.89	0.02	803.00	110.00		
Pro144											
Met145	5	0.67	0.02	0.77	0.01	0.87	0.02	814.00	94.70		
His146	2	0.80	0.03			0.80	0.03	37.20	8.59		
Ile147	4	0.89	0.02			0.89	0.02	47.10	17.40	2.27	0.49

Continued on next page

Table D.2 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Arg148	5	0.77	0.03	0.84	0.02	0.91	0.03	1030.00	550.00		
Leu149											
Ser150	2	0.80	0.01			0.80	0.01	31.20	4.34		
Phe151											
Asn152	2	0.85	0.02			0.85	0.02	30.60	8.39		
Pro153											
Thr154											
Gln155											
Leu156											
Glu157											
Glu158	4	0.85	0.02			0.85	0.02	52.50	10.70	2.98	0.74
Gln159											
Cys160											
His161											
Ile162											

D.3 β-Lg C Model-Free at 320 K

Table D.3 lists the β-Lg C model-free values and respective errors derived at 320 K.

Table D.3 β-Lg C model-free parameters at 320 K.

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Met0											
Leu1											
Ile2											
Val3	5	0.56	0.01	0.84	0.01	0.66	0.01	756.00	29.50		
Thr4	5	0.56	0.01	0.83	0.01	0.68	0.01	628.00	25.70		
Gln5											
Thr6											
Met7	5	0.74	0.01	0.86	0.01	0.85	0.01	885.00	86.80		
Lys8											
Gly9	5	0.45	0.01	0.81	0.01	0.55	0.01	1490.00	74.80		
Leu10	5	0.51	0.02	0.82	0.01	0.62	0.02	914.00	40.70		
Asp11											
Ile12	2	0.79	0.01			0.79	0.01	38.50	3.96		
Gln13	2	0.89	0.01			0.89	0.01	88.50	11.30		
Lys14	5	0.77	0.01	0.86	0.01	0.90	0.01	1360.00	239.00		
Val15	2	0.86	0.01			0.86	0.01	46.30	7.30		
Ala16	2	0.88	0.01			0.88	0.01	46.20	6.02		
Gly17	5	0.79	0.02	0.89	0.01	0.89	0.02	1000.00	939.00		
Thr18	2	0.84	0.01			0.84	0.01	23.80	7.03		
Trp19	1	0.89	0.01			0.89	0.01				
Tyr20	1	0.83	0.01			0.83	0.01				
Ser21	1	0.85	0.02			0.85	0.02				
Leu22	3	0.89	0.01			0.89	0.01			6.20	0.48
Ala23	1	0.93	0.01			0.93	0.01				
Met24											
Ala25	1	0.87	0.01			0.87	0.01				
Ala26	2	0.88	0.01			0.88	0.01	50.00	18.00		
Ser27	1	0.91	0.01			0.91	0.01				
Asp28											
Ile29											
Ser30	1	0.96	0.01			0.96	0.01				
Leu31	5	0.91	0.02	1.00	0.01	0.92	0.02	663.00	167.00		
Leu32	1	0.94	0.01			0.94	0.01				
Asp33											

Continued on next page

Table D.3 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Ala34	5	0.63	0.01	0.76	0.01	0.83	0.01	522.00	54.30		
Gln35											
Ser36	4	0.86	0.01			0.86	0.01	57.30	10.20	0.89	0.30
Ala37											
Pro38											
Leu39	4	0.84	0.02			0.84	0.02	55.20	22.60	10.14	1.03
Arg40											
Val41	1	0.89	0.01			0.89	0.01				
Tyr42											
Val43	1	0.92	0.02			0.92	0.02				
Glu44											
Glu45	5	0.78	0.02	0.87	0.01	0.90	0.01	932.00	178.00		
Leu46	2	0.88	0.01			0.88	0.01	37.10	13.60		
Lys47	1	0.86	0.01			0.86	0.01				
Pro48											
Thr49	5	0.79	0.01	0.87	0.01	0.91	0.01	893.00	157.00		
Pro50											
Glu51											
Gly52	5	0.80	0.02	0.89	0.01	0.90	0.01	768.00	134.00		
Asp53	5	0.79	0.01	0.87	0.01	0.91	0.01	656.00	113.00		
Leu54											
Glu55	2	0.88	0.01			0.88	0.01	67.20	11.40		
Ile56	2	0.88	0.01			0.88	0.01	36.40	13.10		
Leu57											
Leu58											
His59											
Lys60											
Trp61											
Glu62	4	0.90	0.01			0.90	0.01	60.90	19.60	1.94	0.29
Asn63											
Gly64											
Glu65	5	0.78	0.02	0.89	0.01	0.88	0.01	692.00	105.00		
Cys66											
Ala67	4	0.83	0.01			0.83	0.01	70.00	13.40	6.57	0.51
Gln68											
Lys69											
Lys70	2	0.82	0.01			0.82	0.01	60.90	13.70		
Ile71	4	0.87	0.01			0.87	0.01	45.10	9.73	1.33	0.21

