Structure determination of bovine ß-lactoglobulin variants A and B

The structural consequences of point mutations

The structural basis of the Tanford transition

The structural influence of ligand on bovine BLGA

Submitted to Massey University as partial fufilment of the requirements for the degree of Doctor of Philosophy in Chemistry

Bin Qin

Spring, 1998

Abstract

From the structure determination of bovine ß-lactoglobulin variants A and B, the structural consequences of point mutations, the structural basis of the Tanford transition and mode of ligand binding by fatty acids have been elucidated.

Bovine ß-lactoglobulin was isolated first in 1934 by Palmer from skim milk. Bovine ß-lactoglobulin contains 162 amino acid residues with 8 major ß-strands which fold forth and back to form a ß-barrel ——creating a cup-shaped molecule or calyx. As a member of the lipocalin superfamily, this protein demonstrates the ability to bind a variety of small hydrophobic molecules, the most notable of which is retinol. Bovine ß-lactoglobulin has at least six variants and undergoes pH-dependent conformational changes. Around pH 7, the conformational change is named as the Tanford transition and is characterized by the exposure of a buried COOH group above pH 7.

Even though bovine ß-lactoglobulin has a long history of structural studies, several regions of this protein molecule have remained poorly characterized. We applied X-ray diffraction techniques and successfully determined the structures of bovine ß-lactoglobulin for variant A in lattice Z at pH 6.2, 7.1 and 8.2, for variant B in lattice Z at pH 7.1, and for variant A in lattice Z at pH 7.3 with the ligand 12-bromododecanoic acid (BrC12) bound. The structures have resolutions of 2.56, 2.24, 2.46, 2.22, and 2.23 Å, respectively. The corresponding values of R (R_f) are 19.19% (24.03%), 23.35% (27.93%), 23.16% (27.69%), 23.93% (28.62)%, and 23.23% (27.93%). The C and N termini, as well as two disulfide bonds, are clearly defined in these models.

Bovine ß-lactoglobulin can be divided into several portions: the flexible top region which involves loops AB, CD, EF, and GH; the more rigid bottom region comprising loops BC, DE and FG; the calyx handle region which partially covers the ß-barrel; and the ß-barrel which is formed by ß-sheets I and II. While the major portions of bovine ß-lactoglobulin remain unchanged over the pH range 6.2 to 8.2, the loop EF, which contains Glu89, experiences a critical conformational change. This transition causes the side chain of Glu89, which is buried at pH 6.2, to become

exposed at pH 7.1 and 8.2. This conformational change provides a structural basis for a variety of pH-dependent phenomena which are collectively known as the Tanford transition.

Bovine ß-lactoglobulin variant A differs from variant B at two point mutation sites: D64G and V118A. The first point mutation occurs in a rather mobile region of bovine ß-lactoglobulin, loop CD, and results in the side chain of Glu65 adopting a different orientation. The second point mutation occurs in a very rigid region of bovine ß-lactoglobulin, ß-sheet II, and results in no significant conformational difference between two variants. While the major portions of the structures of variants A and B at pH 7.1 remain very similar, the loop EF adopts a different conformation. At pH 7.1, loop EF of variant B is closed, whereas that of variant A is opened.

The crystal structure of the complex of bovine ß-lactoglobulin variant A with 12-bromododecanoic acid (BrC12) reveals that the primary binding site for an aliphatic acid is in the interior of the calyx. The ligand (BrC12) has limited influence on the structure of bovine ß-lactoglobulin A. Compared with unliganded BLGA at pH 7.1, one hundred and fifty six Cα atoms (96.3% out of total of 162 Cα atoms) have a displacement smaller than 0.5 Å; among these Cα atoms, the average displacement is only 0.169 Å. The region of major disturbance is loop GH, which has the maximum displacement of Cα, 1.38 Å.

Comparison of the structures of bovine ß-lactoglobulin (BLG) which crystallised in lattices X (triclinic), Y (orthorhombic), Z (trigonal) reveals that the core portion of bovine ß-lactoglobulin, which includes the "bottom" region, the ß-sheets I and II, is relatively more rigid than other portions of bovine ß-lactoglobulin molecule. The Cα atom trace of the structure of bovine ß-lactoglobulin in lattice X is closer to that in lattice Z than that in lattice Y. The "lock and key" interface of bovine BLG uniquely exists in lattice Z, whereas the dimer interfaces of bovine BLG are found in all three lattices. Bovine BLG in lattice X and in lattice Y have similar "bottom:bottom" interface.

Comparison of the structure of bovine ß-lactoglobulin with retinol-binding protein, odorant-binding protein and bilin-binding protein

reveals a similar fold, typical of the lipocalin family: an eight ß-stranded ß-barrel to which is attached a three-turn α-helix. However, the ß-barrel of bovine ß-lactoglobulin is potentially relatively flexible, as it is comprised of two ß-sheets. This flexibility may reflect the fact that bovine ß-lactoglobulin is able to bind a variety of small and not-so-small hydrophobic molecules.

The conformational transition of loop EF may be relevant to the physiological function of ß-lactoglobulin. The acid- and proteinase-resistant bovine ß-lactoglobulin may hold and protect its ligand inside the calyx in the acidic conditions of stomach. After the holo bovine ß-lactoglobulin passes the stomach, the basic conditions of the intestine cause the conformational change of loop EF, which opens the "door" to release the ligand. This property of bovine ß-lactoglobulin suggests a potential pharmaceutical application —— as a shuttle to convey acid-sensitive medicines.

Acknowledgements

Many people have given me their time and advice during my Ph.D period and if I do not mention your name please don't be offended as I do thank you.

The biggest acknowledgement goes to my supervisors Prof. Ted N. Baker and Assoc. Prof. Geoff B. Jameson who gave endless support and guidance throughout a long and difficult Ph.D period. Their contribution was immeasurable in both knowledge and encouragement. Also Mrs Heather M. Baker for her adept crystallization skills and kindly ear when required, and Dr. Maria Bewley who provided me special help during the initial stage of refinement.

I should say special thanks to Assoc. Prof. Michael Hardman who is the first person I met at Massey University three years ago, and Dr. Kathy Kitson who arranged the first contact with several staff members of Massey University.

To the members of Muppet lab: Peter Metcalf, Clyde Smith, Bryan Anderson, Ross Edwards, etc.

To Dr. Lawrie Creamer and NZDRI for the protein samples and helpful discussions.

I am grateful to the New Zealand government for its student allowance and to Prof. Ted N. Baker for the financial assistance he arranged.

Lastly I would like to thank my family who provide their support through my time at Massey University.

Table of Contents

Abbreviations

5-DSA 5-doxylstearic acid

12-DSA 12-doxylstearic acid

16-DSA 16-doxylstearic acid

α2-UG rat urinary protein

apo-D apo-lipoprotein

BBP bilin-binding protein

BLG ß-lactoglobulin

BLGA bovine ß-lactoglobulin variant A

BLGB bovine ß-lactoglobulin variant B

BrC12 12-bromododecanoic acid

ESP rat androgen-dependent epididymal secretory protein

FFA free fatty acid

HC inter-a-trpsin inhibitor light chain precusor

HCHU HC protein from human

HEPES N-(2-hydroxyethyl)piperazine-N'-2-ethanesulfonic acid)

INCYN tobacco hornworm insectocyanin

MW molecular weight

MUP mouse urea protein

OBP pyrazine-binding protein, odorant-binding protein

PCMB *p*-chloromercuri-benzoic acid

RBP retinol-binding protein

RMS root-mean-square

SDS sodium dodecanoic sulfric acid

SDS-PAGE sodium dodecyl sulfate polyacrylamide gel electrophoresis

TAPSO 3-(N-tris(hydroxymethyl)methyl-3-aminopropane-sulfonic acid

List of tables

List of figures

Related publications

1. Structural basis of the Tanford transition of bovine ß-lactoglobulin from crystal structures at three pH values

Authors: Bin Y. Qin, Maria C. Bewley, Lawrence K. Creamer, Heather M. Baker, Edward N. Baker, Geoffrey B. Jameson

Biochemistry, in press

2. Structural consequences of the point mutations of variants A and B of bovine ß-lactoglobulin

Authors: Bin Y. Qin, Lawrence K. Creamer, Maria C. Bewley, Edward N. Baker, Geoffrey B. Jameson

Protein Science, accepted

3. Bromo-dodecanoic acid binds inside the calyx of bovine beta-lactoglobulin

Authors: Bin Y. Qin, Lawrence K. Creamer, Edward N. Baker, Geoffrey B. Jameson

FEBS letters, in press

4. Bovine beta-lactoglobulin and its variants: a three dimensional structural perspective

Authors: Maria C. Bewley, Bin Y. Qin, Geoffrey B. Jameson, Lindsay Sawyer and Edward N. Baker

Intl. Dairy Fed. Bull. Special Issue. "Milk Protein Polymorphism," IDF, Brussels: 1998, 101-109

Chapter 1 Introduction

Mammals rely on milk to nourish their newborn offspring. In the course of evolution, it can be anticipated that individual species may vary the components of their milk to achieve the best survival strategy for their offspring. Comparison of the composition of the milk of human beings and cows reveals a major difference: ß-lactoglobulin (abbreviated BLG), which is abundant in cow's milk, is absent in human's [Hambling et al., 1992]. What is ß-lactoglobulin? What is the biological function of bovine ß-lactoglobulin? What factors are responsible for the existence of BLG in cow but not in human milk?

Bovine ß-lactoglobulin is a small protein molecule with 162 residues. To biochemists, the abundance and easy preparation of bovine ß-lactoglobulin make it a good candidate as a research subject. Up to 1975, ß-lactoglobulin was one of the most thoroughly investigated proteins and was considered to be typical of most globular proteins [Creamer and MacGibbon, 1996]. Investigations of bovine ß-lactoglobulin reveal that it may be involved in the transportation of retinol and/or the metabolism of aliphatic acids [Pérez and Calvo, 1995]. However, the exact physiological function remains uncertain after more than sixty years of extensive research. It is obvious that better knowledge of ß-lactoglobulin is essential to understand its biological function.

In fact, much biochemical research on bovine ß-lactoglobulin has created more puzzles than explanations concerning its biological function. One distinctive biochemical feature of bovine ß-lactoglobulin is that it undergoes pH-dependent conformational changes, the most prominent one around pH 7 called the N $\Leftrightarrow$ R or the Tanford transition [Tanford et al., 1959]. Another distinctive feature of bovine ß-lactoglobulin involves the wide range of ligands that bind quite strongly [O'Neill and Kinsella, 1987], such as retinol, aliphatic acids, porphyrin species, etc. To date, no crystallographic evidence exists to show exactly where the binding site is.

One of the nice characteristics of bovine ß-lactoglobulin is that this protein crystallises relatively easily in multiple crystal forms. This is an encouraging aspect of bovine ß-lactoglobulin structural research

activities. X-ray diffraction studies have provided the evidence which ascribes this protein to the lipocalin super-family. Members of this super-family have a calyx-like (or cup-like) fold and are capable of binding small hydrophobic molecules. Unfortunately, X-ray diffraction studies have not answered all the puzzles of ß-lactoglobulin. One very recent publication has described this protein as "still an enigmatic lipocalin" [Brownlow *et al.*, 1997].

To add to the complexity, bovine ß-lactoglobulin has multiple variants. Some authors [Green and Aschaffenburg, 1959] have suggested that it is not important to distinguish the different bovine ß-lactoglobulin variants during investigations of the biochemical characteristics. However, it has been observed that milk dominated by different variants behaves differently during industrial processing [Hill *et al.*, 1996]. This is something else that structural biologists can address.

The previous research history of bovine ß-lactoglobulin has demonstrated difficulties in the studies of bovine ß-lactoglobulin. To achieve scientific advances, we can hope that the more structural information we obtain for ß-lactoglobulin, the more knowledge we have about ß-lactoglobulin, and the more we can answer the queries on the structure-function relationships of ß-lactoglobulin. The structure determinations of bovine ß-lactoglobulin A and B in lattice Z at a variety of conditions are the focus points of this research project.

ß-Lactoglobulin research is reviewed in chapter 2. The history of its discovery is described first. Following the history, the biochemical and biological properties of bovine ß-lactoglobulin are discussed, along with its nomenclature, isolation methods, sequence analyses, structure information in solution, and previous X-ray diffraction studies. Chapter 2 ends with a statement of the aims of this project.

The principles of X-ray diffraction techniques are reviewed in chapter 3. The principle of diffraction of X-rays from a crystal, Fourier transformations, the Patterson function, and the principles of refinement are discussed. On the basis of chapter 3, the details of the whole experimental procedure are described in chapter 4, which includes the source of protein material, purification procedure, crystallization

trials, data collection and processing, and structure determination and refinement.

The results are presented in chapter 5, included four fully determined bovine ß-lactoglobulin structures, and one structure with an aliphatic acid bound to bovine ß-lactoglobulin. The structures of bovine ß-lactoglobulin are disassembled into sub-regions and examined. The four major regions are the top region, bottom region, handle region, and the ß-barrel which is further segregated into ß-sheets I and II. The structures are described based on these sub-regions. Some residues and the ligand binding site are examined further.

The structures are discussed in detail in chapter 6. The structural basis of the Tanford transition has been elucidated along with the structural disturbance of bovine BLG after ligand binding, the structural consequences of point mutations of BLG, the structural differences among bovine BLG in different lattice types, and the structural relationship of bovine BLG with lipocalins.

In chapter 7, the final conclusions are made about current research, and suggestions are offered for future research.

Chaper 2 Review

2.1 Discovery and early research history

2.1.1 Discovery and naming of bovine ß-lactoglobulin

The main whey proteins of mature bovine milk are [Bell, 1964]: ß-lactoglobulin (~0.3 g/100 mL), α-lactalbumin (~0.1 g/100 mL), serum albumin (~0.04 g/100 mL) and immunoglobulin (~0.08 g/100 mL). ß-Lactoglobulin is, therefore, the major component of whey protein. The composition of milk was unclear in the 1930s. Albumin was recognised as the major component of the skim milk of cow and was the general subject of much research. Palmer (1934) followed Sjögren and Svedberg's (1930) lactalbumin preparation procedure to isolate albumin, but claimed a new globular protein had been crystallized. Subsequently it was referred to as Palmer's crystalline protein.

Palmer did not name his discovery. A few years later, Cannan *et al.* (1941) suggested a nomenclature "ß-lactoglobulin" for Palmer's crystalline protein. The Greek prefix derived from Pedersen's publication (1936a), where three major components of skim milk were distinguished based on ultracentrifugation studies. Palmer's crystalline protein corresponded to the second component of the skim milk. This term also identifies the source (lacto for milk) and the basic characteristic (globulin for globular protein) of Palmer's crystalline protein, yet distinguishes it from the classical globulin fraction of milk.

2.1.2 Purification of Palmer's crystalline protein

Palmer's crystalline protein was regarded as a purified protein sample for a long time. Observations in support of this conclusion included that: (i) this protein was purified by the method of crystallization; (ii) the free boundary electrophoresis experiment conducted by Pedersen (1936b) showed homogeneity of this protein, and (iii) SΦrensen and Palmer (1938) showed that its solubility was independent of preparation.

However, the fact that different values [Pedersen, 1936a; Bull and Curry, 1946a] had been reported for the molecular weight of Palmer's crystalline

protein, as well as the different contents of certain amino acids [Brand *et al.*, 1945; Chibnall, 1946] suggested that this crystalline sample might not be a pure substance. Choh Hao Li (1946) first provided the evidence by means of high resolution electrophoresis that showed that Palmer's crystalline protein was in fact heterogeneous. No matter how many times Palmer's crystalline protein was re-crystallized, there were at least three electrophoresis boundaries. This observation was supported by the isolation of the sub-components of Palmer's crystalline protein in 1950. Polis *et al.* (1950) demonstrated that at least two different components existed in the ß-lactoglobulin prepared according to Palmer's method. They named them ß1 and ß2. The solubility of the ß1 component is obviously different from Palmer's crystalline protein. The relative amount of the two sub-components was constant in general. The notable exception appeared in the sample from the first-day colostrum of one cow [Polis *et al.*, 1950]. Only a single component was observable from this sample.

Experimental techniques as well as biological theory evolved dramatically in the 1950s. Free-boundary electrophoresis was replaced by solid-support electrophoresis. By paper electrophoresis, Aschaffenburg and Drewry (1955) separated Palmer's crystalline protein clearly into ß1-lactoglobulin and ß2-lactoglobulin. In the mid 1950s, the origin of proteins was related to the expression of individual genes. Their results also demonstrated the relationship between the occurrence of ß1 or ß2-lactoglobulin and the genetic type of cow. From the information which their publication provided, with the aid of modern genetic theory, we can conclude that bovine ß-lactoglobulin is controlled by a single gene site on one of the autosomes. No more than two variants of ß-lactoglobulin can occur in the milk from a single cow. However, if ß-lactoglobulin is a protein which enjoys some flexibility during the course of evolution, more variants can be anticipated from bovine gene pool.

2.2 Terminology

Bovine ß-lactoglobulin is only one member of the widely distributed ß-lactoglobulin family. To classify ß-lactoglobulins, we need to indicate the milk source (secreted by which animal), gene series (some animals

may have more than one ß-lactoglobulin gene, usually labelled as I, II, etc.), and the variant type (each gene may have several point mutants, labelled as A, B, C, etc.). Compared to many other macromolecules, bovine ß-lactoglobulin crystallises rather readily in a variety of crystal forms [Aschaffenburg *et al.*, 1965]. These crystal forms have been assigned a lattice type code. Therefore, for research involving X-ray diffraction experiments, the lattice type code of ß-lactoglobulin crystals is also required, which specifies the crystal cell dimensions and the space group. Note that bovine ß-lactoglobulin is referred to as bovine BLG throughout this thesis.

2.3 The biological properties of ß-lactoglobulin

2.3.1 ß-lactoglobulin and lipocalins

The sequence homology of ß-lactoglobulin with other proteins has been noted by several authors. Pervaiz and Brew (1985) noted the sequence homology among human protein HC (also known as α-1-microglobulin, or inter-α-trypsin inhibitor light chain precusor (protein HC) from human, abbreviated as HCHU), RBP and ß-lactoglobulin. Drayna (1986) included human apo-lipoprotein (apo-D), tobacco hornworm insecticyanin (INCYN) or bilin-binding protein (BBP), rat urinary protein (α2-UG) and rat androgen-dependent epididymal secretory protein (ESP) as proteins which share amino acid sequence similarity with ß-lactoglobulin. Lee *et al.* (1987) added a frog protein BG to this list, where BG is a protein corresponding to an abundant mRNA in the cells of Bowman's gland in the olfactory tissue of frog.

Members of this potential protein family, including BLG, have a mass around 20,000 Dalton, and are able to bind small hydrophobic molecules. Later this protein family was classified as a protein super-family and titled as "lipocalin" [Pervaiz and Brew, 1987]. Now, at least 68 individual lipocalins can be browsed in the SWISS-PROT data base. While the binding of hydrophobic molecules is the basic property of members of this super-family, the functional importance of binding by individual proteins is quite vague. For instance, the physiological function of ß-lactoglobulin remains uncertain after many years of research.

Lipocalins generally exhibit only a low level of residue identity, often no more than 20% between pairs of sequences. This is below the 25% level where structural similarities frequently go undetected [Scader and Schneider, 1991]. In cases of such low similarity, the sequence alignment can lose statistical significance. Identifying the existence of lipocalins by sequence information is practically based on the motifs of protein sequence [Flower *et al.*, 1993]. The heart of the motif-based sequence analysis (Parry-Smith and Attwood, 1992) is that a protein family can be characterized by one or more conserved motifs, that the order of motifs within a sequence is maintained, and, optionally, that the maximum distance (in residues) between the motifs is restricted.