Continued on next page

Table D.3 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Ile72	4	0.84	0.01			0.84	0.01	49.10	10.70	2.23	0.26
Ala73	4	0.88	0.01			0.88	0.01	70.80	20.40	5.23	0.49
Glu74	1	0.86	0.01			0.86	0.01				
Lys75											
Thr76	5	0.76	0.02	0.86	0.01	0.88	0.01	1140.00	777.00		
Lys77	5	0.69	0.02	0.82	0.01	0.84	0.01	710.00	69.90		
Ile78	5	0.77	0.01	0.87	0.01	0.89	0.01	868.00	88.00		
Pro79											
Ala80	2	0.81	0.01			0.81	0.01	37.30	6.09		
Val81	4	0.84	0.01			0.84	0.01	43.60	4.13	0.40	0.16
Phe82	4	0.85	0.01			0.85	0.01	37.60	7.57	0.53	0.20
Lys83	4	0.82	0.01			0.82	0.01	34.10	5.85	0.76	0.19
Ile84											
Asp85											
Ala86	4	0.81	0.01			0.81	0.01	27.80	8.90	1.08	0.25
Leu87											
Asn88	4	0.91	0.01			0.91	0.01	58.40	23.30	8.48	0.48
Glu89	2	0.87	0.01			0.87	0.01	58.00	19.30		
Asn90											
Lys91											
Val92											
Leu93	1	0.87	0.01			0.87	0.01				
Val94	2	0.81	0.01			0.81	0.01	23.30	6.73		
Leu95	2	0.89	0.01			0.89	0.01	57.10	15.20		
Asp96	5	0.75	0.04	0.84	0.02	0.89	0.03	1010.00	239.00		
Thr97											
Asp98	2	0.89	0.01			0.89	0.01	51.20	10.30		
Tyr99	4	0.88	0.01			0.88	0.01	52.90	12.20	0.88	0.27
Lys100	4	0.81	0.01			0.81	0.01	43.70	5.16	0.56	0.17
Lys101	2	0.88	0.01			0.88	0.01	54.80	10.60		
Tyr102	2	0.88	0.01			0.88	0.01	35.80	7.84		
Leu103	1	0.91	0.01			0.91	0.01				
Leu104											
Phe105											
Cys106	4	0.83	0.02			0.83	0.02	37.50	8.08	0.93	0.25
Met107	2	0.85	0.01			0.85	0.01	35.00	11.00		
Glu108	4	0.84	0.01			0.84	0.01	44.00	7.39	0.77	0.22
Asn109	4	0.83	0.02			0.83	0.02	40.60	16.90	6.84	0.96

Continued on next page

Table D.3 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Ser110	2	0.82	0.01			0.82	0.01	30.10	6.07		
Ala111											
Glu112	2	0.80	0.01			0.80	0.01	35.00	3.98		
Pro113											
Glu114	5	0.76	0.01	0.84	0.01	0.90	0.01	545.00	76.90		
Gln115											
Ser116	5	0.77	0.02	0.89	0.01	0.87	0.02	825.00	115.00		
Leu117											
Ala118	4	0.86	0.01			0.86	0.01	36.90	10.90	1.49	0.24
Cys119	3	0.85	0.01			0.85	0.01			1.16	0.24
Gln120	3	0.88	0.01			0.88	0.01			1.05	0.38
Cys121											
Leu122	1	0.92	0.01			0.92	0.01				
Val123											
Arg124											
Thr125	5	0.81	0.02	0.87	0.01	0.93	0.01	876.00	221.00		
Pro126											
Glu127	4	0.79	0.01			0.79	0.01	54.00	5.51	2.82	0.189
Val128											
Asp129	2	0.81	0.01			0.81	0.01	55.40	12.90		
Asp130											
Glu131	5	0.80	0.02	0.88	0.01	0.91	0.01	883.00	473.00		
Ala132											
Leu133	1	0.93	0.02			0.93	0.02				
Glu134											
Lys135											
Phe136											
Asp137											
Lys138	1										
Ala139	4	0.83	0.01			0.83	0.01	221.00	28.80	0.00	0.03
Leu140	5	0.71	0.03	0.84	0.02	0.84	0.02	920.00	147.00		
Lys141	2	0.88	0.01			0.88	0.01	73.90	14.70		
Ala142											
Leu143	5	0.74	0.01	0.84	0.01	0.88	0.01	658.00	81.30		
Pro144											
Met145											
His146	1	0.83	0.01			0.83	0.01				
Ile147	4	0.88	0.01			0.88	0.01	37.80	13.50	3.23	0.32

Continued on next page

Table D.3 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Arg148	1	0.89	0.01			0.89	0.01				
Leu149											
Ser150	4	0.80	0.01			0.80	0.01	20.70	7.27	1.09	0.24
Phe151	4	0.80	0.02			0.80	0.02	64.60	14.60	3.51	0.42
Asn152	2	0.83	0.01			0.83	0.01	61.80	12.60		
Pro153											
Thr154											
Gln155											
Leu156											
Glu157											
Glu158	4	0.83	0.01			0.83	0.01	59.00	9.88	1.87	0.23
Gln159											
Cys160											
His161											
Ile162											

D.4 β-Lg A Model-Free at 305 K

Table D.4 lists the β-Lg A model-free values and respective errors derived at 305 K.

Table D.4 β-Lg A model-free parameters at 305 K.

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Met0											
Leu1											
Ile2											
Val3	5	0.63	0.01	0.90	0.02	0.70	0.02	700.00	49.80		
Thr4	5	0.60	0.01	0.82	0.01	0.74	0.01	648.00	35.50		
Gln5	5	0.54	0.01	0.78	0.01	0.70	0.02	746.00	46.00		
Thr6	4	0.59	0.03			0.59	0.03	19.90	7.28	6.12	1.30
Met7	5	0.77	0.01	0.87	0.01	0.89	0.01	657.00	90.30		
Lys8											
Gly9	5	0.54	0.01	0.83	0.01	0.65	0.01	1450.00	103.00		
Leu10	5	0.58	0.01	0.84	0.01	0.69	0.01	926.00	49.20		
Asp11	2	0.80	0.01			0.80	0.01	33.30	7.07		
Ile12	2	0.77	0.01			0.77	0.01	15.00	4.87		
Gln13	2	0.87	0.01			0.87	0.01	31.80	7.01		
Lys14	5	0.78	0.01	0.84	0.01	0.93	0.01	746.00	165.00		
Val15	2	0.84	0.01			0.84	0.01	27.90	6.62		
Ala16	4	0.84	0.02			0.84	0.02	25.70	5.71	0.70	0.33
Gly17	1	0.83	0.01			0.83	0.01				
Thr18											
Trp19	1	0.80	0.01			0.8	0.01				
Tyr20	1	0.85	0.01			0.85	0.01				
Ser21	1	0.88	0.02			0.88	0.02				
Leu22	3	1.00	0.04			1.00	0.04			27.06	8.08
Ala23	2	0.82	0.01			0.82	0.01	24.40	6.68		
Met24	1	0.83	0.01			0.83	0.01				
Ala25	1	0.78	0.01			0.78	0.01				
Ala26	1	0.82	0.01			0.82	0.01				
Ser27	1	0.83	0.02			0.83	0.02				
Asp28											
Ile29	1	0.82	0.02			0.82	0.02				
Ser30											
Leu31	1	0.90	0.02			0.90	0.02				
Leu32	1	0.86	0.02			0.86	0.02				
Asp33	1	0.83	0.02			0.83	0.02				