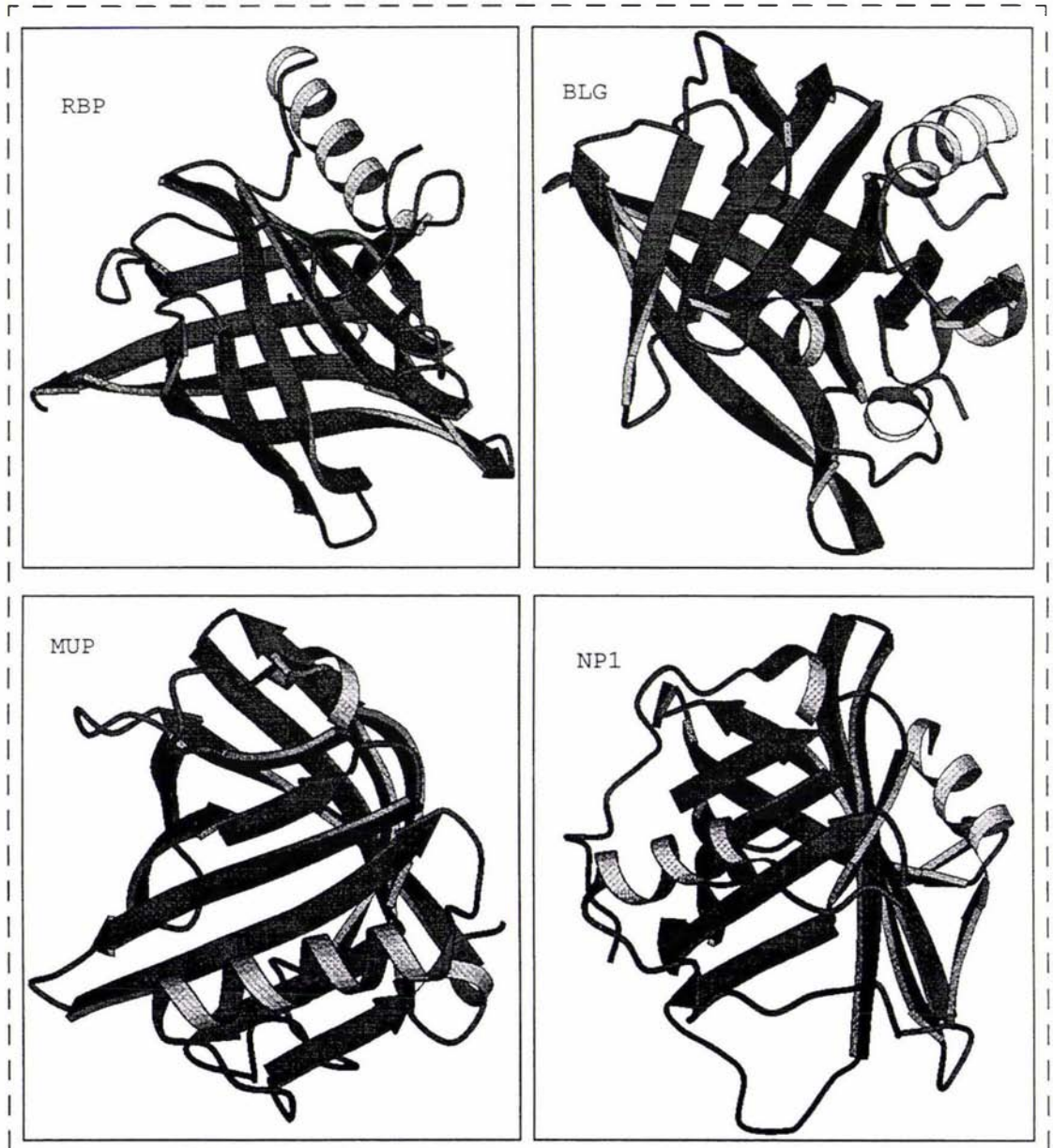


Figure 231-1: The lipocalin fold, exemplified by RBP, etc.
Figure prepared by MOLSCRIPT [Kraulis, 1991].

The motif bas-X-X-G-X-W-aro (bas=basic residue, aro=aromatic residue) occurs near residue 20 (in the bovine BLG sequence). A second motif T-D-Y-X-X-aro appears around residue 100 (in the bovine BLG sequence). Between these two segments is a cysteine that may form a disulfide bond with another cysteine at the C-terminal end. A third motif R-X-P occurs around residue 124. Each motif corresponds to an important structurally conserved region of the lipocalin fold. Under such criteria, we are encountering a rapidly burgeoning protein family. The analysis shows the lipocalin protein family to be composed of a core set of quite closely related proteins, the kernel lipocalins, which form by far the largest self-consistent subset within the whole set of related sequences. The kernel lipocalins share the three conserved sequence motifs. A smaller number of more divergent sequences, the outlier lipocalins, share only two of the three motifs. Bovine ß-lactoglobulin has all three motifs, which are "KVAGTWY", "TDYKKY" and "RTP", respectively. A disulfide bond is formed between residue 66 and residue 160. Therefore, bovine ß-lactoglobulin belongs to the kernel part of the lipocalin super-family.

The structures of several lipocalins have been determined by X-ray diffraction techniques: plasma RBP [Cowan *et al.*, 1990; Zanotti *et al.*, 1993], BBP [Huber *et al.*, 1992], OBP [Tegoni *et al.*, 1996], BLG [Green *et al.*, 1979; Brownlow *et al.*, 1997; Qin *et al.*, 1998], MUP [Bocskei *et al.*, 1992], ERBP [Newcomer, 1993] and NP1 protein [Weichsel, *et al.*, 1998]. The similarity in three-dimensional structure is obvious enough to include them into a single protein family, which is characterized by a common fold: a calyx-like central 8-stranded ß-barrel, plus one long three-turn α-helix. The calyx folds of the lipocalins RBP, MUP, NP1 and BLG are displayed in Figure 231-1.

2.3.2 Gene structure of two lipocalins: BLG and MUP

Analysis of the gene structure of ß-lactoglobulin from sheep also supports the idea that these related proteins form a protein family [Ali and Clark, 1988]. The primary structures of sheep ß-lactoglobulin and mouse major urinary protein (MUP, Unterman *et al.*, 1981) have limited amino-acid sequence homology. However, the genes which encode these two proteins have similar intron and exon patterns. The analogous introns are in phase with respect to the reading frame. Both of these genes encode

a 162-residue molecule with six exons with one untranslated exon. The analogous exons in the genes have similar size and function (see Table 232-1).

Table 232-1: Similar gene structure of two lipocalins

	BLG	MUP
Exon I	136 bp	161 bp
Exon II	140 bp	134 bp
Exon III	74 bp	74 bp
Exon IV	111 bp	111 bp
Exon V	105 bp	102 bp
Exon VI	Stop codon	Stop codon
Exon VII	Untranslated	Untranslated

2.3.3 Bovine lipocalins

Retinol-binding protein (RBP) and odorant-binding protein (OBP) are two structurally defined lipocalins which exist in cow. Retinol-binding protein [Zanotti *et al.*, 1993] is the highly specific carrier of vitamin A *in vivo*. It is biosynthesized in the liver and circulated in the blood, where it forms a complex with transthyretin (also called pre-albumin). This macromolecule makes the insoluble retinol molecule transferable in the blood and delivers retinol to the cells with appropriate receptors. RBP is very sensitive to the conformation of retinol, binding only the all-trans conformation of retinol [Jang and Swaisgood, 1990]. RBPs are found not only in bovine blood, but also in the blood of human [Rask *et al.*, 1979], rat and rabbit [Sundelin, 1985], which are species without ß-lactoglobulin. The sequences of these RBPs demonstrate high identity, and most of the sequence changes are concentrated in restricted regions of the molecule [Pelosi, 1989].

OBP (also called pyrazine-binding protein) is isolated in high concentration in the nasal mucosa homogenates of cows [Bignetti *et al.*, 1985]. This protein was first identified by its ability to bind 2-isobutyl-3-methoxy-pyrazine [Pelosi, 1989]. Subsequently it was found that many other hydrophobic volatile molecules can be bound to its active site [Pelosi, 1989; Pevsner, 1990]. This protein is a dimer with a single binding site per monomer for hydrophobic ligands. After bovine OBP [Tirindelli *et al.*, 1989] and mouse OBP [Pevsner and Reed, 1988] primary

structures were determined, the sequence homology with ß-lactoglobulin suggested that OBP belongs to the RBP-BLG protein family [Lee, 1987]. The subsequent tertiary structure determinations of these three proteins support this idea [Monaco and Zanotti, 1992].

The sequence alignment of these three proteins provides useful information, as summarised in Table 233-1. We can trace the lipocalin motifs on both ß-lactoglobulin and retinol-binding protein. Indeed, both of them are favourable retinol carriers. Comparing the sequence identity of OBP to these two proteins, we found that OBP has a close relationship to RBP. Two of the lipocalin motifs (TDYxxaro and RXP) disappear in the OBP. But lipocalins are based on their folding similarity more than their sequence identity.

Table 233-1: Sequence alignment of structurally defined bovine lipocalins by MULTALIN [Corpet, 1988]

	1*	34
BLG	MKCLLLALAL TCGAQALIVT QTMKGLDIQK VAGTWYSLAM AASDISLLDA	
RBP	........ER DCRVSSFRVK EN...FDKAR FAGTWYAMAK KDPEGLFLQD	
OBP	....AQEEEAE QNLSEL.... .SGPWRTVYI GSTNPEKIQE	
Consensus	.... .c.aq...v. qn...ld... .aGtWy..a. ....lq.	
	35	83
BLG	QSAP.LRVYV EELKPTPEGD LEILLQKWEN GECAQKKIIA EKTKIPAVFK	
RBP	NIVAEFSVDE NGHMSATAKG RVRLLNNWDV ..CADMVGTF TDTEDPAKFK	
OBP	NGP..FRTYF RELVFDDEKG TVDFYFSVKR DGKWKNVHVK ATKQDDGTYV	
Consensus	n....frvy. .el....ekg .v.ll..w.. ..ca..v... ..t.dpa.fk	
	84	122
BLG	IDALNENKVL ....VL DTDYKKYLL. F.CMENSAEP EQSLACQ.CL	
RBP	MKYWGVASFL QKGNDDHWII DTDYETFAVQ YSCRLLNLDG TCADSYSFVF	
OBP	ADY....... .EGQNVFKIV SLS.RTHLVA HNINVDKHGQ TTELTELFV.	
Consensus	.dy......l ..g.....i. dtdy.t.lv. ..c..... t..l...fv.	
	123	162
BLG	VRTPEVDDEA LEKFDKALKA ...LPMHIRL SFNPT..QLE EQCHI.	
RBP	ARDPSGFSPE VQKIVRQRQE ELCLARQYRL IPHNGYCDGK SERNIL	
OBP	..KLNVEDED LEKF.WKLTE DKGIDKKNVV ....NFLENE DHPHPE	
Consensus	.r.p.v.de. leKf...l.e ...l....rl ....e ...hi.	

* Sequence numbers for the mature (post-translation) form of bovine BLG are given. The lipocalin motifs are underlined.

2.3.4 Distribution of ß-lactoglobulin

Since the 1960s, the existence of ß-lactoglobulin in other animals has been investigated in several laboratories [Lyster *et al.*, 1966; Liberatori *et al.*, 1979b; Pervaiz and Brew, 1986; etc.]. Accumulated data demonstrate that this secreted globular protein is widely distributed, as indicated in Table 234-1. Most ß-lactoglobulins occur as dimers, while monomeric ß-lactoglobulins exist in some non-ruminant species such as pig, horse, cat, etc. No similar macromolecule has been found in rodents, such as mouse and rabbit. It is believed that no BLG exists in human milk [Bell and McKenzie, 1964; Liberatori *et al.*, 1979a].

Table 234-1: Distribution of ß-lactoglobulin among different species

Species	Type	Reference
Cow (*Bos taurus*)	dimer	Bull, 1946a,b
Oxen(*Bos javanicus*)	dimer	Bell, 1981
Yak(*Bos grunniens*)	dimer	Grosclaude, 1966
Zebu(*Bos indicus*)	dimer	Lyster, 1966
Buffalo(*Bubalis bubalis*)	dimer	Lyster, 1966
Bison(*Bison bison*)	dimer	Lyster, 1966
Musk ox(*Ovibos moschatus*)	dimer	Lyster, 1966
Eland(*Taurotragus oryx*)	dimer	Lyster, 1966
Goat(*Capra hircus*)	dimer	Askonas, 1954
Sheep(*Ovis aries*)	dimer	Bell, 1967
Mouflon(*Ovis ammon musimon*)	dimer	Liberatori, 1979b
Red deer(*Cervus elaphus L.*)	dimer	Lyster, 1966
European elk(*Alees alees L.*)	dimer	Lyster, 1966
Reindeer(*Rangifer tarandus L.*)	dimer	Lyster, 1966
White-tailed deer (*Odocoileus virginianus*)	dimer	Lyster, 1966
Fallow deer(*Dama dama*)	dimer	Lyster, 1966
Caribou(*Rangifer arcticus*)	dimer	Lyster, 1966
Giraffe(*Giraffa camelopardalis*)	dimer	Lyster, 1966
Okapi(*Okapia johnstoni*)	dimer	Lyster, 1966

(continued)

Table 234-1: Distribution of ß-lactoglobulin among different species

Species	Type	Reference
Prong-horn antelope (*Antilocapra americana*)	dimer	Lyster, 1966
Camel(*Camelus dromedarius*)	monomer	Liberatori, 1979
Peccary(*Pecari tajacu*)	monomer	Lyster, 1966
Pig(*Sus scrofa domestica*)	monomer	Liberatori, 1979b
Horse(*Equus caballus*)	monomer	Liberatori, 1979b
Zebra(*Equus quagga*)	monomer	Lyster, 1966
Rhinoceros(*Diceros bicornis*)	monomer	Lyster, 1966
Donkey(*Equinus asinus*)	monomer	Godovac-Zimmermann, 1988
Dolphin(*Tursiops truncatus*)	monomer	Pervaiz, 1986
Manatee (*Trichechus manatus latirostris*)	monomer	Pervaiz, 1986
Beagle(*Canis familiaris*)	monomer	Pervaiz, 1986
Cat(*Felis cutus*)	monomer	Halliday, 1991
Grey kangaroo (*Macropus giganteus*)	monomer	McKenzie, 1983
Red kangaroo(*Macropus rufus*)	monomer	McKenzie, 1983
Llama(*Lama glaba L.*)	none	Fernandez, 1988
Mouse(*Mus musculus*)	none	Clark, 1987
Rat(*Ratus norvegicus*)	none	Hennighausen, 1982
Guinea-pig(*Cavia porcellus*)	none	Brew, 1967
Rabbit(*Oryctolagus cuniculus*)	none	Lyster, 1966
Human(*Homo sapiens*)	none	Bell, 1964

The ruminant animals have dimeric ß-lactoglobulin and show very high sequence identity to bovine ß-lactoglobulins. Genetic polymorphism of ß-lactoglobulin is more common in ruminant animals than in the other species. Cow has several variants as mentioned before. Donkey possess two variants [Godovac-Zimmermann *et al.*, 1990], whereas sheep have variants A, B and C [Erhardt *et al.*, 1989], which are different from the A, B and C variants for cows. Horse [Clark *et al.*, 1987; Halliday *et al.*, 1991], cat [Halliday *et al.*, 1990] and dog [Halliday *et al.*, 1993] have more than one form of the ß-lactoglobulin gene and they are not the products of an autosomal allele (a pair of genes in the cell irrespective of the sex). In other words, they are not phenotypes of one gene. The

SWISS-PROT protein data bank now has at least 14 sequences available for ß-lactoglobulin, excluding variants.

Sequence alignment of available ß-lactoglobulins is displayed in Table 234-2. The first 16 residues of bovine ß-lactoglobulin are the signal peptide, which do not appear in the mature protein. In contrast to cow, pig's ß-lactoglobulin retains the signal peptide. The sequence motifs underlined are characteristic of the lipocalin super-family. The high sequence identity among ß-lactoglobulins is obvious, except for Kangaroo. In the consensus sequence, there are 21 perfect matched residues (11.35 %). The two disulfide bonds between residues 66-160 and 106-119 are conserved in all ß-lactoglobulins. Residue 89 (bold letter), a glutamic acid, is also conserved throughout all ß-lactoglobulins. Figure 617-1 displays the location of the conserved residues of BLG in a stereodiagram.

Table 234-2: Sequence alignment of ß-lactoglobulins, by CLUSTAL W [Thompson *et al.*, 1994]

ID	1	10	20	30	40[a]
LACB_BOVIN	MKCLLLALALT--CGAQALIVTQTMKGLDIQKVAGTWYSLAMAASDISLLDAQSAPLR-VY				
LACB_BUBAR	----------IIVTQTMKGLDIQKVAGTWYSLAMAASDISLLDAQSAPLR-VY				
LACB_CAPHI	MKCLLLALGLALACGIQAIIVTQTMKGLDIQKVAGTWYSLAMAASDISLLDAQSAPLR-VY				
LACB_SHEEP	MKCLLLALGLALACGVQAIIVTQTMKGLDIQKVAGTWHSLAMAASDISLLDAQSAPLR-VY				
LACB_PIG	MRCLLLTLGLALLCGVQAVEVTPIMTELDTQKVAGTWHTVAMAVSDVSLLDAKSSPLK-AY				
LACA_CANFA	----------IVVPRTMEDLDLQKVAGTWHSMAMAASDISLLDSETAPLR-VY				
LACC_CANFA	----------IVIPRTMEDLDLQKVAGTWHSMAMAASDISLLDSETAPLR-VY				
LACB_FELCA	----------ATVPLTMDGLDLQKVAGMWHSMAMAASDISLLDSETAPLR-VY				
LACC_FELCA	----------ATVPLTMDGLDLQKVAGTWHSMAMAASDISLLDSEYAPLR-VY				
LACB_EQUAS	----------TNIPQTMQDLDLQEVAGKWHSVAMAASDISLLDSEEAPLR-VY				
LACB_HORSE	MKCLLLALGLALMCGIQATNIPQTMQDLDLQEVAGKWHSVAMAASDISLLDSESAPLR-VY				
LACA_EQUAS	----------TDIPQTMQDLDLQEVAGRWHSVAMVASDISLLDSESAPLR-VY				
LACA_HORSE	MKCLLLALGLSLMCGNQATDIPQTMQDLDLQEVAGRWHSVAMVASDISLLDSESVPLR-VY				
LACA_FELCA	----------ATLPPTMEDLDIRQVAGTWHSMAMAASDISLLDSETAPLR-VY				
LACB_MACEU	MKFLLLTVGLALIGAIQAVENIRSKNDLGVEKFVGSWYLREAAKT----MEFSI-PLFDMD				
LACB_MACGI	----------VENIRSKNDLGVEKFVGSWYLREAAKT----MEFSI-PLFDMD				
Q29146	MKFLLLTVGLTSICAIQAIENIHSKEELVVEKLIGPWYRVEEAKA----MEFSI-PLFDMN				
Consensus[b]		* * *			**

(continued)

Table 234-2: Sequence alignment of ß-lactoglobulins, by CLUSTAL W [Thompson *et al.*, 1994]

```
ID              43        53        63        73        83          93
----------------------------------------------------------------------
LACB_BOVIN      VEELKPTPEGDLEILLQKWENGECAQKKIIAEKTKIPAVFKIDALNENKVLVLDTDY
LACB_BUBAR      VEELKPTPEGDLEILLQKWENGECAQKKIIAEKTKIPAVFKIDALNENKVLVLDTDY
LACB_CAPHI      VEELKPTPEGNLEILLQKWENGECAQKKIIAEKTKIPAVFKIDALNENKVLVLDTDY
LACB_SHEEP      VEELKPTPEGNLEILLQKWENGECAQKKIIAEKTKIPAVFKIDALNENKVLVLDTDY
LACB_PIG        VEGLKPTPEGDLEILLQKRENDKCAQEVLLAKKTDIPAVFKINALDENQLFLLDTDY
LACA_CANFA      IQELRPTPQDNLEIVLRKWEDGRCAEQKVLAEKTEVPAEFKINYVEENQIFLLDTDY
LACC_CANFA      IQELRPTPQDNLEIVLRKWEDNRCVEKKVFAEKTELAAKFSINYVEENQIFLLDTDY
LACB_FELCA      VQELRPTPRDNLEIILRKWEDNRCVEKKVLAEKTECAAKFNINYLDENELIVLDTDY
LACC_FELCA      VQELRPTPRDNLEIILRKWEQKRCVQKKILAQKTELPAEFKISYLDENELIVLDTDY
LACB_EQUAS      IEKLRPTPEDNLEIILREGENKGCAEKKIFAEKTESPAEFKINYLDEDTVFALDSDY
LACB_HORSE      IEKLRPTPEDNLEIILREGENKGCAEKKIFAEKTESPAEFKINYLDEDTVFALDTDY
LACA_EQUAS      VEELRPTPEGNLEIILREGANHVCVERNIVAQKTEDPAVFTVNYQGERKISVLDTDY
LACA_HORSE      VEELRPTPEGNLEIILREGANHACVERNIVAQKTEDPAVFTVNYQGERKISVLDTDY
LACA_FELCA      VQELRPTPRDNLEIILRKRENHACIEGNIMAQRTEDPAVFMVDYQGEKKISVLDTDY
LACB_MACEU      IKEVNLTPEGNLELVLLEKAD-RCVEKKLLLKKTQKPTEFEIYISSESASYTFSVMETDY
LACB_MACGI      IKEVNLTPEGNLELVLLEKTD-RCVEKKLLLKKTKKPTEFEIYISSESASYTFCVMETDY
POSSUM          IKEVNRTPEGNLELIVLEQTD-SCVEKKFLLKKTEKPAEFEIYIPSESASYTLSVMETDY
Consensus             **   **         *         *     *      *           **
                100       110       120       130       140       150
LACB_BOVIN      KKYLLFCMENSAEPEQ-SLACQCLVRTPEVDDEALEKFDKALKALPMHIRLSFNPTQLEE
LACB_BUBAR      KKYLLFCMENSAEPEQ-SLACQCLVRTPEVDDEALEKFDKALKALPMHIRLSFNPTQLEE
LACB_CAPHI      KKYLLFCMENSAEPEQ-SLACQCLVRTPEVDKEALEKFDKALKALPMHIRLAFNPTQLEG
LACB_SHEEP      KKYLLFCMENSAEPEQ-SLACQCLVRTPEVDNEALEKFDKALKALPMHIRLAFNPTQLEG
LACB_PIG        DSHLLLCMENSASPEH-SLVCQSLARTLEVDDQIREKFEDALKTLSVPMRIL--PAQLEE
LACA_CANFA      DNYLFFCEMNADAPQQ-SLMCQCLARTLEVDNEVMEKFNRALKTLPVHMQLL-NPTQAEE
LACC_CANFA      DNYLFFCMENANAPQQ-SLMCQCLARTLEVNNEVIGKFNRALKTLPVHMQLL-NPTQVEE
LACB_FELCA      ENYLFFCLENADAPDQ-NLVCQCLTRTLKADNEVMEKFDRALQTLPVHVRLFFDPTQVAE
LACC_FELCA      ENYLFFCLENADAPGQ-NLVCQCLTRTLKADNEVMEKFDRALQTLPVDVRLFFDPTQVAE
LACB_EQUAS      KNYLFLCMKNAATPGQ-SLVCQYLARTQMVDEEIMEKFRRALQPLPGRVQIVPDLTRMAE
LACB_HORSE      KNYLFLCMKNAATPGQ-SLVCQYLARTQMVDEEIMEKFRRALQPLPGRVQIVPDLTRMAE
LACA_EQUAS      AHYMFFCVGPCLPSAEHGMVCQYLARTQKVDEEVMEKFSRALQPLPGHVQIIQDPSGGQE
LACA_HORSE      AHYMFFCVGPPLPSAEHGMVCQYLARTQKVDEEVMEKFSRALQPLPGRVQIVQDPSGGQE
LACA_FELCA      THYMFFCMEAPAPGTENGMMCQYLARTLKADNEVMEKFDRALQTLPVHIRIILDLTQGKE
LACB_MACEU      DSYFLFCLYNISDREK--MACAHYVRRIE-ENKGMNEFKKILRTLAMPYTVI--EVRTRD
LACB_MACGI      DSYFLFCLYNISDREK--MACAHYVRRIE-ENKGMNEFKKILRTLAMPYTVI--EVRTRD
Q29146          DNYILGCLENVNYREK--MACAHYERRIE-ENKGMEEFKKIVRTLTIPYTMI--EAQTRE
Consensus             *           *    *          *      *
                                                          (continued)
----------------------------------------------------------------------
```

Table 234-2: Sequence alignment of ß-lactoglobulins, by CLUSTAL W [Thompson *et al.*, 1994]

ID	159	Species	Reference
LACB_BOVIN	QCHI	Cow	[Braunitzer, 1973]
LACB_BUBAR	QCHV	Buffalo	[Kolde, 1981]
LACB_CAPHI	QCHV	Goat	[Preaux, 1979]
LACB_SHEEP	QCHV	Sheep	[Ali, 1990]
LACB_PIG	QCRV	pig	[Alexander, 1992]
LACA_CANFA	QCLI	Dog	[Halliday, 1993]
LACC_CANFA	QCLV	Dog	[Halliday, 1993]
LACB_FELCA	QCRI	Cat	[Halliday, 1990]
LACC_FELCA	QCRI	Cat	[Halliday, 1991]
LACB_EQUAS	RCRI	Donkey	[Godovac, 1988]
LACB_HORSE	RCRI	Horse	[Halliday, 1991]
LACA_EQUAS	RCGF	Donkey	[Godovac, 1990]
LACA_HORSE	RCGF	Horse	[Godovac, 1985]
LACA_FELCA	QCRV	Cat	[Halliday, 1991]
LACB_MACEU	MCHV	Wallaby	[Collet, 1995]
LACB_MACGI	MCHV	Kangaroo	[Godovac, 1987]
Q29146	MCRV	Possum	[Hunter, 1995]
Consensus	*		

[a] The sequence number is from the mature bovine ß-lactoglobulin

[b] Consensus length 185

Number of perfect matches = * 21 => 11.35 %

2.3.5 Variants of bovine ß-lactoglobulin

As mentioned earlier, if ß-lactoglobulin enjoys some flexibility in the course of evolution, a number of variants may be found in the bovine gene pool. Bell (1962) reported a new variant of bovine ß-lactoglobulin, which was found in Australia after screening milk samples. For convenience, Bell denoted ß1-lactoglobulin as variant A, ß2-lactoglobulin as variant B and his new discovery as variant C, according to the order of appearance on the starch gel electrophoresis from the positive electrode to the negative electrode.

Extensive investigation reveals that ß-lactoglobulin variants occur in different bovine sub-species frequently. It is believed that the cow (*Bos*

taurus) has at least six variants: ß-Lactoglobulin A, B, C and D [Brignon and Ribadeau-Dumas, 1969, 1973], H [Conti *et al.*, 1988] and W [Godovac-Zimmermann *et al.*, 1990]. ß-Lactoglobulins labelled as E, F and G [Bell *et al.*, 1981] have been identified in *Bos javanicus*. The sub-species *Bos indicus* has the variants A and B [Aschaffenburg, 1968]. Modification after translation occurs on some sub-species, as exemplified by the occurrence of the variant Dr [Bell *et al.*, 1970] which is an N-glycosylated variant and found in the Droughtmaster breed in Australia.

Each sub-species is dominated by only a few kinds of variants. Comparison of sequence information reveals that only a few sites are subjects of mutation, as shown in Table 231-1. For *Bos javanicus*, there are two mutation sites, at residues 50, 78. Each pair of variants has at most three mutation sites.

Table 235-1: The known mutation sites for bovine BLG

variants	residue sites/type								sub-species
	45	50	56	59	64	78	118	158	*Bos*
A	E	P	I	Q	D	I	V	E	*taurus*
B	E	P	I	Q	G	I	A	E	*taurus*
C	E	P	I	H	G	I	A	E	*taurus*
D	Q	P	I	Q	G	I	A	E	*taurus*
W	E	P	L	Q	G	I	A	E	*taurus*
E	E	P	I	Q	G	I	A	G	*javanicus*
F	E	S	I	Q	G	I	A	G	*javanicus*
G	E	P	I	Q	G	M	A	G	*javanicus*

2.4 Isolation of bovine ß-lactoglobulin

The isolation of ß-lactoglobulin from the milk of cows involves removal of fat, caseins and several other whey proteins. In early studies, Palmer (1934) made use of skim milk as the starting material. Casein was removed by adding hydrochloric acid at pH 4.6, which precipitated casein. After removing the precipitant, sodium sulphate (Na_2SO_4), which the author claimed as a better substitute for ammonium sulphate, was then used to

fractionate the rest of the milk. ß-Lactoglobulin was obtained after dialysis and recrystallization.

Aschaffenburg and Drewry (1957) improved Palmer's method. Anhydrous Na_2SO_4 (20g/100mL) was added to the cow's milk at 40 °C. The filtrate of this mixture contains α-lactalbumin and ß-lactoglobulin. Addition of HCl, reducing pH to 2 caused the precipitation of all proteins except for ß-lactoglobulin. The supernatant solution was brought back to pH 6 by addition of ammonia solution, then ammonium sulphate (20g/100mL) was added to precipitate bovine ß-lactoglobulin. Further purification can be achieved by recrystallization.

Monaco *et al.* (1987) introduced a procedure for bovine ß-lactoglobulin isolation, which avoided the harsh conditions (described above), to provide ß-lactoglobulin suitable for high quality crystals. The fat was removed from the milk by centrifuging twice for 1 hour at 9000 revs/min. The casein was removed by introducing $CaCl_2$ at pH 6.6. The remaining solution was dialysed against 20 mM Tris-HCl at pH 7.2. After removal of precipitated protein, the supernatant was then applied to a DEAE-cellulose column, which was pre-equilibrated at the same conditions: 20 mM Tris-HCl, pH 7.2. Bovine ß-lactoglobulin was adsorbed onto the column and eluted with 0.4 M NaCl. Final purification was done by gel filtration on G-100.

2.5 Properties of bovine ß-lactoglobulin

2.5.1 Physical properties

A number of variants exist for bovine ß-lactoglobulin. Characteristics of each variant are not exactly the same. For example, the solubility of variant B is five times larger than variant A [Hambling *et al.*, 1994]. Except for a few such deviations, in general, the different variants share similar physical and chemical properties [Green and Aschaffenburg, 1959]. Some basic physico-chemical properties of bovine ß-lactoglobulin are given in Table 251-1. We notice that the axial ratio is 2:1, and the molecular weight is 36,000 Dalton or two times larger than that obtained from the primary structure. These data imply that bovine ß-lactoglobulin is a dimer in solution. The sedimentation coefficient of bovine ß-

lactoglobulin varies slightly around pH 7 [Pedersen, 1936b]. The extinction coefficient is rounded to 1 mg/mL/O.D. in the later experiments for determining protein concentration for crystallization.

Table 251-1: Physical properties of bovine ß-lactoglobulin

Parameter	Value	(unit)	Reference
Molecular weight	36,000	Dalton	[Townend, 1957]
Extinction coefficient (278nm)	0.96	litres/g/cm	[Townend, 1960]
Axial ratio	2:1		[Green, 1959]
Sedimentation coefficient	2.83	$S^{O}_{20w} \times 10^{13} s^{-1}$	[Cecil, 1949]
Intrinsic viscosity	2.9	mL/g	[Townend, 1960]

2.5.2 Binding properties

Bovine ß-lactoglobulin belongs to the lipocalin family. The basic character of this family is the ability to bind small hydrophobic molecules. ß-Lactoglobulin is very distinct in such an ability, as documented in Table 252-1. Several biologically important ligands in this table are the free fatty acids (FFA), which are prevalent in milk, retinol and hemin. FFA may be bound to ß-lactoglobulin in a manner similar to sodium dodecyl sulfate (SDS) [Spector and Fletcher, 1970]. FFA reversibly bind at a single site per monomer with a binding constant of $\sim 10^5$. Other biologically important ligands, retinol, hemin and protoporphyrin IX, all bind to bovine BLG with relatively high affinity. Besides these physiologically important hydrophobic molecules, a variety of small hydrophobic molecules can be bound to ß-lactoglobulin with lower affinities. This suggests that ß-lactoglobulin may have evolved as a non-specific binder of physiologically important hydrophobic molecules.

Among all the binding constants, the one of retinol is distinctive. The corresponding affinity for RBP, which is a specific retinol carrier *in vivo*, is only 5.9×10^6, i.e. an order of magnitude smaller than that of bovine BLG. This suggests that bovine ß-lactoglobulin's biological function may be involved in transport of retinol more so than of other small hydrophobic molecules. Retinol, see Figure 252-1, has long been known to be important to vision [Blomhoff *et al.*, 1990]. The influence

of retinol and its derivatives on the regulation of cell growth and survival has also been recognised more recently [Schubert and LaCorbiere, 1985; Schubert *et al.*, 1986]. Why does ß-lactoglobulin bind retinol far better than does RBP *in vitro*? Is that the same *in vivo*? Is retinol binding the only reason for the existence of ß-lactoglobulin? These questions may be relevant to the functional puzzle of ß-lactoglobulin.

Table 252-1: Ligands binding to bovine ß-Lactoglobulin[a]

Ligand	Binding constant	Reference
Retinol	5.0×10^{7}	[Fugate, 1980]
ß-Ionone	1.47×10^{6}	[Dufour, 1991]
Stearate[b]	1.7×10^{5}	[Spector, 1970]
Palmitate[b]	6.8×10^{5}	[Spector, 1970]
SDS	3.1×10^{5}	[O'Neill, 1987]
Hemin	4.0×10^{6}	[Dufour, 1991]
Protoporphyrin IX	2.5×10^{6}	[Dufour, 1991]
Toluene	4.5×10^{2}	[Robillard, 1972a]
p-Nitrophenol	1.9×10^{5}	[Farrell, 1987]
2-Heptanone	2.0×10^{2}	[O'Neill, 1987]
Laurate[b]	0.5×10^{5}	[Spector, 1970]
Oleate[b]	0.4×10^{5}	[Spector, 1970]
Butane	1.7×10^{3}	[Wishnia, 1966]
Pentane	7.1×10^{3}	[Wishnia, 1966]
Iodobutane	2.8×10^{3}	[Wishnia, 1966]
2,6-MANS	3.4×10^{5}	[Lovrien, 1969]
Methyl orange	0.2×10^{4}	[O'Neill, 1987]
n-Octylbenzene-p-sulphonate	6.3×10^{4}	[O'Neill, 1987]
p-Nitrophenol	1.9×10^{4}	[Farrell, 1987]
p-Nitrophenylacetate	3.0×10^{4}	[Farrell, 1987]
p-Nitrophenyl-ß-glucuronide	1.6×10^{4}	[Farrell, 1987]
p-Nitrophenyl sulphate	2.0×10^{3}	[Farrell, 1987]
p-Nitrophenyl pyridoxal phosphate	3.1×10^{3}	[Farrell, 1987]
2-Octanone	0.5×10^{3}	[O'Neill, 1987]
2-Nonanone	2.4×10^{3}	[O'Neill, 1987]
Trifluorotoluene	4.2×10^{2}	[Robillard, 1972b]
Hexafluorobenzene	1.6×10^{3}	[Robillard, 1972a]

[a] Hambling *et al.*, 1994

[b] taking into account the primary binding site only

There are two putative models for the binding of retinol by ß-lactoglobulin. The first model was based on the 2.8 Å tertiary model of ß-lactoglobulin [Papiz *et al.*, 1986], which revealed the folding

similarity between ß-lactoglobulin and retinol-binding protein. By direct analogy to the way in which RBP binds retinol, the retinol was fitted in the centre of the BLG calyx, the position that RBP accepts its substrate, see Figure 654-2. This binding mode involves Trp 19 as a key residue to host the ß-ionone moiety of retinol, which by itself binds to ß-lactoglobulin with high affinity. Site-directed mutations of ß-lactoglobulin [Cho *et al.*, 1994; Katakura *et al.*, 1994] have been made and studied to test this model. The results of the site mutations W19A [Cho *et al.*, 1994] and W19Y [Katakura *et al.*, 1994] on ß-lactoglobulin are different. The site mutation W19A indicated that Trp 19 is quite important to bovine ß-lactoglobulin's retinol-binding ability, while the site mutation W19Y shows only a slight change in the dissociation constant. With the more reliable ß-lactoglobulin model, which is described in chapter 5, it is seen that accessibility of the ß-ionone moiety to Trp19 is blocked by the side chain of two aliphatic amino acid residues.

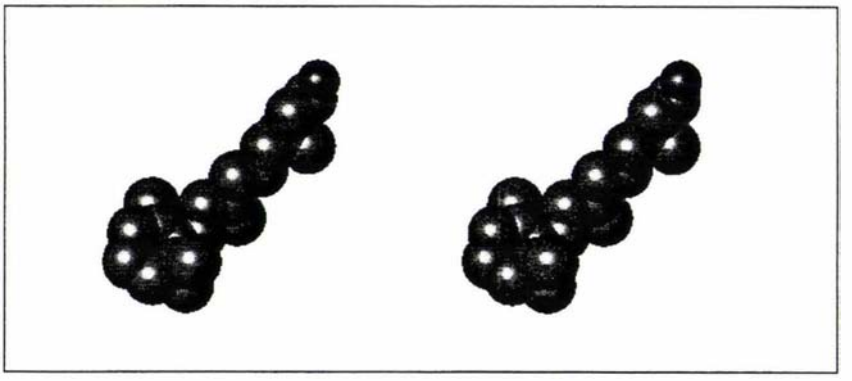


Figure 252-1: Conformation of all trans retinol

Other experimental data suggest that the binding manner of bovine ß-lactoglobulin to retinol may be different from that of RBP. First, alcohol dehydrogenase can rapidly degrade retinol bound to bovine ß-lactoglobulin at pH 7.4, but has no effect on retinol bound to RBP [Futterman and Heller, 1972]. Second, retinol bound to bovine ß-lactoglobulin can be replaced by other ligands. This phenomenon does not happen in the case of RBP [Hemley *et al.*, 1979]. Third, measurements of the rotational relaxation time suggest that retinol binds in a region of bovine ß-lactoglobulin that is flexible with respect to the bulk of the bovine ß-lactoglobulin molecule [Fugate and Song, 1980]. Therefore, Monaco *et al.* (1987) presented an alternative retinol-binding model for ß-lactoglobulin based on a 2.5 Å structure. This model involves two key residues Phe136 and Lys141 on the outside of the calyx. Site mutation

experiments, F136A and K141M, do not support this idea [Cho *et al.*, 1994]. Due to the errors in Monaco *et al.*'s ß-lactoglobulin model, there is no substantive evidence that retinol binds in this manner. The reasons are discussed in full in Chapter 6.

2.5.3 Biochemical properties and potential physiological function

ß-Lactoglobulin is a good nutritious component in milk. All essential amino acids are included in ß-lactoglobulin, and it is also a major component of whey protein. However, nutritious purpose alone does not explain its special fold, acidic stability and resistance to gastric proteolysis [Futterman and Heller, 1972; Hemley *et al.*, 1979]. Other profound functions may be responsible for the existence of ß-lactoglobulin. As a member of the lipocalin super-family, and deduced directly from its ability to bind a variety of hydrophobic substrates, the classic explanation for the function of ß-lactoglobulin is that it protects and transports small hydrophobic molecules *in vivo*. A more sophisticated physiological function of ß-lactoglobulin may involve the transfer of a passive immunity [Fugate and Song, 1980].