Continued on next page

Table D.4 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Ala34	5	0.53	0.02	0.70	0.02	0.75	0.02	981.00	90.00		
Gln35											
Ser36	2	0.85	0.02			0.85	0.02	40.80	37.30		
Ala37											
Pro38											
Leu39	1	0.83	0.01			0.83	0.01				
Arg40	1	0.84	0.02			0.84	0.02				
Val41	1	0.80	0.01			0.80	0.01				
Tyr42	1	0.87	0.02			0.87	0.02				
Val43											
Glu44											
Glu45	1	0.81	0.01			0.81	0.01				
Leu46	1	0.81	0.01			0.81	0.01				
Lys47	2	0.78	0.01			0.78	0.01	25.60	5.05		
Pro48											
Thr49	5	0.75	0.01	0.82	0.01	0.91	0.02	894.00	215.00		
Pro50											
Glu51											
Gly52	2	0.83	0.01			0.83	0.01	33.40	8.44		
Asp53	5	0.78	0.01	0.85	0.01	0.92	0.02	704.00	428.00		
Leu54	2	0.82	0.01			0.82	0.01	30.30	6.37		
Glu55	5	0.80	0.01	0.86	0.01	0.92	0.01	511.00	105.00		
Ile56											
Leu57	2	0.87	0.02			0.87	0.02	36.10	14.30		
Leu58											
Gln59	2	0.80	0.01			0.80	0.01	36.00	7.76		
Lys60	2	0.88	0.01			0.88	0.01	45.40	13.90		
Trp61	3	1.00	0.03			1.00	0.03			3.82	0.94
Glu62	4	0.86	0.02			0.86	0.02	47.30	13.10	1.12	0.37
Asn63											
Asp64	2	0.77	0.01			0.77	0.01	47.40	5.62		
Glu65	2	0.79	0.01			0.79	0.01	42.40	5.54		
Cys66	4	0.74	0.01			0.74	0.01	27.60	4.48	0.48	0.19
Ala67	4	0.83	0.03			0.83	0.03	35.40	15.30	15.41	1.51
Gln68											
Lys69	5	0.73	0.01	0.83	0.01	0.89	0.01	674.00	112.00		
Lys70	5	0.71	0.01	0.81	0.01	0.88	0.02	683.00	129.00		
Ile71	2	0.82	0.01			0.82	0.01	28.60	6.39		

Continued on next page

Table D.4 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Ile72	2	0.79	0.01			0.79	0.01	17.50	6.44		
Ala73	1	0.82	0.01			0.82	0.01				
Glu74	2	0.80	0.00			0.80	0.00	31.10	7.57		
Lys75	1	0.81	0.01			0.81	0.01				
Thr76	5	0.75	0.01	0.82	0.01	0.91	0.02	1100.00	634.00		
Lys77	5	0.72	0.01	0.80	0.01	0.91	0.02	741.00	156.00		
Ile78	2	0.82	0.02			0.82	0.02	30.50	6.82		
Pro79											
Ala80	1	0.79	0.01			0.79	0.01				
Val81	1	0.86	0.01			0.86	0.01				
Phe82	2	0.83	0.01			0.83	0.01	24.30	9.31		
Lys83	2	0.81	0.02			0.81	0.02	31.70	9.03		
Ile84	2	0.86	0.01			0.86	0.01	31.80	8.71		
Asp85	1	0.83	0.01			0.83	0.01				
Ala86	2	0.77	0.01			0.77	0.01	22.40	5.07		
Leu87											
Asn88	2	0.85	0.02			0.85	0.02	22.40	10.20		
Glu89	2	0.83	0.01			0.83	0.01	29.10	7.13		
Asn90	1	0.84	0.02			0.84	0.02				
Lys91											
Val92											
Leu93	1	0.82	0.02			0.82	0.02				
Val94	1	0.81	0.01			0.81	0.01				
Leu95	2	0.88	0.01			0.88	0.01	36.70	16.60		
Asp96	1	0.77	0.01			0.77	0.01				
Thr97											
Asp98	1	0.83	0.01			0.83	0.01				
Tyr99	1	0.92	0.01			0.92	0.01				
Lys100	4	0.78	0.02			0.78	0.02	24.90	4.90	2.51	0.26
Lys101	2	0.85	0.01			0.85	0.01	41.70	10.60		
Tyr102	2	0.87	0.01			0.87	0.01	26.50	9.63		
Leu103	2	0.82	0.01			0.82	0.01	24.00	6.76		
Leu104	2	0.85	0.01			0.85	0.01	23.20	10.80		
Phe105	2	0.90	0.02			0.90	0.02	38.90	19.60		
Cys106	2	0.84	0.01			0.84	0.01	23.20	9.29		
Met107	1	0.82	0.01			0.82	0.01				
Glu108	1	0.85	0.01			0.85	0.01				
Asn109	1	0.88	0.02			0.88	0.02				