The existence of a receptor for the holo ß-lactoglobulin-retinol complex in the intestines of new-born calves supports the idea that bovine ß-lactoglobulin is involved in the transport of retinol [Papiz *et al.*, 1986]. The ^{125}I-labelled bovine ß-lactoglobulin-retinol complex can be bound to the microvilli prepared from the lower segment of the intestinal tract of one-week old calves. That the bound radioactivity can be reduced 70% by the unlabelled ß-lactoglobulin-retinol complex suggests that the binding is quite specific.

2.5.4 Gelation of ß-lactoglobulin

If the protein concentration is high enough, around 10%, pH and salt concentrations are suitable, ß-lactoglobulin can form heat-induced gels. Above pH 6 and in the absence of calcium, such gels are moderately tough and transparent [Mulvihill and Kinsella, 1987]. The mechanism of gelation has been investigated by several authors. Cho *et al.* (1994) have applied genetic engineering techniques to introduce cysteine residues at positions 104 or 132, which are close in space to the free thiol moiety

of Cys121. On oxidation, one more disulfide bond is formed in ß-lactoglobulin. This modified ß-lactoglobulin lacks the ability to form gels. Moreover, the introduction of additional thiols at other sites, either chemically [Kim *et al.*, 1990] or by genetic engineering [Lee *et al.*, 1993] enhances gelation. These experimental data stress the role of the free thiol group of ß-lactoglobulin in gelation.

2.5.5 Resistance to gastric environment

Bovine ß-lactoglobulin is resistant to hydrolysis *in vitro* by proteases such as pepsin, trypsin, and chymotrypsin [Reddy, 1988]. Resistance to proteolysis is diminished after cleavage of disulfide bonds. At low pH, for example 2.0, Kella and Kinsella (1988) concluded that the increased internal hydrogen bonds will increase the resistance to peptic digestion by pepsin. In contrast, ß-lactoglobulin is less resistant to trypsin, and chymotrypsin, which require a basic pH environment. This may result from the fact that ß-lactoglobulin undergoes a conformational change at pH 7.5, which may expose the strategic cleavage points for these enzymes [Chobert *et al.*, 1991; Reddy *et al.*, 1988]. Similarly, ß-lactoglobulin is quite resistant to gastric digestion *in vivo* and is reported to remain mostly intact after it passes through the stomach [Yvon *et al.*, 1984]. This can be evidenced by the absorption of some intact bovine ß-lactoglobulin by humans, as this protein can be detected in trace amounts in maternal milk from women who have consumed milk from cows [Monti *et al.*, 1989].

2.5.6 Effects of bovine BLG phenotype on physical properties

The composition of milk varies and is related to the ß-lactoglobulin phenotype of individual cows. Aschaffenburg and Drewry (1955) noted that the A variant of ß-lactoglobulin was associated with a higher yield of the protein in the milk. It has also been suggested that the yields of non-protein components of milk are correlated with the ß-lactoglobulin variant [Jakob and Puhan, 1992; Jakob *et al.*, 1995; Jakob, 1994; Mclean *et al.*, 1987]. In one large-scale investigation, Hill *et al.* (1994) reported that the variant A of bovine ß-lactoglobulin was associated with a 25-30% increase in the concentration of bovine BLG in milk and was also associated with a statistically significant decrease in both fat and

total solid contents. On the other hand, the milk produced by cows with the pure ß-lactoglobulin BB phenotype contains more fat.

The difference in primary structure of ß-lactoglobulin variants results in different charge properties. For A, B and C the isoelectric points in 0.1 M KCl are 5.26, 5.34 and 5.33 respectively [McKenzie, 1971]. The titration curves are indistinguishable between pH 2 and pH 4.5. At higher pH, the curves diverge at pH 7.0 and above. The A and C variants, respectively, carry one more and one less negative charge than the B variant. However, all variants undergo the so-called Tanford transition and all contain an "anomalous" carboxyl group that has an apparent pKa of 7.3 instead of ~4.7.

Different phenotypes are reported to have different self-association reactions. The A variant forms both dimers and octamers under appropriate conditions, whereas the B and C variants do not form octamers [Timasheff, 1966], but form dimers more readily than the variant A [McKenzie, 1972]. The self-association of ß-lactoglobulin has been studied by several groups using ultracentrifugation and light-scattering techniques. The general picture is that the A, B and C variants all form dimers in the pH range form 3.5 to 6.5. Within this range (pH ~3.5 - 5.1), the variant A dimer tetramerizes, particularly at low temperatures. The relative ease of dimer dissociation has been measured for the ß-lactoglobulin variants. At pH 7.5, 20.5 °C and 0.1 M ionic strength, K_d was < 5 µM for variant C, K_d ~ 8 µM for variant B, and K_d ~ 60 µM for variant A. At pH 2.6, K_d is ~50 µM for variant B and ~130 µM for variant A [Thresher *et al.*, 1994].

2.5.7 Interaction of ß-lactoglobulin with proteins

The interaction of ß-lactoglobulin with the other milk proteins has been investigated extensively. ß-Lactoglobulin is known to interact with α-lactalbumin [Hunziker and Tarassuk, 1965], and several caseins. For example, it reacts with κ-casein to form a 3:1 complex. This complex is based on hydrophobic interactions and is further stabilized by covalent bonding. The conformational changes make the disulphide bridges less susceptible to attack [Haque *et al.*, 1987].

ß-Lactoglobulin can interact with some non-milk proteins, for example,

cytochrome-c. Cytochrome-c binds to form a 1:1 complex. Above pH 7.5, the conformation of bovine ß-lactoglobulin in the complex is altered, enabling the free sulfydryl of BLG to reduce the iron in cytochrome-c [Brown and Farrell, 1978]. If porcine ß-lactoglobulin is used, only non-reduced cytochrome-c is found in the complex. It is not clear which residues are involved in the interaction. The bovine ß-lactoglobulin variants A, B and C all form complexes with cytochrome-c but with different affinities.

Screening ß-lactoglobulin from different milk sources results in understanding about the interaction of ß-lactoglobulin with immunoglobulins. Antisera against bovine ß-lactoglobulin cross-react with ruminant ß-lactoglobulins to different extents [Jenness *et al.*, 1967], but evidence is equivocal for cross-reactions also occurring with the monomeric ß-lactoglobulins from pig [Liberatori *et al.*, 1979c], horse [Liberatori *et al.*, 1979c] and camel [Liberatori *et al.*, 1979b]. Antibodies towards bovine BLG have been shown to cross-react with human milk. Removal of bovine milk protein from the mother's diet caused the cross-reaction to fall to non-detectable amounts. It appears that this is due to the tranfer of bovine BLG or antigenic peptides from the mother's diet to human milk [Axelsson *et al.*, 1986; Jakobsson *et al.*, 1985; Jakobsson *et al.*, 1985]. Monoclonal antibodies towards bovine ß-lactoglobulin have been used to investigate the denaturation/ renaturation process of ß-lactoglobulin [Hattori *et al.*, 1993].

2.6 Crystal forms

Since the 1950s, X-ray diffraction techniques have become the standard method to determine protein structure. Bovine ß-lactoglobulin crystallises readily, and was the subject of X-ray diffraction experiments as early as 1938. Long before the appearance of phase-determination methods for protein crystals, two forms of ß-lactoglobulin crystals had already been identified by Crowfoot and Riley (1938). One of crystal forms had the space group $P2_12_12_1$ with unit cell dimensions close to lattices now identified as Q or N [Aschaffenburg *et al.*, 1965]. The unit cell dimensions measured by Crowfoot and Riley were orthorhombic with a=b=63.5 Å and c=145 Å. The density of this crystal form (1.257 g/cm^3) was obtained by flotation in sugar solutions. Applying the upper

limit for density of 1.31 g/cm^3 which had been deduced for crystalline insulin, the authors estimated the molecular weight for ß-lactoglobulin was 35,300-36,500 dalton, reasonably close to twice the number calculated from the sequence of bovine BLG, consistent with a dimer.

Repeated nearly three decades after Crowfoot and Riley's pioneering study, Aschaffenburg *et al.* published the parameters (1965) for a number of ß-lactoglobulin crystal forms. These crystals were formed over a pH range from 3.5 to 7.6. All seven crystal systems except cubic are represented, as shown in Table 26-1. The subsequent structure determinations have involved primarily lattices X, Y and Z.

Table 26-1: Crystal forms of ß-lactoglobulin

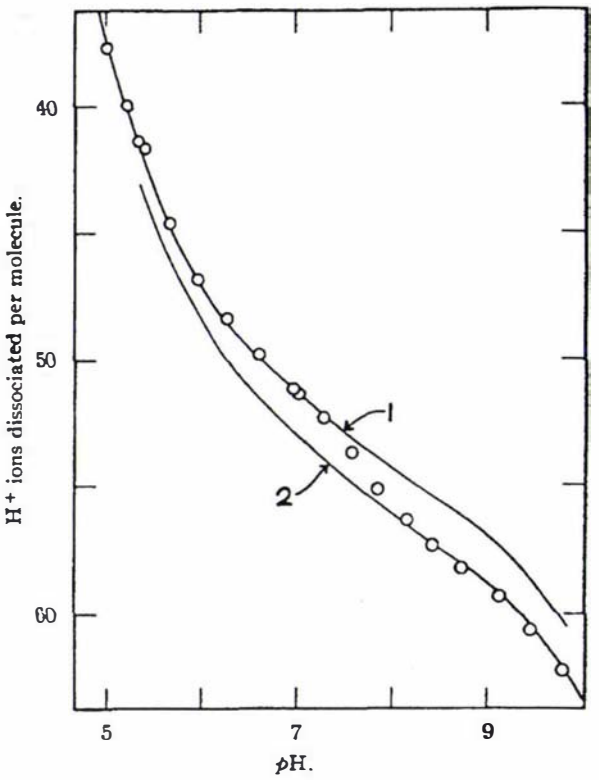
Variant	pH	type code	space group	a (Å)	b (Å)	c (Å)	α	ß	γ	Volume (10^5 Å^3)
?	3.5	P	$P6_3$	67	67	141				5.46
B	5.2	R	$P2_1$	36.1	127.5	36.0		106°05'		1.64
B	5.2	S	$P2_1$	36.4	127.6	36.4		98°12'		1.68
A	5.2	W	$P2_1$	36.4	68.2	72.4		92°12'		1.80
A	5.2	N	$P2_12_12_1$	68.6	70.2	137.8				6.63
A/B	5.2	Q	$P2_12_12_1$	69.3	70.7	157.5				7.74
B	5.2	T	$P4_12_12$	69.2	69.2	138.8				6.65
?	5.2	U	$P4_22_12$	67.5	67.5	133.5				6.05
A	6.9	X	P1	31.8	49.7	56.6	122°45'	97°31'	104°05'	0.83
A/B	7.6	Y	$C222_1$	55.7	81.2	67.2				3.04
A/B	7.6	Z	$P3_12_1$	54.3	54.3	117.0				2.99

2.7 Conformational changes of ß-lactoglobulin in solution

2.7.1 pH-Induced reversible conformational change

Several BLG pioneer researchers [Cannan *et al.*, 1941; Tanford *et al.*, 1959; Nozaki *et al.*, 1959; Brignon and Ribadeau-Dumas, 1969; Ghose *et al.*, 1968] monitored the titration behaviour of bovine ß-lactoglobulin as a function of pH. A reversible transition of ß-lactoglobulin was found to occur at around pH 7.5 (see Figure 271-1). McKenzie and Sawyer (1967) reported the reversible transition in specific variants by monitoring the optical rotation. The different variants have demonstrated different

behaviour, see Figure 271-2. Together with information from circular dichroism [Townend *et al.*, 1967] and optical rotatory dispersion [McKenzie and Sawyer, 1967], bovine ß-lactoglobulin is believed to undergo three pH-dependent conformational transitions which can be

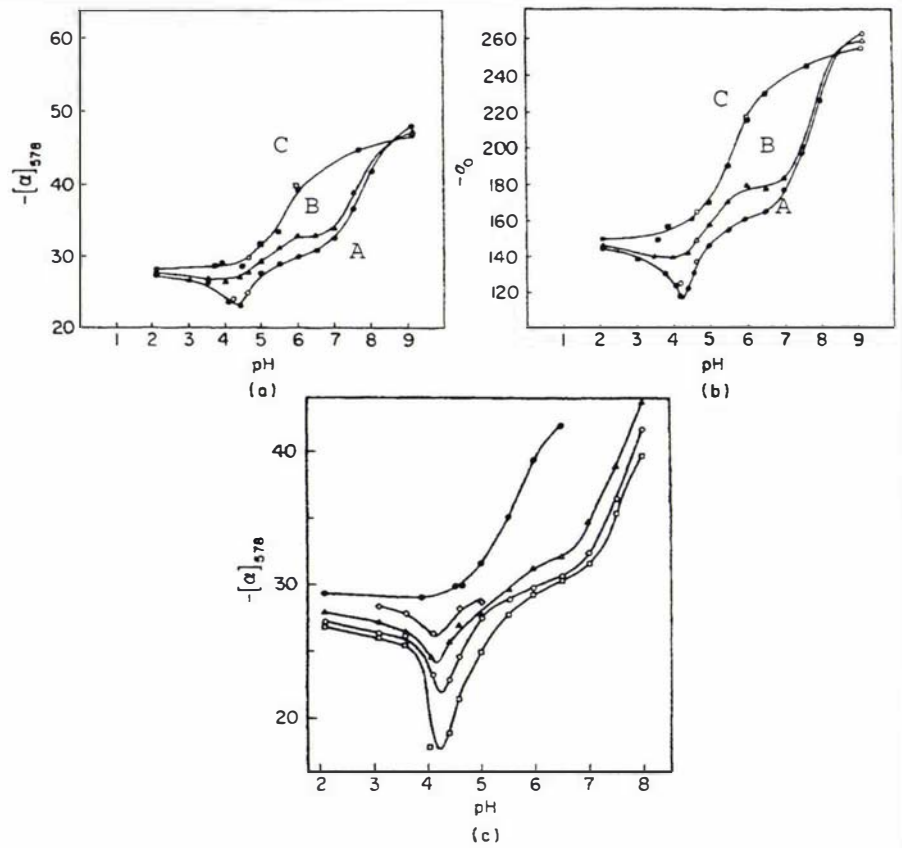


-The neutral region of the titration curve at ionic strength 0.15. Curves 1 and 2 are calculated curves, differing only in the number of carboxyl groups assumed available for titration (see text). Almost all of the experimental points are superpositions of points from forward and reverse titration curves, at identical pH values. The agreement between them was always 0.1 group titrated or better.

Figure 271-1: Titration behavior of a BLGA and BLGB mixture, [Tanford *et al.*, 1959].

summarised as: Q ⇔ N ⇔ R ⇒ S. The first transition between the Q state and the N state occurs from pH 4 to 6 and is reversible; the protein shrinks slightly in volume and the sedimentation coefficient increases a little. From pH 6 to 8, bovine ß-lactoglobulin undergoes a second reversible transition N ⇔ R, also called the Tanford transition. Above pH 9, the protein molecule undergoes irreversible conformational change.

The Tanford transition between the N state and the R state is one of the most extensively investigated biochemical properties of bovine ß-lactoglobulin. Centrifugation experiments by Pedersen (1936) revealed



Conformational transitions dependent on pH for bovine β-lactoglobulins A, B and C. (a) The specific rotation at 578 nm ($[\alpha]_{578}$) at 20°C. (b) The parameter, α_0, of the Moffitt-Yang equation (see Volume I, Chapter 8, Section IX.H), at 20°C. Symbols for (a) and (b): ●, A (variant); ▲, B; ■, C; solid symbols, direct pH titration; open symbols, buffered solutions (from McKenzie and Sawyer, 1967). (c) The effect of temperature on $[\alpha]_{578}$ for the A variant near pH 4.5. The C variant is shown for comparison. For β-C: ●, 20°C. For β-A: ◇, 45°C; △, 30°C; ○, 20°C; and □, 10°C. (From McKenzie *et al.*, 1967.)

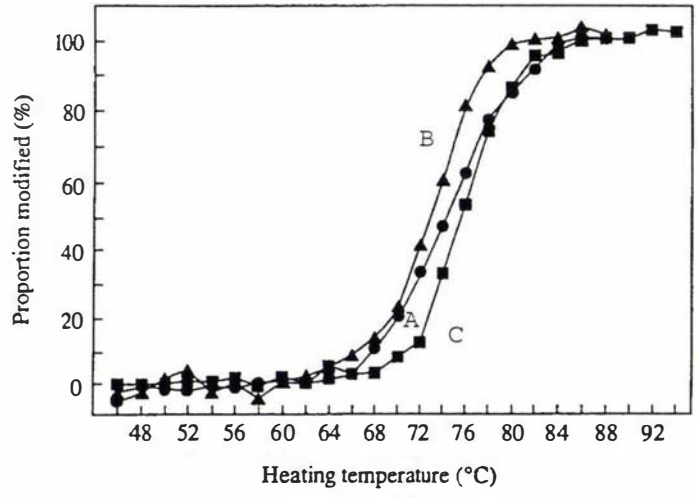
Figure 271-2: Conformational changes of bovine ß-lactoglobulin monitored by optical rotation [McKenzie *et al.*, 1967].

that the sedimentation coefficients of BLG decreased with increasing pH in the range of pH 6 to pH 8. Groves *et al.* (1951) reported an increase in optical laevorotation of BLG with increasing pH; McKenzie and Sawyer (1967) extended this study to specific BLG variants. Tanford *et al.* (1959) found different titration behaviour for BLG in acidic and basic environments as mentioned above. It was proposed that this difference arose from exposure of a buried COOH group at basic pH; otherwise, this COOH group can not be accessed during titration at acidic pH. Qi *et al.* (1995) observed different thermal behaviour of bovine BLG at different pH: a peak in the thermalgram present pH 6.75, disappears at pH 8.05. Another pH-related phenomenon, which has been described by Dunnill and

Green (1965), is that the free sulfydryl group of BLG can react with reagents containing Hg^{2+} much more readily above pH 6.7. Collectively, those pH-dependent changes in physical, chemical and spectroscopic properties of BLG are known as the Tanford transition.

2.7.2 Heat-induced conformational change

Heating will increase the entropy of protein molecules, make the molecule more dynamic, and result in some new conformations beyond the energy barriers. Casal *et al.* observed (1988) that bovine BLGB undergoes multiple temperature-dependent comformational changes from -100 °C to +90 °C. One change, which occurs abruptly between 58 and 60 °C, leads to the same structure as that found in the alkaline denaturation process of bovine BLGB, as evidenced by Fourier-transform infra-red spectroscopy.



Denaturation of β-lactoglobulin A, B and C in phosphate buffer at pH 6.7. Samples (3 mg/ml) were heated rapidly, held for 12.5 min, and cooled rapidly. The circular dichroism signal at 293 nm was determined and the results were normalized. Symbols: ●, β-lactoglobulin A; ▲, β-lactoglobulin B; ■, β-lactoglobulin C.

Figure 272-1: Thermal denaturation of bovine BLGA, BLGB and BLGC in phosphate buffer at pH 6.7 [Manderson *et al.*, 1995].

The three most common phenotypes of bovine ß-lactoglobulin have slightly different heat-denaturation profiles [Manderson *et al.*, 1995]; see Figure 272-1. The denaturation process of bovine BLGC happens at a higher temperature than bovine BLGB. The denaturation curve for bovine BLGA lies between those of variants B and C.

2.8 Structural knowledge

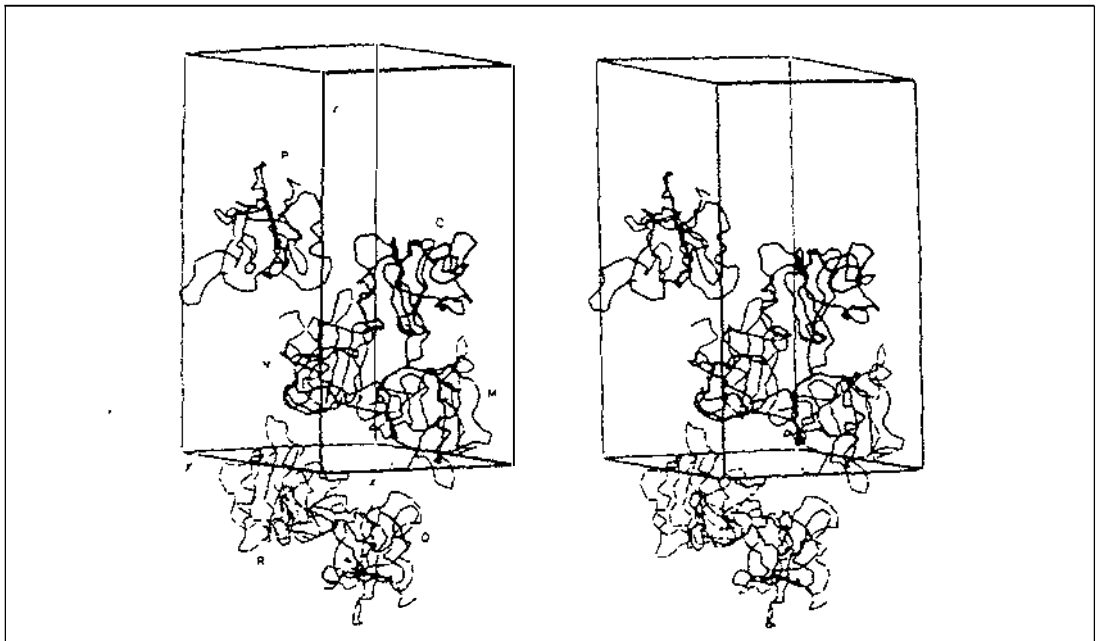
The structure determination of bovine ß-lactoglobulin was begun in the mid-1950s and reported at a low resolution of 6 Å at the end of the 1970s (Green *et al.*, 1979). Three native crystal forms of ß-lactoglobulin were involved: Lattices X,Y and Z. It is impossible to trace the secondary structure at such low resolution. But the electron density of lattice X revealed the existence of a molecular dyad in the dimer of bovine ß-lactoglobulin. This axis is not necessarily coincident with a crystallographic symmetry axis. They assumed that lattice X (pH 6.5), lattice Y (pH 7.5), and lattice Z (pH 8.1) corresponded to the conformational shift N ⇔ R with respect to the pH. Their reports showed that the free sulfydryl site in lattice X is a distance 3 Å away from that in the lattices Y and Z. This deduction was based on the heavy atom site which can be deduced relatively accurately even at low resolution.

A medium resolution model was reported first in 1986 for ß-lactoglobulin in the orthorhombic lattice Y. This 2.8 Å resolution model [Papiz *et al.*, 1986] demonstrated the conformational similarity of bovine BLG with that of retinol-binding protein (RBP). The molecule consists of an anti-parallel ß-sheet, formed by 9 strands wrapped round to form a flattened cone or calyx. A three-turn α-helix is located towards the C-terminal end. The authors suggested that the ninth strand, ß-strand I, is involved in the formation of the dimer by making anti-parallel interactions with a dyad-related strand. Although this model demonstrated the basic structure of ß-lactoglobulin, difficulties in refinement ensued.

A second medium resolution structure of bovine BLGA (2.5 Å) was obtained by molecular replacement by Monaco *et al.* (1987) in the trigonal lattice Z. Although the BLGA molecules in the Z or Y crystal forms showed no great difference in the ß-barrel topology, differences in several areas were reported. Monaco *et al.* (1987) reported that all residues were identified and that the two disulphide bonds were clearly visible. However, this medium ß-lactoglobulin model had serious problems. For example, it was refined without a R_{free} monitor, and the average rms deviation of bond lengths from their ideal values is very high at 0.08 Å. Moreover, the BLGA molecules are packed in monomers not dimers in the

crystal lattice, as shown in Figure 28-1, which is inconsistent with other observations in solution or in crystal structures [Townend *et al.*, 1969; Green *et al.*, 1979; Papiz *et al.*, 1986; Brownlow *et al.*, 1997; Qin *et al.*, 1998].

High resolution models for bovine BLG in lattice Y [Bewley *et al.*, 1998] and lattice X [Brownlow *et al.*, 1997] have also been determined. The high resolution model (1.8 Å) of BLGA in lattice Y (space group $C222_1$) contains residues 1-62, 66-110, 115-154 and 10 water molecules, R_f=27%, R=22%. One disulfide bridge between Cys106 and Cys119 is clearly visible, but the other disulfide bond is poorly defined. Bovine ß-lactoglobulin in lattice X was also determined at high resolution (1.9 Å). The C and N-termini are not visible and the loop CD is not well defined according to the authors [Brownlow *et al.*, 1997]. These two structures convincingly correct the errors of Monaco's model. Lack of dimeric packing and mis-threading of the polypeptide chain in Monaco's model are the key points of criticism.



Crystal packing of BLG in the trigonal cell. Molecules are related by the following symmetry relations: M (x, y, z); N ($x-y$, $1-y$, $1/3-z$); O ($1+y-x$, $1-x$, $1/3+z$); P ($1-x$, $1+y-x$, $2/3-z$); Q ($1-y$, $x-y$, $z-1/3$); R (y, x, $-z$). Molecules Q and R are translated by a unit cell parameter along z to show the contacts around the central molecule.

Figure 28-1: Bovine BLGA packed in lattice Z, defined by Monaco *et al.* (1987).

2.9 Goals of the project

A number of features of bovine ß-lactoglobulin have been described in previous pages, ranging from the biological properties to biochemical characteristics. This protein is a dominant component of milk of cow, but we do not know yet its actual biological function. It is able to bind a variety of hydrophobic molecules, but we do not know their actual binding site(s). It has pH-dependent conformational changes, but we do not know how its conformation is changed. The different variants of bovine ß-lactoglobulin have a substantial bearing on the processing of milk [Hill *et al.*, 1996], and for this reason are an important focus of research in the development of New Zealand's dairy industries. The previous structural determinations have revealed the lipocalin fold of bovine ß-lactoglobulin, but offered little to help understand these important aspects of bovine ß-lactoglobulin.

While many valid purposes for this project can be summarised today, in the initial stages, the purpose of this project was quite vague to me. The project might just have served as a training case, because the high resolution structure of BLGA in lattice Y was almost completed. I could do little to push this project ahead. However, after the structure for BLGB in lattice Y was determined including all residues, this training case seemed likely to convert to a research project. Ted Baker expressed interest in understanding the structure of BLG at different pH, and suggested a possible condition for preparing crystals of BLG at low pH for lattice X. This was the immediate reason I designed the crystal screen of BLG in a 4 × 6 matrix, the 4 serves to vary the pH, and 6 to vary the ammonium sulfate concentration.

Unfortunately, I have never obtained any BLG crystals in lattice X, or even in lattice Y. All the structures I have determined are based on lattice Z. At this stage, what I wanted was to determine the structure of bovine ß-lactoglobulin in full. I achieved this purpose first in the structure of BLGA in lattice Z at pH 8.2, and then in the structures of BLGA in lattice Z at pH 7.1 and 6.2.

In about March of 1997, I knew that one loop in BLGA varied according to pH. I did not understand the meaning of the conformational change of this

loop until August of 1997. The Tanford transition was also recognised as an old story to which I had never paid attention. One day when I packed up the rubbish paper on my desk, a printed email made a remark about Glu89, and indicated it was a buried residue. This conflicted with my memory: in my memory only Glu108 is a buried glutamate residue in the bovine BLG structures. I checked the bovine BLG structures in lattice Z again, I was sure Glu89 was not buried at pH 7.1 and 8.2, but buried at pH 6.2. Later, Lindsay Sawyer's paper made a tentative suggestion about the role of Glu89 in the Tanford transition. After carefully reading the publication of Tanford *et al.* (1959), I thought I knew the key to the Tanford transition and announced my conclusion in a group meeting in August of 1997.

The structure of BLGB in lattice Z at pH 7.1 was determined later. The difference in properties and structures of BLG variants has been a topic of interest for a long time by researchers of the New Zealand Dairy Research Institute. So, the structural consequences of point mutations of bovine BLG became one of the foci of later research, especially as the lattice Y structures were poorly defined in the region of one of the point mutations.

The crystal structure of bovine BLG with bound ligand was one of the experimental goals at an early stage. SDS, retinol, hemin and protoporphyrin IX were tested, and no good results were obtained. Crystals grown in the presence of SDS diffracted to a low resolution ~5.5 Å. Crystals grown in the presence of retinol were very small. The slightly red-colored crystals were obtained in the presence of hemin, but no hemin could be found. I ran out of sample for a half year and was not interested in these ligand-binding trials. I thought I had enough material for my thesis.

However, Geoff Jameson insisted that further investigation of some new samples would be desirable, when I modified my first version of thesis. These ligand-BLG samples came from NZDRI, ~ 20 mg of BLG-retinol and a similar amount of BLG-BrC12. I immediately set up the modified crystal screen for the fresh ligand-BLG samples. A few days later the crystals grew. After one week waiting for X-ray facilities, I collected the data set, again lattice Z, resolution very good at 2.23 Å —— not hopeful

signs for ligand existing in the crystal. However, the density map revealed the existence of the ligand BrC12, 12-bromododecanoic acid. It was less than 20 days from the time we obtained the ligand-BLG samples to the time we calculated the density map. This was a pleasant surprise in our BLG research activities.

The BLG can be crystallized into several crystal forms. Establishing the differences of BLG in lattices X, Y and Z was one of the early goals; this goal has now been achieved.

Chapter 3 Principles

3.1 Principles of single-crystal X-ray diffraction

When electromagnetic radiation, for example, X-rays, impinges on matter, the electric and magnetic components of the radiation cause the electrons in the matter to oscillate with the same frequency as the incident radiation. The oscillating electron acts as a radiation scatterer and emits radiation of the same frequency as the incident X-ray radiation. In any system, the scattering of the incident X-ray beam is the summation of the scattering from all the electrons. The diffraction of X-rays from single crystals is analysed below by first summarizing the scattering from a single electron, then from an atom, a unit cell, and finally from a crystal.

To achieve the summation of all the scattering waves of electrons, the phases of all the individual scattering waves from single electrons must be calibrated according to a reference scattering wave. Figure 31-1 shows a simple two-electron system. The scattering wave from the white electron

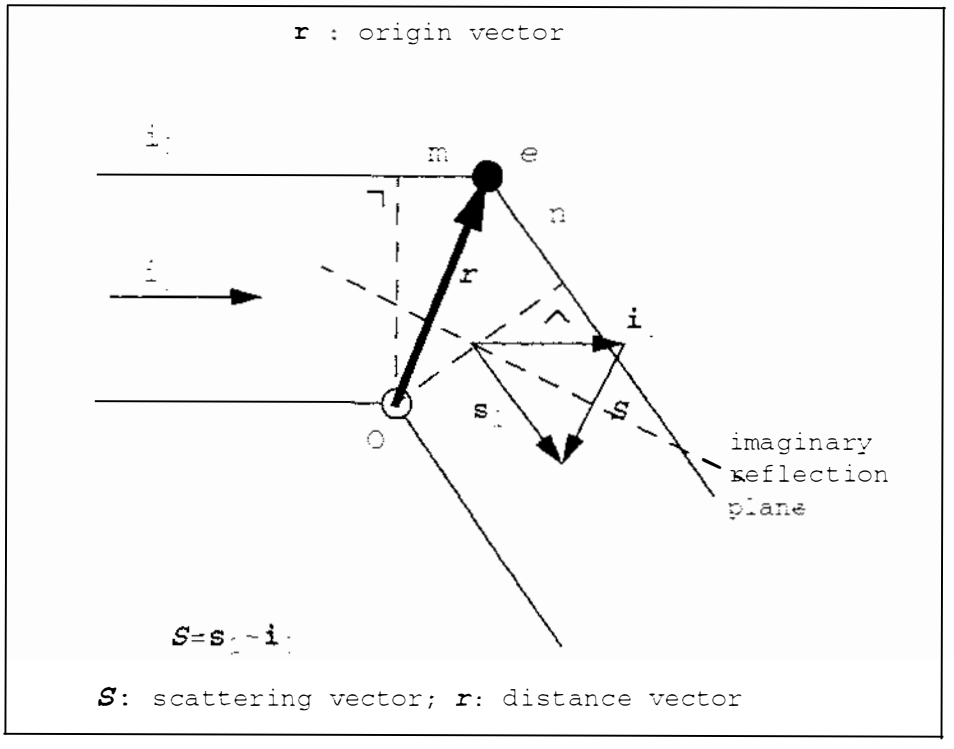


Figure 31-1: Scattering phase-shift in a two-electron system

is chosen as the reference scattering wave and referred to as O, the

origin. The black electron, e, which scatters the incident X-ray beam, i_0 $(1/\lambda)$. Now, the phase shift of the scattered beam from the black electron can be related to its path difference with the the scattered beam from the white electron as follows:

$$2\pi(m+n)/\lambda = 2\pi r \bullet S \tag{31-1}$$

Then, the scattered X-ray beam of the black electron relative to the origin can be expressed as:

$$e_s \exp\{2\pi i r \bullet S\} \tag{31-2}$$

e_s: Scattering ability of a single electron

For an atom, the scattering is achieved by the electron cloud around the nucleus, as shown in Figure 31-2. The scattering of X-rays by an atom requires the summation of scattering over the electron cloud. The nucleus is chosen as the origin. The scattering at site r has a phase relative to the origin:

$$\exp[2\pi i r \bullet S] \tag{31-3}$$

At site r, the electron density is $\rho(r)$, and the scattering ability electron cloud over this position is $\rho(r)dr$. Then the atomic scattering factor f is the integration over the volume of the electron cloud:

$$f = \int_r \rho(r) \exp[2\pi i r \bullet S]\, dr \tag{31-4}$$

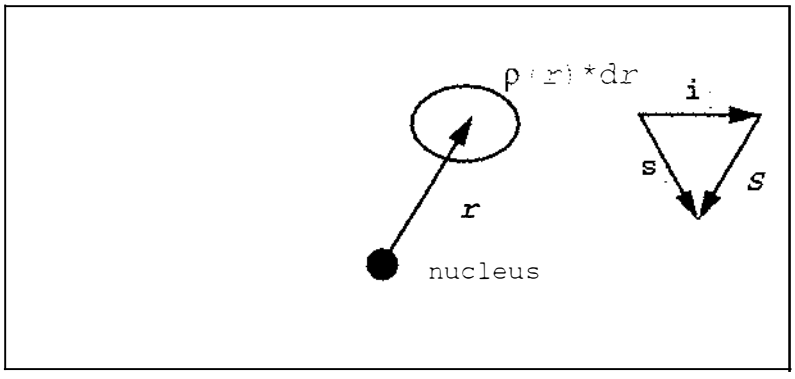


Figure 31-2: X-ray scattering by an atom

One unit cell of a crystal contains a number of atoms, as illustrated in Figure 31-3. To simplify the discussion, the origin is chosen as the

origin of the unit cell, point O. Similarly, the scattering wave from the jth atom in the unit cell has a phase shift relative to the origin of:

$$\exp\,[2\pi i r_j \bullet S] \tag{31-5}$$

Next, summation, $U(\boldsymbol{S})$, is made over the n atoms in the unit cell, each with atomic scattering factor $\boldsymbol{f}_j$:

$$U(\boldsymbol{S}) = \sum_{j}^{n} f_j \exp\,[2\pi i r_j \bullet S] \tag{31-6}$$

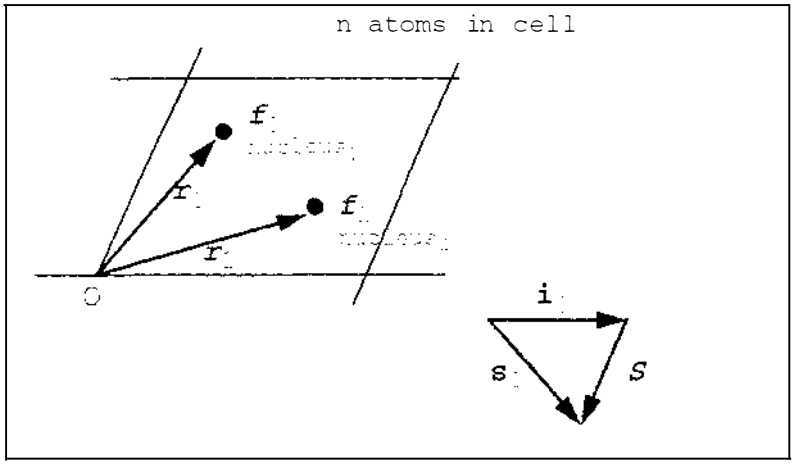


Figure 31-3: X-ray scattering by a unit cell

The crystal is the 3-dimensional translation lattice of the unit cell, as detailed in section 3.2. Scattering from a crystal is the result of scattering from effectively an infinite number of unit cells. Summation must be over all unit cells with phase shifts with respect to the crystal origin. To one individual unit cell, U_x, the shift of origin will introduce a phase shift to $U(\boldsymbol{S})$. Therefore, scattering from this unit cell is:

$$U_x(S) = U(S) \exp\,[2\pi i (ta + ub + vc) \bullet S] \tag{31-7}$$

where $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ are the translational vectors of crystal and t, u, v are integers which define the position of U_x in the cell.

The scattering $F(\boldsymbol{S})$ from the crystal is the summation of scattering over all the unit cells in the crystal. $U(\boldsymbol{S})$ is the same for each cell. Suppose each direction has l, m, n unit cells in the crystal, then:

$$F(S) = U(S) \sum_{t=0,u=0,v=0}^{l,m,n} \exp[2\pi i(ta + ub + vc) \bullet S] \qquad (31\text{-}8)$$

$$= U(S)\left(\sum_{t=0}^{l} \exp[2\pi ita \bullet S] \cdot \sum_{u=0}^{m} \exp[2\pi iub \bullet S] \cdot \sum_{v=0}^{n} \exp[2\pi ivc \bullet S]\right) \qquad (31\text{-}9)$$

$F(\boldsymbol{S})$ is called the structure factor.