Continued on next page

Table D.4 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Ser110	1	0.80	0.01			0.8	0.01				
Ala111	5	0.47	0.01	0.77	0.01	0.61	0.01	981.00	32.30		
Glu112	2	0.75	0.01			0.75	0.01	35.10	3.91		
Pro113											
Glu114											
Gln115	5	0.71	0.01	0.85	0.02	0.84	0.02	746.00	110.00		
Ser116	1	0.87	0.02			0.87	0.02				
Leu117											
Val118	1	0.87	0.01			0.87	0.01				
Cys119	1	0.81	0.01			0.81	0.01				
Gln120	3	0.81	0.02			0.81	0.02			1.17	0.31
Cys121	1	0.87	0.01			0.87	0.01				
Leu122	1	0.92	0.02			0.92	0.02				
Val123	2	0.84	0.01			0.84	0.01	44.20	9.67		
Arg124											
Thr125	2	0.82	0.01			0.82	0.01	30.50	5.64		
Pro126											
Glu127	4	0.79	0.02			0.79	0.02	38.20	6.67	1.59	0.31
Val128	4	0.83	0.03			0.83	0.03	66.40	28.30	12.55	1.52
Asp129	2	0.80	0.02			0.80	0.02	49.40	13.70		
Asp130	4	0.86	0.03			0.86	0.03	45.90	20.10	6.26	0.47
Glu131	5	0.78	0.01	0.85	0.01	0.91	0.01	836.00	327.00		
Ala132											
Leu133	2	0.85	0.01			0.85	0.01	38.10	10.60		
Glu134	2	0.84	0.01			0.84	0.01	29.50	5.87		
Lys135											
Phe136											
Asp137											
Lys138	2	0.84	0.01			0.84	0.01	29.40	5.88		
Ala139											
Leu140	5	0.72	0.02	0.83	0.02	0.86	0.02	1070.00	192.00		
Lys141											
Ala142	5	0.66	0.01	0.81	0.01	0.82	0.02	648.00	67.30		
Leu143	5	0.74	0.01	0.83	0.01	0.88	0.01	532.00	61.40		
Pro144											
Met145	5	0.67	0.02	0.77	0.02	0.86	0.02	943.00	239.00		
His146	5	0.71	0.02	0.81	0.02	0.88	0.03	952.00	675.00		
Ile147	1	0.95	0.02			0.95	0.02				

Continued on next page

Table D.4 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Arg148	1	0.78	0.01			0.78	0.01				
Leu149											
Ser150	5	0.72	0.02	0.81	0.02	0.88	0.03	946.00	237.00		
Phe151	2	0.85	0.01			0.85	0.01	41.50	7.60		
Asn152	1	0.81	0.01			0.81	0.01				
Pro153											
Thr154	5	0.60	0.01	0.80	0.01	0.75	0.01	806.00	40.20		
Gln155											
Leu156	3	0.94	0.04			0.94	0.04			14.57	4.41
Glu157											
Glu158	4	0.79	0.02			0.79	0.02	42.20	10.40	1.20	0.42
Gln159											
Cys160											
His161											
Ile162											

D.5 β-Lg B Model-Free at 305 K

Table D.5 lists the β-Lg B model-free values and respective errors derived at 305 K.

Table D.5 β-Lg B model-free parameters at 305 K.

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Met0											
Leu1											
Ile2											
Val3											
Thr4											
Gln5											
Thr6											
Met7	2	0.84	0.01			0.84	0.01	42.47	6.31		
Lys8	2	0.78	0.02			0.78	0.02	62.95	8.76		
Gly9	5	0.58	0.01	0.84	0.01	0.68	0.01	1392.80	55.44		
Leu10	5	0.62	0.03	0.86	0.02	0.72	0.02	967.35	56.51		
Asp11	2	0.83	0.02			0.83	0.02	19.05	5.35		
Ile12											
Gln13	4	0.85	0.02			0.85	0.02	20.34	5.15	1.00	0.35
Lys14	2	0.83	0.01			0.83	0.01	15.12	3.56		
Val15	2	0.85	0.01			0.85	0.01	16.22	4.33		
Ala16	1	0.88	0.02			0.88	0.02				
Gly17	2	0.85	0.02			0.85	0.02	17.89	7.18		
Thr18											
Trp19	2	0.81	0.01			0.81	0.01	25.41	5.23		
Tyr20	1	0.83	0.02			0.83	0.02				
Ser21	1	0.89	0.02			0.89	0.02				
Leu22	3	1.00	0.08			1.00	0.08			39.75	14.09
Ala23	1	0.89	0.03			0.89	0.03				
Met24	1	0.76	0.01			0.76	0.01				
Ala25	1	0.78	0.01			0.78	0.01				
Ala26	1	0.79	0.01			0.79	0.01				
Ser27	1	0.81	0.02			0.81	0.02				
Asp28	5	0.74	0.03	0.84	0.02	0.88	0.03	1169.60	769.63		
Ile29	5	0.74	0.04	0.91	0.03	0.81	0.04	1633.70	957.27		
Ser30	5	0.78	0.04	0.92	0.02	0.85	0.05	2319.20	1745.40		
Leu31											
Leu32	1	0.83	0.03			0.83	0.03				
Asp33	5	0.69	0.07	0.86	0.04	0.80	0.05	2068.20	530.17		

Continued on next page

Table D.5 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Ala34	5	0.51	0.02	0.71	0.02	0.71	0.02	1239.80	80.16		
Gln35											
Ser36	5	0.75	0.02	0.89	0.02	0.84	0.03	1365.00	264.24		
Ala37	5	0.74	0.02	0.90	0.02	0.82	0.03	1801.00	354.86		
Pro38											
Leu39	2	0.90	0.03			0.90	0.03	40.67	235.72		
Arg40											
Val41	1	0.91	0.05			0.91	0.05				
Tyr42	2	0.82	0.02			0.82	0.02	29.05	46.94		
Val43											
Glu44	1	0.97	0.03			0.97	0.03				
Glu45	2	0.84	0.01			0.84	0.01	21.33	4.84		
Leu46	1	0.85	0.02			0.85	0.02				
Lys47	2	0.80	0.01			0.80	0.01	22.07	3.78		
Pro48											
Thr49	2	0.84	0.01			0.84	0.01	19.79	4.96		
Pro50											
Glu51	2	0.76	0.02			0.76	0.02	35.65	4.91		
Gly52	2	0.88	0.02			0.88	0.02	31.64	8.79		
Asp53	2	0.87	0.01			0.87	0.01	38.51	6.45		
Leu54											
Glu55	2	0.85	0.01			0.85	0.01	23.14	5.73		
Ile56	2	0.84	0.02			0.84	0.02	16.64	6.77		
Leu57	1	0.88	0.01			0.88	0.01				
Leu58											
Gln59	2	0.87	0.03			0.87	0.03	37.28	11.75		
Lys60	2	0.91	0.02			0.91	0.02	48.39	37.48		
Trp61	4	0.94	0.04			0.94	0.04	596.24	412.10	5.57	1.05
Glu62	4	0.82	0.02			0.82	0.02	34.54	6.08	2.35	0.58
Asn63	2	0.74	0.01			0.74	0.01	40.73	3.24		
Gly64	2	0.78	0.02			0.78	0.02	73.30	13.45		
Glu65	2	0.83	0.02			0.825	0.018	42.50	7.12		
Cys66	4	0.74	0.02			0.742	0.018	19.63	2.72	1.00	0.32
Ala67	4	0.86	0.05			0.858	0.045	31.59	31.07	20.03	3.50
Gln68											
Lys69											
Lys70	2	0.78	0.01			0.78	0.014	32.25	5.26		
Ile71	2	0.83	0.02			0.83	0.02	21.98	5.47		