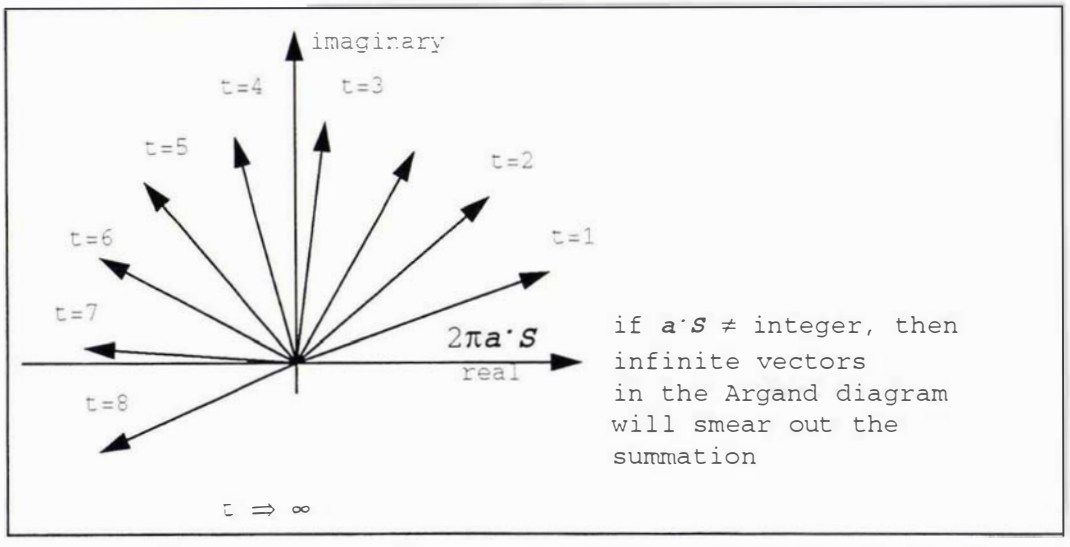


Figure 31-4: Reflection condition, illustrated in an Argand diagram

If $\boldsymbol{a}\cdot\boldsymbol{S}$ is not equal to an interger, the summation from 0 to ∞ for t will go to be zero (see Figure 31-4); similarly for $\boldsymbol{b}\cdot\boldsymbol{S}$ and $\boldsymbol{c}\cdot\boldsymbol{S}$. These are known as the Laue diffraction conditions. They specify the directions for diffraction from a crystal.

3.2 Relationship between crystal and diffraction pattern

A protein crystal is a three-dimensional periodic arrangement of protein molecules. The minimal repeating unit, the unit cell is characterized the vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$, with the angles α ($\boldsymbol{b}\wedge\boldsymbol{c}$), ß ($\boldsymbol{a}\wedge\boldsymbol{c}$) and γ ($\boldsymbol{a}\wedge\boldsymbol{b}$) between the vectors. The three-dimensional stack of unit cells defines the crystal lattice.

In last section, the Laue conditions for diffraction from a single crystal were stated, and are conventionally expressed as:

$$a \bullet S = h \qquad b \bullet S = k \qquad c \bullet S = l \qquad (32\text{-}1,2,3)$$

h, k, l are integers. The Laue conditions can be arranged as:

$$\frac{a}{h} \bullet S = 1 \qquad \frac{b}{k} \bullet S = 1 \qquad \frac{c}{l} \bullet S = 1 \qquad (32\text{-}3,4,5)$$

The scalar product of the two vectors $\boldsymbol{a}/h$ and $\boldsymbol{S}$ is equal to 1. In other words, the vector $\boldsymbol{a}/h$ is projected on to the vector $\boldsymbol{S}$ and multiplied with $|\boldsymbol{S}|$ gives unity. This is also true for b/k and c/l. Therefore, the endpoints of the vectors $\boldsymbol{a}/h$, $\boldsymbol{b}/k$, and $\boldsymbol{c}/l$ form a plane perpendicular to vector $\boldsymbol{S}$, as shown with a two-dimensional example in Figure 32-1.The lattice symmetry then generates a set of lattice planes, denoted (*hkl*), with equal interplanar spacing, denoted d_{hkl}.

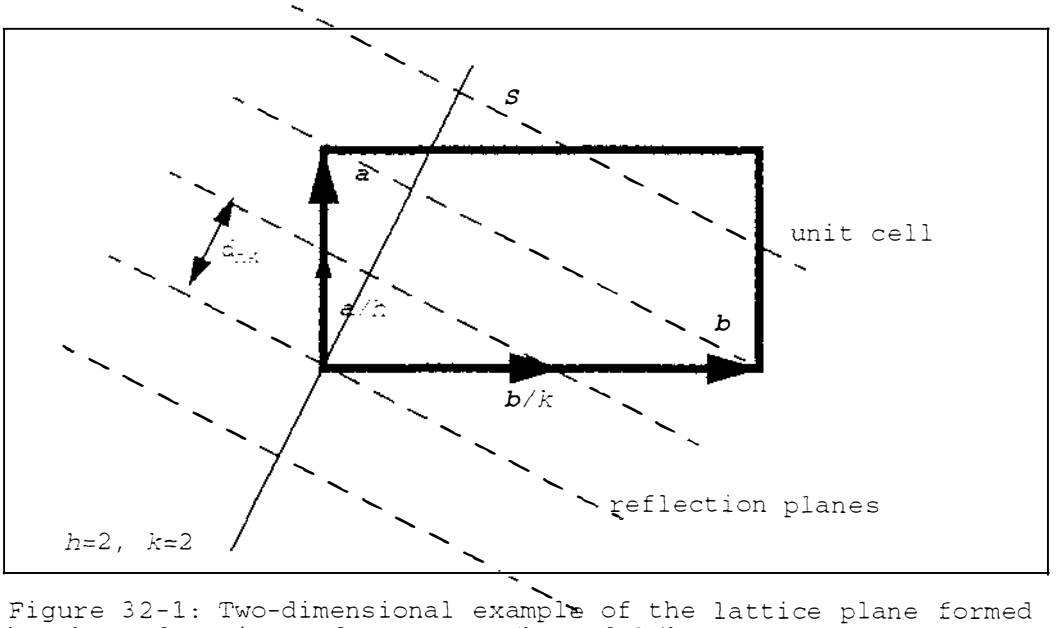


Figure 32-1: Two-dimensional example of the lattice plane formed by the end points of vectors $\boldsymbol{a}/h$, and $\boldsymbol{b}/k$.

In the X-ray diffraction experiment, this set of planes is conveniently treated as a set of imaginary *reflection* planes.

Within each unit cell, symmetry operations such as rotations, reflections, screw axes and glide planes may relate molecules to each other. The combination of lattice translations and symmetry operations constitutes the space group. The asymmetric unit is the minimum volume

of the unit cell or the minimum number of atoms or molecules from which the contents of the entire unit cell can be generated by application of the symmetry operations. The asymmetric unit of the unit cell can be part of a molecule, a molecule or more than one molecule. Bovine ß-lactoglobulin crystallises in a variety of crystal forms. Bovine BLG's lattice Z is a crystal form with space group $P3_121$. In this case, the unit cell has six asymmetric units which are symmetrically related; see Figure 32-2. Each asymmetric unit corresponds to a bovine ß-lactoglobulin monomer.

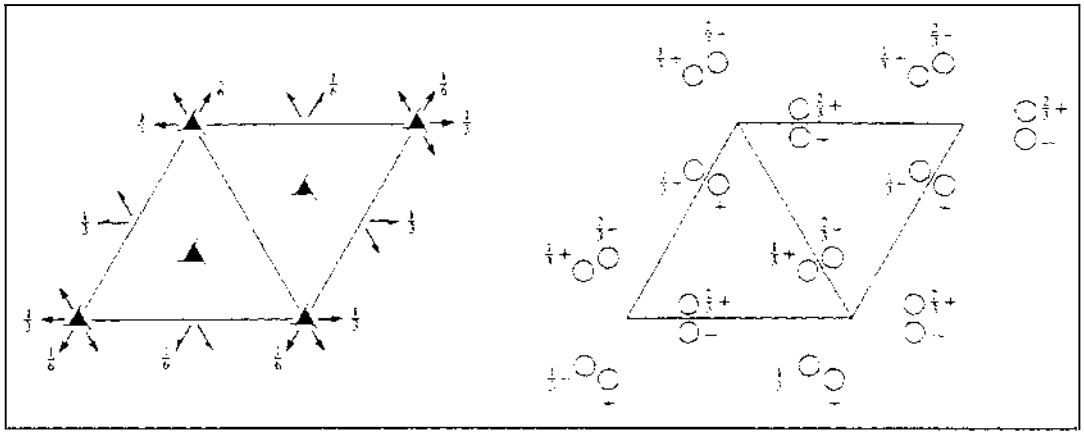


Figure 32-2: International Tables representation of space group $P3_121$ [Hahn, 1995].

We have summarised the crystal space also known as real space. The X-ray diffraction pattern defines reciprocal space or diffraction space. When the Laue conditions for diffraction are satisfied, the diffraction associated with integers *hkl* is loosely termed the reflection (*hkl*). The reflection (*hkl*), with associated structure factor F_{hkl}, is conveniently represented as a reflection of X-rays from lattice planes *hkl*. Each reflection generated by a set of imaginary reflection planes (*hkl*) is represented by a diffraction vector which has its end point, a distance $1/d_{hkl}$ from the origin. The set of all reflections defines the reciprocal lattice, which is represented by three vectors ***a*, b*, c**** [Barlow, 1897]. The Ewald construction [Ewald, 1921] effectively relates each reflection or diffraction spot to its reflection plane (*hkl*). When a reciprocal lattice point lies on the surface of the Ewald sphere, the Laue conditions for diffraction are satisfied, as shown in Figure 32-3. Technically, this is called the index of the diffraction.

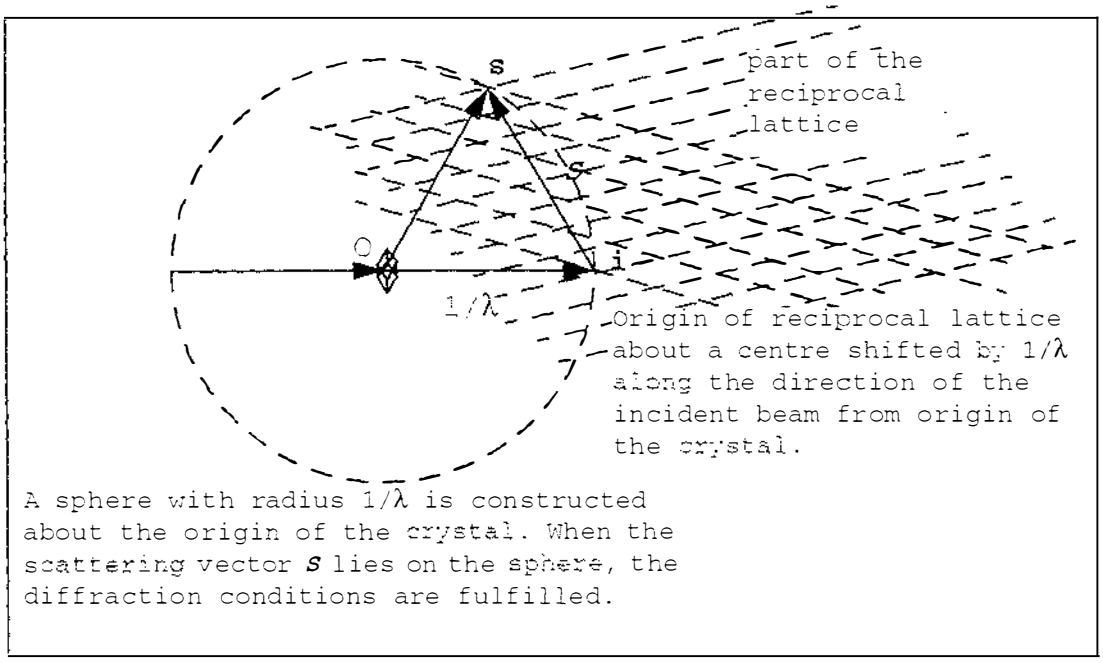


Figure 32-3: Ewald construction, showing conditions for diffraction

In the oscillation method of data collection, the crystal is rotated over a small angle, ~ 0.5 to ~ 1.5 °, about an axis perpendicular to the X-ray beam. The corresponding rotation of the reciprocal lattice brings more diffraction spots to the diffraction condition (i.e., cross the Ewald sphere). Careful choice of the oscillation range to avoid the overlap of diffraction spots on the recorded image plate allows each exposure to record more data. Because real crystals are not perfect, the diffraction conditions are fulfilled over a small angular range around the reciprocal lattice point, i.e. the reciprocal lattice point has volume. When the crystal is oscillated, it is possible that only part of the reciprocal lattice point fulfills the diffraction conditions. Such partial diffractions increase the difficulty in the indexing.

3.3 From diffraction data image to molecular structure image

3.3.1 Phase angle of diffraction

The indexed diffraction spots are not the final result of the X-ray diffraction experiment. What we want is the image of molecules, the building components of the crystal. The next thing is the reconstruction of the molecular image from the diffraction data. In section 3.1, the

summation equation (equation 31-6) of scatterings for a unit cell was made in terms of atoms. Instead of summing over all individual atoms, we can integrate over all electron clouds in the unit cell, similar to the manner of equation 31-4:

$$U(S) = \int_{cell} \rho(r) \exp[2\pi i r \bullet S]\, dv \qquad (33\text{-}1)$$

where:

$$dv = V dx dy dz \qquad (33\text{-}2)$$

$\rho(r)$: electron density at point r

V: volume of unit cell

and:

$$r \bullet S = (ax + by + cz) \bullet S \qquad (33\text{-}3)$$

$$= (a \bullet S)x + (b \bullet S)y + (c \bullet S)z \qquad (33\text{-}4)$$

x, y, z are fractional coordinates in the unit cell, which define a point with electron density $\rho(r)$ or $\rho(xyz)$. Applying the Laue diffraction conditions, equations 32-1,2,3, gives:

$$r \bullet S = hx + ky + lz \qquad (33\text{-}5)$$

Therefore, on replacing $U(S)$ by $F(hkl)$ and integrating over xyz, we obtain:

$$F(hkl) = \int_{xyz}^{111} \rho(xyz) \exp[2\pi i (hx + ky + lz)]\, dx dy dz \qquad (33\text{-}6)$$

This is the Fourier transform of $\rho(xyz)$ and the reverse is also true. Taking into account the Laue diffraction conditions that the diffraction is discrete, the continuous integration of $\rho(xyz)$ can be replaced by the summation of $F(hkl)$:

$$\rho(xyz) = \frac{1}{V} \sum_h \sum_k \sum_l F(hkl) \exp[-2\pi i (hx + ky + lz)] \qquad (33\text{-}7)$$

Because only the amplitude of the structure factor $F(hkl)$ is measurable, we replace F_{hkl} ($F(hkl)$) with $|F_{hkl}|\exp(i\alpha_{hkl})$ to get:

$$\rho(xyz) = \frac{1}{V}\sum_x\sum_y\sum_z |F_{hkl}| \exp\left[-2\pi i(hx+ky+lz) + i\alpha_{hkl}\right] \qquad (33\text{-}8)$$

$|F_{hkl}|$ is obtained directly as the square root of the intensity of diffraction beam; α_{hkl} is the phase of corresponding to diffraction beam, which can not be obtained directly from the diffraction pattern.

There are several ways to obtain the phase angles of reflections. The isomorphous replacement method requires the attachment of heavy atoms to the protein. The multiple wavelength anomalous diffraction methods depend on the presence of sufficiently strong anomalously scattering atoms in the protein structure itself. In the molecular replacement method the similarity of the unknown structure to an already known structure is a prerequisite. All three methods rely on the Patterson-function [Patterson, 1934; Patterson, 1935]

3.3.2 Patterson function

The Patterson function [Patterson, 1934; Patterson, 1935] is the convolution of the structure with itself, defined as:

$$P(u) = \int_r \rho(r) \times \rho(r+u)\, dv \qquad (33\text{-}9)$$

r, u are defined in Figure 332-1. By means of the convolution theorem, the equation 33-9 can be reformed as: [Drenth, 1995]

$$P(u) = \frac{1}{V}\sum_S |F(S)|^2 \cos\left[2\pi u \bullet S\right] \qquad (33\text{-}10)$$

Thus, the Patterson function can be calculated directly from the

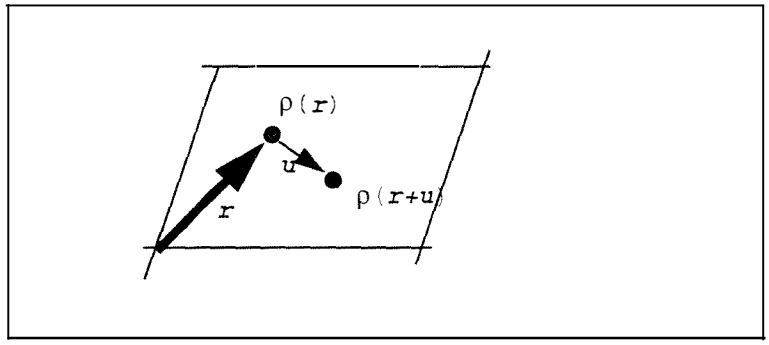


Figure 332-1: Vector relationships in a two-dimensional unit cell with two atoms

diffraction pattern. The physical significance of the Patterson function is that maxima in $P(\boldsymbol{u})$ correspond to vectors, $\boldsymbol{u}$, between atoms. Unless electron density is present at both $\rho(\boldsymbol{r})$ and $\rho(\boldsymbol{r}+\boldsymbol{u})$, the Patterson map has a value of zero at $P(\boldsymbol{u})$. Therefore, only vectors between pairs of atoms in the crystal structure show up as vectors from the origin to maxima in the Patterson map. Vectors between heavy atoms (eg, Hg...Hg) will dominate vectors between light atoms(eg, C, N, O) in the Patterson map due to high electron density. Thus, heavy atom ...heavy atom vectors between symmetry-related atoms can easily be identified and the positions of the heavy atoms in the crystal lattice can be deduced.

3.3.3 Isomorphous replacement methods

The protein crystal usually contains large water channels which allow the heavy atoms to travel and attach to the protein molecules. The attachment of heavy atoms often does not seriously alter the conformation of the protein and also leaves a reasonable isomorphism between the crystal unit cell parameters of the native protein and the heavy-atom derivative. The near-identity of structure between the native protein crystal and the heavy-atom protein crystal leads to the same basic diffraction pattern but with small differences in intensities between corresponding reflections. The phase angles of reflections for the native protein are calculated from these differences in intensities.

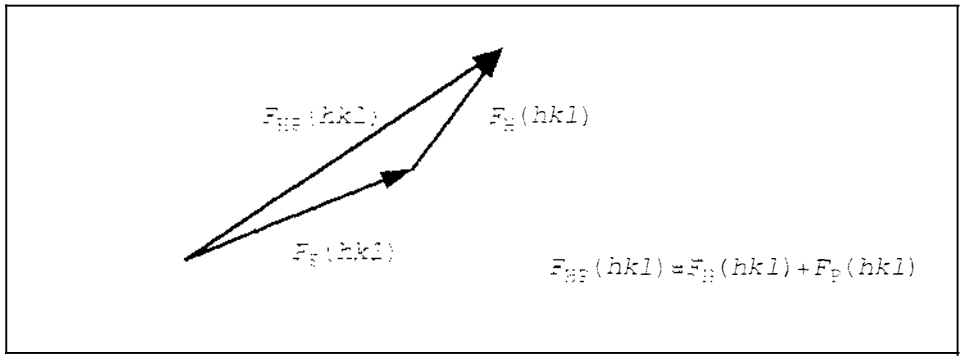


Figure 333-1: Change in structure factor when a heavy atom is attached to protein molecule.

The structure factor of the heavy-atom derivative ($F_{HP}(hkl)$) is the vector sum of the structure factor of the heavy atom ($F_H(hkl)$)and the structure factor of the protein ($F_P(hkl)$); see the Argand diagram, Figure 333-1. What we know from the diffraction patterns of both the native

protein crystal and the heavy atom derivative are the intensities which correspond to $|F_{HP}(hkl)|$, and $|F_P(hkl)|$.

The initial step of isomorphous replacement methods is to find out the position(s) of the attached heavy atoms. The attached heavy atoms are electron-rich moieties compared to the protein molecule, which is usually comprised of the light atoms C, N, O, etc. Thus, peaks in the Patterson map [Patterson, 1934], due to the heavy atoms, are relatively easy to identify. The positions of heavy-atom vectors in the Patterson map [Patterson, 1935] may be deconvoluted to give the positions of the heavy atoms in the unit cell. Once we know the positions of heavy atoms in the crystal unit cell, we can calculate the structure factor of the heavy atom, $F_H(hkl)$, through the equation 33-6.