Continued on next page

Table D.5 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Ile72	2	0.80	0.01			0.80	0.01	19.71	4.08		
Ala73	1	0.87	0.01			0.87	0.01				
Glu74	2	0.82	0.03			0.82	0.03	22.45	6.89		
Lys75	2	0.83	0.02			0.83	0.02	25.92	5.61		
Thr76	1	0.83	0.01			0.83	0.01				
Lys77	5	0.75	0.02	0.82	0.01	0.91	0.02	763.47	146.78		
Ile78	2	0.84	0.01			0.84	0.01	24.75	3.77		
Pro79											
Ala80	2	0.82	0.02			0.82	0.02	27.74	5.81		
Val81	2	0.85	0.01			0.85	0.01	18.17	3.85		
Phe82	1	0.85	0.02			0.85	0.02				
Lys83	2	0.84	0.02			0.84	0.02	20.65	7.12		
Ile84											
Asp85											
Ala86	2	0.79	0.01			0.79	0.01	19.41	2.79		
Leu87											
Asn88	1	0.89	0.02			0.89	0.02				
Glu89	1	0.88	0.01			0.88	0.01				
Asn90	1	0.86	0.04			0.86	0.04				
Lys91											
Val92											
Leu93	1	0.82	0.01			0.82	0.01				
Val94	1	0.79	0.01			0.79	0.01				
Leu95	1	0.91	0.02			0.91	0.02				
Asp96	2	0.80	0.01			0.80	0.01	17.61	4.07		
Thr97	2	0.82	0.02			0.82	0.02	20.92	4.92		
Asp98	2	0.88	0.03			0.88	0.03	28.80	14.24		
Tyr99	1	0.98	0.02			0.98	0.02				
Lys100											
Lys101	4	0.81	0.02			0.81	0.02	20.93	5.26	2.47	0.41
Tyr102	4	0.84	0.02			0.84	0.02	15.91	4.74	1.68	0.42
Leu103	1	0.85	0.01			0.85	0.01				
Leu104	1	0.84	0.01			0.84	0.01				
Phe105	1	0.91	0.02			0.91	0.02				
Cys106	1	0.84	0.03			0.84	0.03				
Met107	2	0.83	0.01			0.83	0.01	19.37	6.23		
Glu108											
Asn109	1	0.92	0.02			0.92	0.02				

Continued on next page

Table D.5 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Ser110	2	0.82	0.01			0.82	0.01	16.23	5.24		
Ala111											
Glu112	2	0.79	0.02			0.79	0.02	32.21	5.20		
Pro113											
Glu114											
Gln115											
Ser116	5	0.84	0.02	0.92	0.02	0.91	0.02	981.82	244.88		
Leu117	1	0.86	0.02			0.86	0.02				
Ala118	1	0.87	0.02			0.87	0.017				
Cys119	1	0.87	0.04			0.87	0.04				
Gln120	1	0.84	0.02			0.84	0.02				
Cys121	1	0.83	0.02			0.83	0.02				
Leu122	1	0.91	0.02			0.91	0.02				
Val123	1	0.88	0.02			0.88	0.02				
Arg124											
Thr125	2	0.85	0.01			0.85	0.01	27.77	4.63		
Pro126											
Glu127	4	0.76	0.02			0.76	0.02	27.18	4.59	3.30	0.38
Val128	4	0.78	0.03			0.78	0.03	37.41	11.27	15.67	2.34
Asp129	2	0.87	0.03			0.87	0.03	68.33	67.28		
Asp130											
Glu131	2	0.88	0.02			0.88	0.02	36.85	9.55		
Ala132	1	0.87	0.01			0.87	0.01				
Leu133											
Glu134											
Lys135											
Phe136											
Asp137	1	0.86	0.03			0.86	0.03				
Lys138											
Ala139	5	0.77	0.02	0.87	0.02	0.89	0.02	1486.50	787.21		
Leu140	5	0.71	0.02	0.81	0.02	0.87	0.03	1664.10	922.33		
Lys141	5	0.75	0.01	0.88	0.02	0.86	0.02	1324.00	191.19		
Ala142	5	0.63	0.01	0.80	0.02	0.79	0.02	953.63	93.34		
Leu143	2	0.79	0.01			0.79	0.01	31.39	3.41		
Pro144											
Met145	5	0.63	0.01	0.75	0.02	0.84	0.02	1282.90	165.90		
His146											
Ile147	1	0.91	0.02			0.91	0.02				

Continued on next page

Table D.5 – continued from previous page

Residue	Model	S^2	δS^2	S^2_f	δS^2_f	S^2_s	δS^2_s	τ_e (ps)	$\delta\tau_e$ (ps)	R_{ex} (s^{-1})	δR_{ex} (s^{-1})
Arg148	1	0.83	0.03			0.83	0.03				
Leu149											
Ser150	1	0.75	0.01			0.75	0.01				
Phe151											
Asn152											
Pro153											
Thr154	5	0.60	0.01	0.81	0.01	0.74	0.01	855.76	39.13		
Gln155	2	0.86	0.03			0.86	0.03	35.65	20.96		
Leu156											
Glu157											
Glu158	1	0.89	0.02			0.89	0.02				
Gln159											
Cys160											
His161											
Ile162											