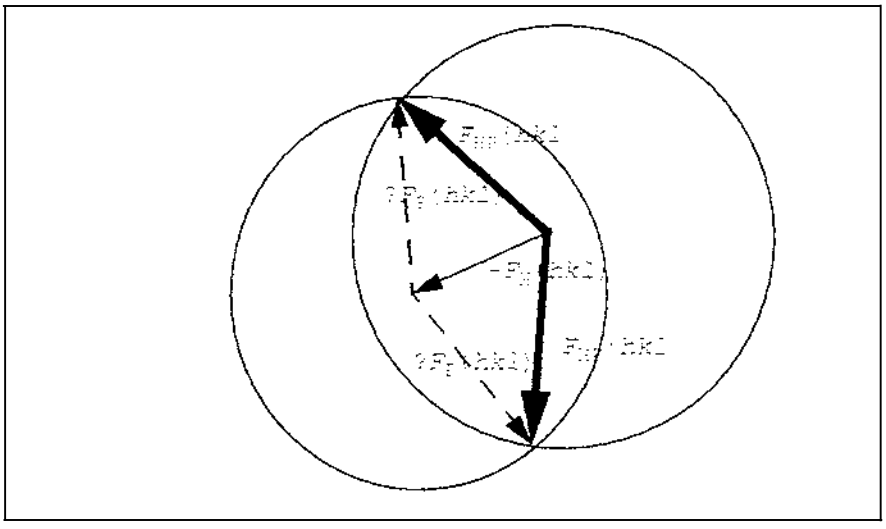


Figure 333-2: Determination of the phase angle for reflection (hkl)

At this stage, we get three numbers: $|F_{HP}(hkl)|$, $|F_P(hkl)|$, and $F_H(hkl)$. The protein phase angle can be determined as follows [Harker, 1956]: In the Argand diagram, first, draw a circle centered at the origin with radius $|F_P(hkl)|$; second, draw the vector $-F_H(hkl)$ from the origin of the Argand diagram; at the head of the vector $F_H(hkl)$, draw a circle with radius $|F_{HP}(hkl)|$. See Figure 333-2 for detail. At this stage, we obtain two possible vectors for $F_P(hkl)$. To remove the ambiguity, we need at least two isomorphous heavy-atom derivatives.

3.3.4 Molecular replacement methods

If we are sure that similarity of the unknown structure to an already known structure is present, molecular replacement is a much more efficient method. The essence of molecular replacement is to transfer the known protein structure from its crystalline arrangement into the crystal lattice of the protein for which the structure is not yet known. Two steps are involved, rotation and translation. In practice, both involve the Patterson function. After this, further refinement may give a correct solution to the unknown protein structure.

The vectors generated by the pairs of atoms in the same molecule in the Patterson map are in general relatively short compared to the vectors generated by pairs of atoms in different molecules. The former vectors are concentrated around the origin of the Patterson map. Except for orientation with respect to unit cell axes, the inner part of the Patterson map would be identical for the same molecule in different crystal structures (ignoring the noise from some intermolecular vectors), irrespective of the rotation (rotation does not modify the relationship of atoms in the molecule). For homologous molecules, the inner part of the Patterson map is not equal but very similar. Therefore, the inner part of the Patterson map (self-Patterson vectors) can supply us with the rotational relationship between the known and the unknown molecular structures. After this stage the position of the molecule in the crystal can be deduced from the translation function, in which the longer intermolecular vectors calculated by equation 33-10 are matched to those calculated as the correctly oriented molecule is moved around in the unit cell. The best match of vectors corresponds then to a correctly oriented and positioned molecule.

3.4 Structure refinement

The primary phases for the observed structure factor amplitudes are obtained by multiple isomorphous replacement (MIR) and/or Patterson search methods and molecular replacement. No matter which method, the starting phases contain errors. The electron density map calculated using such phases provides only the broad features of the protein molecule. While generally correct, the reconstructed protein model may have many

errors in the details. To monitor the agreement between the structure factors calculated from protein model and those observed, the following index is calculated:

$$R = \frac{\sum_{hkl} ||F_{obs}| - k|F_{calc}||}{\sum_{hkl} |F_{obs}|} \times 100\% \qquad (34\text{-}1)$$

Even though the R factor gives an indication of how good the protein structure is, the important thing is how to get a good protein structure.

The difference between the protein model and observed data can be expressed as:

$$D = \sum_{hkl} (|F_{obs}(hkl)| - |F_{calc}(hkl)|)^2 \qquad (34\text{-}2)$$

where the second item of equation 34-2, $|F_{calc}(hkl)|$, is given by equation 33-6, and can be modified to include a term allowing for motions of atoms:

$$|F_{calc}(hkl)| = \sum_j f_j \exp\left(-B_j \lambda^{-2} \sin^2\theta\right) \exp\left(2\pi i (hx_j + ky_j + lz_j)\right) \qquad (34\text{-}3)$$

Each atom j is described by three positional parameters and one thermal or atomic displacement parameter, respectively, x_j, y_j, z_j and B_j. The goal of protein structure refinement is to adjust the protein parameters set $P(x_j, y_j, z_j, B_j, j=1,\ldots,$ n (maximum number of atoms in structure)) to minimize the D value. Once the global minimum of D is achieved, we possibly have the best protein structure. At this stage, the R factor will be also minimized.

The equations 34-2 and 34-3 can be noted as:

$$D = f(p_1, p_2, p_3, \ldots, p_j, \ldots, p_{4n}) \qquad (34\text{-}4)$$

where p_1, p_2, p_3, up to p_{4n} are the individual parameters of P giving a total of 4×n parameters. At the minimum of D, we have:

$$\frac{\partial D}{\partial p_j} = g_j(p_1, p_2, p_3, \ldots, p_{4n}) = 0 \qquad (34\text{-}5)$$

Equation 34-5 involves a set of equations. Unfortunately, all the minima, as well the maxima and saddle points satisfy these conditions. This is why we seek the global minimum of *D*. To avoid the local minima, manual intervention in the refinement procedures is required. Thus, cycles of refinement are interspersed with calculations of electron density maps by means of which the model is manually adjusted, so-called model rebuilding.

In almost all protein structural determinations, the parameter set is very large, due to the number of atoms in a protein molecule. Therefore, the ratio of the observed reflections to the number of adjustable or variable parameters is often relatively small. For example, for BLG, these ratios are only ~2.0. As a result, overrefinement is possible during the refinement procedure. To avoid overrefinement, we can put more information into the refinement, such as standard bond lengths, bond angles, etc., as derived very precisely from small molecule crystal structure analyses. There are two methods of refinement which absorb the small molecules' data: constrained (rigid) or restrained (flexible) refinements.

One way to monitor possible overrefinement is to apply the R_f monitor [Brünger, 1992]. Before the beginning of any refinement, ~5% or several hundreds of reflections are set aside and used to calculate an R factor called R_f. Because these reflections are not used at any stage in the refinement, R_f is a valuable monitor on the reliability of the whole refinement procedure.

Chapter 4 Experimental procedures

4.1 Source of material

4.1.1 Source of bovine ß-lactoglobulin

The national herd of New Zealand is about 18% Jersey, 57% Holstein/ Friesian, and 16% Jersey-Friesian cross [Hill *et al.*, 1996]. Figure 41-1 illustrates a typical dairy herd of New Zealand. As the dairy industry of New Zealand is a major participant of the national economy, the factors related to the industrial processing of milk remain a key research interest, for example, ß-lactoglobulin. The samples of bovine ß-lactoglobulin variant A (BLGA) and variant B (BLGB) were isolated from the milk of homozygous cows and were identified by the New Zealand Dairy Research Institute. The variants were detected by electrophoresis under near-native conditions using paper, agar gels, or starch gels as supporting material. Other techniques have also been used in the detection of the variants. Separate preparations of ß-lactoglobulin from the milk of cows previously typed as ß-lactoglobulin AA or ß-lactoglobulin BB were conducted in later isolation procedures.

Figure 41-1: Dairy herd of New Zealand

4.1.2 Preparations of whey

Method-I: Late-season lactation milk was collected from two Friesian/ Jersey cows, which had been previously typed for genetic variants ß-lactoglobulin AA and BB (Anema and Creamer, 1993), immediately prior to the commencement of purification. A total volume of 2.5 L was collected from each cow and then cooled to 4 °C. The cold (4 °C) milk was defatted by centrifugation (2×10,000g, 20 mins at 4 °C, Sorvall RC2-B superspeed centrifuge, Sorvall Inc., Norwalk, Connecticut). Skim milk was then frozen and stored at -22 °C. When required, samples were thawed at room temperature and 250 mL aliquots of skim milk were used for the preparation of bovine ß-lactoglobulin.

The skim milk was then warmed to 40 °C and the pH was adjusted to 6.6. A solution of 2 M $CaCl_2$ (BDH, AR) was added slowly with constant stirring of the skim milk up to 60 mL per litre of skim milk. While stirring, the pH of the skim milk was maintained at 6.6 with 1 M HCl (BDH, AR). The temperature of the skim milk was kept at 40 °C to aid agglomeration of casein particles. The preparation was left to stand for approximately 10 mins. Then the casein aggregates were removed by centrifugation (12,000 g, 30 mins at 4 °C). The supernatant (whey) was retained, and the pH was adjusted to 5.9 and concentrated in a 250 mL ultra-filtration cell (model 8400, prod no 5124, Amicon, Inc., Beverly, MA) at 4 °C, from 225 mL to 50 mL for the A variant preparation and 225 mL to 130 mL for the B variant preparation. A YM10 membrane was used (10,000 M_r cut-off).

Method-II: Mid-late season milk was collected from homozygous (i.e. ß-lactoglobulin AA or BB) Friesian/Jersey cross-bred cows typed for all genetic variants (Anema and Creamer, 1993). Milk was collected both the evening before and the morning of the commencement of the purification. The initial volumes of whole milk were, from AA cows, 15 L; and from BB cows, 20 L. Warm (40 °C) whole milk from individual cows was defatted using a separator (model 103 AE, Alfa-Laval Separator Co., NZ). Milk from different cows was not pooled until gel electrophoresis results confirmed that all cows were homozygous with respect to ß-lactoglobulin. Caseins were precipitated from the skim milk *via* preparation of "acid whey"

[Mailliart and Ribadeau-Dumas, 1988]. Warm (40 °C) skim milk (to aid agglomeration of casein particles) was titrated to pH 4.6 with 1 M HCl (BDH, AR). The precipitated caseins were removed by gravity filtration through two-layer polyterylene cloth, which retained all but the smallest casein particles. Whey was then concentrated by a freeze/thaw technique (detailed as follows) followed by ultra filtration.

Freeze/Thaw: The above whey was frozen at -16 °C, then thawed at room temperature, taking care to keep containers stationary during thawing thus allowing a protein concentration gradient to form. The denser protein-rich portion was located towards the bottom of the vessel and the lighter protein-defficient portion was located towards the top. The protein-rich portion was then carefully siphoned off, taking the bottom 1/4- 1/3 of the total volume. Lactose crystals and casein aggregates were observed in the bottom of all whey-containing vessels once thawed (thawed overnight at room temperature). Five freeze/thaw/siphon cycles were performed on the whey from each cow. After freeze/thaw concentration, volumes for each preparation had been reduced.

4.1.3 Isolation of bovine ß-lactoglobulin A and B

The whey preparation methods described above gave identical bovine ß-lactoglobulin. ß-Lactoglobulin A was isolated from whey by gel filtration chromatography on a column (50 × 600 mm, model XK 50/60, Pharmacia) of Superdex-75 (prep grade, Pharmacia) in 20mM phosphate (BDH,AR) buffer, pH 6.0 with 30 mM NaCl (BDH, AR) at a flow rate of 5 mL/min. Samples of 5 mL (80 mg protein) were loaded onto the column. Fractions of 5 mL were collected and those containing pure ß-lactoglobulin were identified using gel electrophoresis (Anema and Creamer, 1993). Pure ß-lactoglobulin fractions were then pooled. The Pharmacia FPLC system used for gel filtration consisted of two P500 pumps, three MV8 valve selectors, one IMV7 automated injector valve, one MV7 manual injector valve, a Frac 20 fraction collector and a P1 pump. All chromatographic operations were carried out at ambient temperature.

ß-Lactoglobulin B was purified from whey using ion-exchange chromatography. Whey (150ml) was dialysed against four changes of piperazine buffer (20 mM, pH 6.0, Sigma) with 30 mM NaCl and ion-exchange

chromatography was conducted on a Q-Sepharose fast-flow (prod no 17-0510-01, Pharmacia) column (Ø50 × 150 mm, model XK 50/20, Pharmacia). The column was equilibrated with piperazine buffer (20mM, pH 6.0, Sigma). Samples (50 mL) were loaded onto the column and whey species were eluted at a flow rate of 10 mL/min with a 0-0.35 M NaCl gradient over a period of 1 h (2 column volumes). Fractions were collected and pooled.

4.2 Evaluation of the purity and concentration of protein samples

4.2.1 Principle of SDS electrophoresis

A given protein molecule has on its surface a number of amino acid side chains which are negatively or positively charged. Except at their isoelectric point, protein molecules are charged particles and are mobile in an electric field. Besides the electrical charge, the protein molecule's shape also influences the speed at which the protein molecule moves through the supporting gel matrix. SDS is an amphipathic molecule and can be bound to protein molecules, disrupting their structure. When so denatured, the influence of the molecule's shape and the electric charge on the mobility of protein molecule in the electric field is minimized. As a result, the protein molecules will move at a speed almost exclusively determined by their molecular weight. Such mobility has an inverse linear relationship with the logarithm of the protein molecular weight [Weber and Osbom, 1969; Maurer and Dati, 1972]. Inclusion of SDS in the sample decreases the resolution of electrophoresis. To get high quality data, discontinuous pH and gel conditions are used to form the sample loading and resolving gel system. The protein will be greatly concentrated in the loading gel before moving into the resolving gel. This technique improves the resolution of SDS polyacrylamide gel electrophoresis (SDS-PAGE).

4.2.2 Procedure of SDS electrophoresis

The SDS-PAGEs (100×80×1.5 mm) were made up according to Table 422-1. The corresponding stock solutions are summarised in Table 422-2. The gels were made at least overnight before the commencement of electrophoresis. The amount of sample loaded was sufficient to allow each separate band to contain 1 ~ 10 µgram of protein. This amount is neccessary for the

subsequent staining of the gel. A PS 500XT DC power supply was used to provide the electric field. Cooling water was supplied during whole electrophoresis procedure with the conditions: 10 mA, 50V, 4hrs. After the electrophoresis, the gel was stripped off the gel mould. The gel was submerged in the stain solution and heated in a microwave oven for 4 mins. The stained gels were then cleaned with distilled water. The heavily stained gels were submerged in destaining solution I and heated in a microwave oven for 4 mins, then cleaned with distilled water. Destained gels were then submerged in the destain solution II for further destaining. The second destain process can be as long as convenient.

Table 422-1: Resolving-gel and stacking-gel recipes

	resolving(mL)	stacking(mL)
Monomer solution	10	3.6
Resolving gel buffer	5	
Stacking gel buffer		10.4
Detergent	0.2	0.2
H2O	4.67	11
**********	Degas for 15 mins	**********
Initiator	0.1	0.15
TEMED	0.01	0.016
**********	Apply	**********
Water saturated n-butanol	4 × 0.1 mL	

Table 422-2: Stock solutions for SDS electrophoresis

Monomer solution, w/v	30% Acrylamine, 2.7% Bis
Resolving gel buffer	1.5M Tris-HCl pH 8.8
Stacking gel buffer	0.5 M Tris-HCl pH 6.8
Detergent, w/v	10% SDS
Initiator, w/v	10% ammonium persulfate
Resolving Gel Overlay	0.375 M Tris-HCl pH 8.8, 0.1% SDS
2 × Treatment buffer	0.125 M Tris-HCl pH 6.8, 4% SDS, 20% glycerol, 10% 2-mercaptoethanol
Tank buffer	0.025 M Tris pH 8.3,0.192 M glycine, 0.1% SDS
Stain solution, v/v	0.125% Coomassie Blue R-250, 50% methanol, 10% acetic acid
Destaining solution I, v/v	50% methanol, 10% acetic acid
Destaining solution II, v/v	7% acetic acid, 5% methanol
TEMED	
Water saturated n-butanol	

4.2.3 Purity of samples on the SDS-PAGE.

The purity of the bovine ß-lactoglobulin sample was satisfactory according to SDS-PAGE. A typical result is shown in Figure 42-1. The molecular weight is ~18,000 dalton according to the mobility of ß-lactoglobulin in the SDS-PAGE, a result corresponding to the monomer ß-lactoglobulin. The molecular weight markers are phosphorylase b, bovine serum albumin, ovalbumin, carbonic anhydrase, soybean trypsin inhibitor, and α-lactalalbumin with molecular weights of 94,000, 67,000, 43,000, 30,000, 20,100 and 14,400 dalton, respectively.

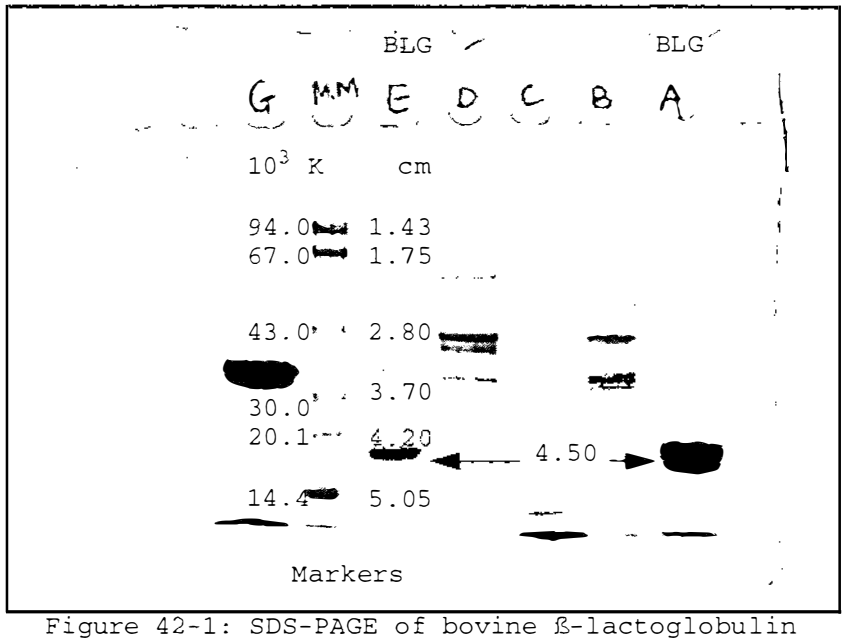


Figure 42-1: SDS-PAGE of bovine ß-lactoglobulin

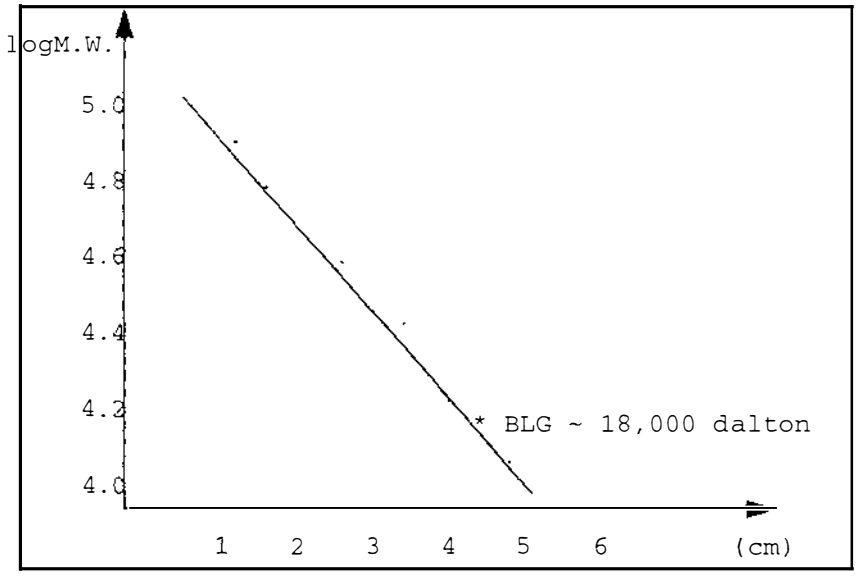


Figure 42-2: Mobility *versus* logarithm of molecular weight

4.2.4 Optical spectra of bovine ß-lactoglobulin samples

Most protein molecules have several amino acid residues which are UV active due to their side chains containing an aromatic moiety. Townend and Timesheff (1960) determined the BLG extinction coefficient as 0.96 (278 nm) litres/g/cm, see Table 251-1. To simplify the data analysis, a value of 1 O.D. cm^{-1}/mg/mL (280 nm) was used as the extinction coefficient in the experiments. The concentration of ß-lactoglobulin sample was measured by a Hitachi UV spectroscanner in a cell with optical length of 1 cm. Figure 42-3 shows a typical spectrum of bovine BLG, from 230 nm to 700 nm. At 280 nm, O.D. ~ 1, corresponding to a low concentration of ~1 mg/mL. This storage concentration is not suitable for crystallization experiments. The native BLG sample demonstrated the normal protein absorption behaviour.

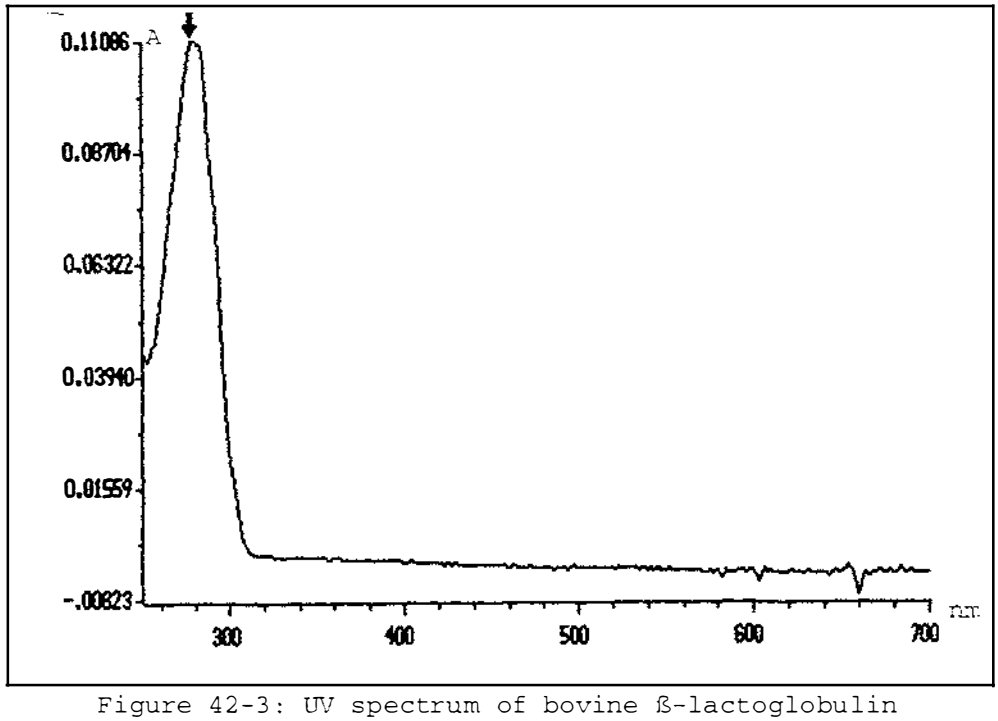


Figure 42-3: UV spectrum of bovine ß-lactoglobulin

The spectrum of bovine ß-lactoglobulin is different after binding molecules that absorb in the UV-visible region, such as retinol, heme, and protoporphyrin IX. The binding of retinol to bovine BLG gives a slight yellow colour to the solution, and some extra peaks in the spectrum (Figure 42-4). The putative heme-BLG and protoporphyrin IX-BLG complexes show different optical characteristics (Figure 42-5). The SDS-BLG complex shows no change in optical properties with respect to the native BLG solution. These putative bovine BLG-ligand complexes have also

been subjected to crystal-screen trials.

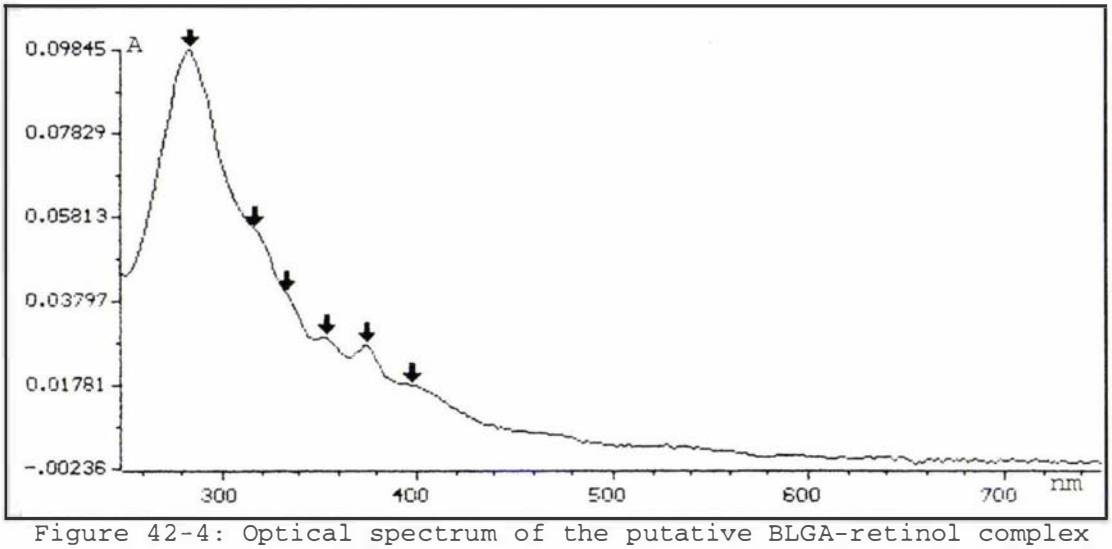


Figure 42-4: Optical spectrum of the putative BLGA-retinol complex

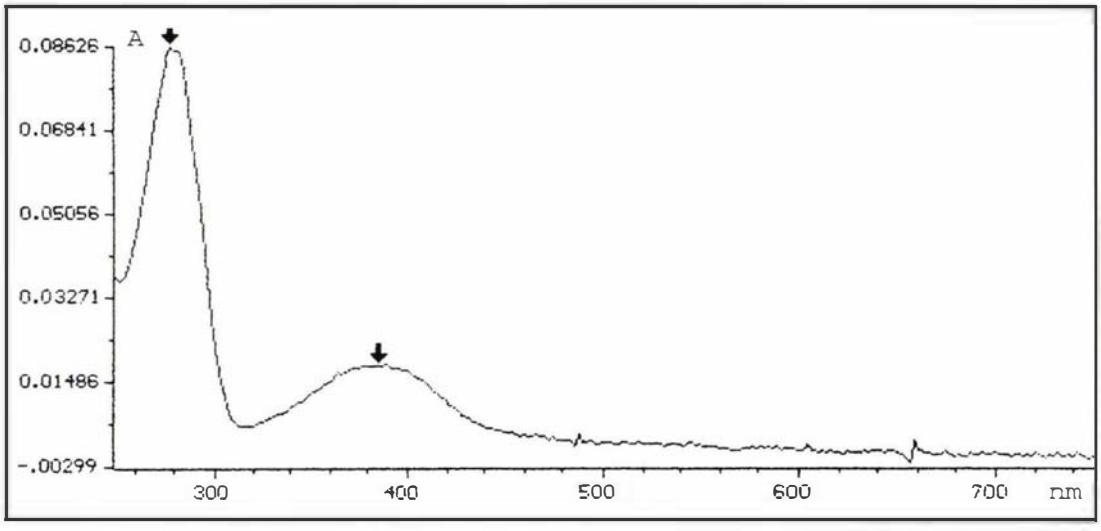


Figure 42-5: Optical spectrum of the putative BLGA-heme complex

4.3 Final sample preparation

The SDS-PAGE results show acceptable purity of bovine BLG samples. However, the concentration and supporting media are not suitable for crystallization. The concentration of bovine BLG required for crystallization is relatively high at 20-30 mg/mL. In addition, the storage buffer for ß-lactoglobulin sample may vary with the preparation procedure. It is important then that before settting up the crystallization screen, bovine ß-lactoglobulin is suitably concentrated and in a proper buffer, for example, 10 mM HEPES, pH 7.4.

Pure samples (1 mg/mL) were pooled in a Centricon-10 (Figure 43-1a, Amicon, Inc. MW cut-off 10,000) with 0.5 ~ 1.5 mL for each centrifugation cycle (5,000 rpm, rotor SORVALL, SS-34, Dufour Inc.), subjected to 3 ~ 5 cycles of washing and concentrating, to achieve a final volume of 0.1-0.2 mL with bovine BLG concentrated by a factor of around 20 ~ 30. This resulted in final protein concentrations of 20 ~ 30 mg/mL. The concentrated samples were washed by HEPES buffer (10 mM, 0.02% azide, pH 7.4). A total of 3 ~ 5 mL of HEPES buffer was required for the washing process. The final volume was kept at 0.1-0.2 mL. This washing and concentrating also desalted the bovine ß-lactoglobulin samples. Before storage of the concentrated bovine ß-lactoglobulin sample, the sample was filtered by a Micropure filter (Amicon, Inc. cut-off 0.1 μm) in the cold room at a speed of 5,000 rpm to remove any possible microoganisms (see Figure 43-1b). The final concentration was around 30 mg/mL, and the sample was stored at 4 °C.

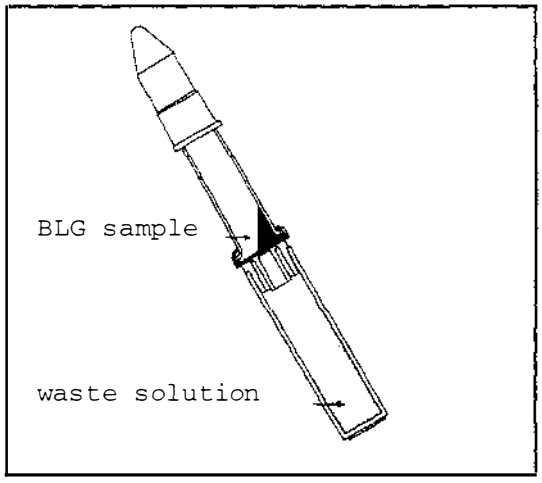


Figure 43-1a: Filtration of sample (Centricon filter) by centrifugation

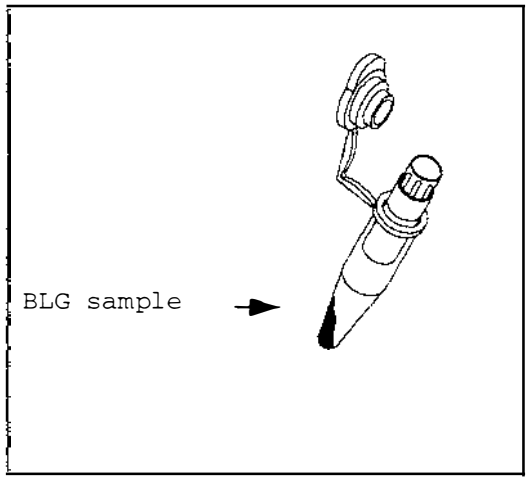


Figure 43-1b: Filtration of sample (Micropure filter) by centrifugation

4.4 Crystallization and pH measurements

4.4.1 Crystallization of bovine BLG

When molecules precipitate from solution, they attempt to reach the lowest free energy state. This is often accomplished by packing themselves in a regular way, i.e. to form crystals. Compared to small molecules, protein molecules contain a polypeptide backbone, which is not as rigid as covalent bonds between atoms in small molecules, and can

be disturbed by many subtle influences, such as salts, crystal packing, etc. As a result, it is more difficult for protein molecules, compared to the small molecules, to adopt a uniform conformation, a requirement for regular packing. Protein molecules are also covered by a solvent mask, which largely prevents the contact of protein molecules with each other and hinders crystallization. In general, for successful crystallization, it is necessary to remove at least in part the solvent mask around the protein molecules (by adding precipitant) and then to refine the experimental conditions to achieve uniform conformation for individual protein molecules. Several precipitants, PEG, MEG, and ammonium sulfate, have been tested in the bovine BLG crystallization trials. Only ammonium sulfate gave suitable crystals for X-ray diffraction experiments.

4.4.2 Crystallization screen and trials

Ammonium sulfate traditionally has been the precipitant for crystallization of bovine ß-lactoglobulin and has provided the best crystals in our X-ray diffraction experiments. The pH of screen solutions were measured before and after adding ammonium sulfate. After adding ammonium sulfate, the pH of screen solutions were shifted a little, since the concentration of ammonium sulfate is high in the crystallization trials and the ammonium ion is very weakly acidic. Each screen solution contains 1.8 ~ 2.0 mL of 1M buffer, 5.5 ~ 8.2 mL of 3.4 ~ 4.0M ammonium sulfate, 0 ~ 2.5 mL of Milli-Q water, made up to a total volume of 10.0 mL; see Tables 44-2 and 44-3. The buffer was varied to produce a 4 × 6 screen matrix, covering the range of pH of the Tanford transition.

Table 44-1: Stock solutions of screen solutions

Chemicals	pH	concentration (M)
Acetic Acid/KOH	5.01	1.0
Cacodylic Acid/KOH	6.10	1.0
HEPES/KOH	7.40	1.0
TAPSO/KOH I	8.71	1.0
Ammonium sulfate		4.0
Bis-tris/HCl	6.90	1.0
Tris/HCl	7.70	1.0
TAPSO/KOH II	8.50	1.0

The pH of the buffers were measured at room temperature (25 °C), the pH meter was calibrated by standard pH buffer with pH 7.00. Table 44-1 gives the stock solutions for the preparation of screen solutions.

Table 44-2: Crystallization screen matrix I. Volumes (ml) of each stock solution are entered in the table.

Buffers	Concentration of ammonium sulfate (3.4M)					
(1M, 1.8 mL)	2.20	2.32	2.44	2.56	2.68	2.80
Acetic Acid/KOH	6.47	6.82	7.18	7.53	7.88	8.20
Cacodylic Acid/KOH	6.47	6.82	7.18	7.53	7.88	8.20
HEPES/KOH	6.47	6.82	7.18	7.53	7.88	8.20
TAPSO/KOH I	6.47	6.82	7.18	7.53	7.88	8.20

Table 44-3: Crystallization screen matrix II. Volumes (ml) of each stock solution are entered in the table.

Buffers	Concentration of ammonium sulfate (4M)					
(1M, 2.0 mL)	2.20	2.32	2.44	2.56	2.68	2.80
Cacodylic Acid/KOH	5.5	5.8	6.1	6.4	6.7	7.0
Bis-tris/HCl	5.5	5.8	6.1	6.4	6.7	7.0
Tris/HCl	5.5	5.8	6.1	6.4	6.7	7.0
TAPSO/KOH II	5.5	5.8	6.1	6.4	6.7	7.0

BLGA and BLGB were crystallized by the screen matrix I. To minimise the protein sample consumed and to refine the experimental conditions efficiently, the hanging drop method was applied, where each trial required only 2 μL protein sample with concentration 20-30 mg/mL. The well solutions with nominated pH 6.1, 7.4 and 8.7, which produced the crystals suitable for data collection, have their pH shifted to 6.2, 7.1 and 8.2 by the ammonium sulfate, respectively. In general, both BLGA and BLGB will produce crystals in lattice Z within two to four days. Figure 44-1 illustrates the typical crystal habit.

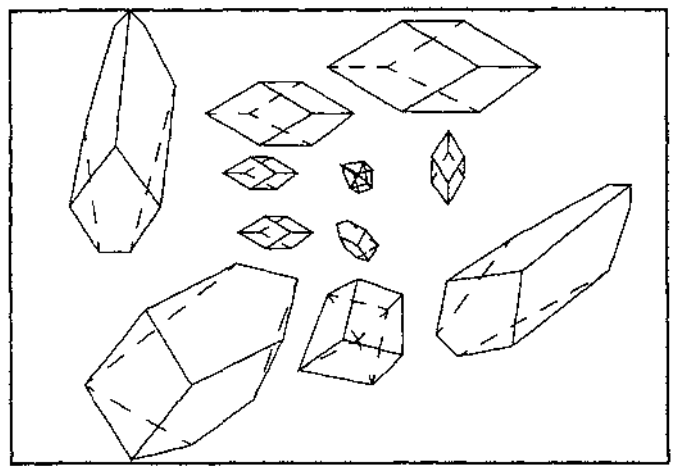


Figure 44-1: Sketch of BLGA crystals in lattice Z

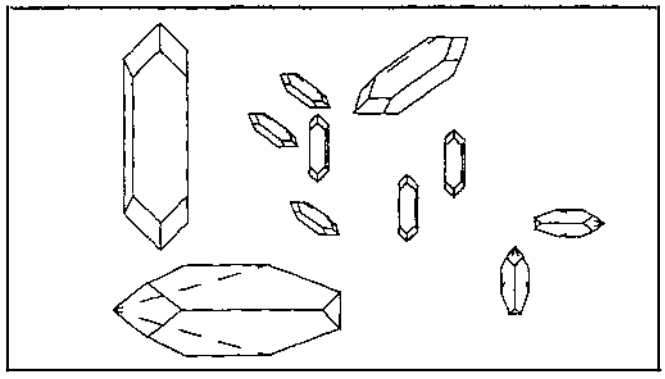


Figure 44-2: Sketch of BLGA-SDS crystals

Heme, protoporphyrin IX and retinol were added to the BLG sample in the molar ratios of 1.2:1 individually. These ligand-BLG were subjected to crystal screen matrix I. No corresponding ligand-BLG crystal was found. The SDS derivative of BLG was prepared according to the method of McMeekin *et al.* (1949). A different crystal form (space group I4) was produced, as shown in Figure 44-2. These crystals diffracted to a low resolution of 5.5 Å.

12-bromododecanoic acid was dissolved in ethanol (14 mM) and added to the BLGA solution (~ 1.1 mM) in the molar ratio of 1.5:1. At room temperature, such a BLGA derivative was subjected to crystal screen matrix II (hanging drop method, 2 µL *versus* 2 µL) and successfully cocrystallised in lattice Z at nominated pH 7.7 and 8.5. A different crystal form was found at pH 6.1 and 6.9. Due to the extreme fragility of these crystals, no diffraction data were collected. The pH value of the well solution which produced the BLGA-BrC12 crystal for data collection was shifted from nominated pH 7.7 to pH 7.3.

4.5 Diffraction system and data collection

4.5.1 Data collection system

The Rigaku "Rotaflex" RU-200 series is a rotating anode X-ray generator which can yield a maximum output of 12 kW (60 kV/200 mA). The heat generated by electron impact on the rotating anode is dissipated by a closed-circuit water cooling system. The X-ray beam from the anode is monochromated by a graphite monochromator. The protein crystal is mounted on the goniometer and centered in the path of the X-ray beam, see Figure 45-1. The diffracted X-ray beams are recorded by a Rigaku RAxis IIC image plate. The image plate is scanned by a He-Ne laser beam, and the luminescence stimulated from the image plate is amplified by a photomultiplier, and the signals are sent to a computer, a VAX station 3100, which is also the control unit of whole system.

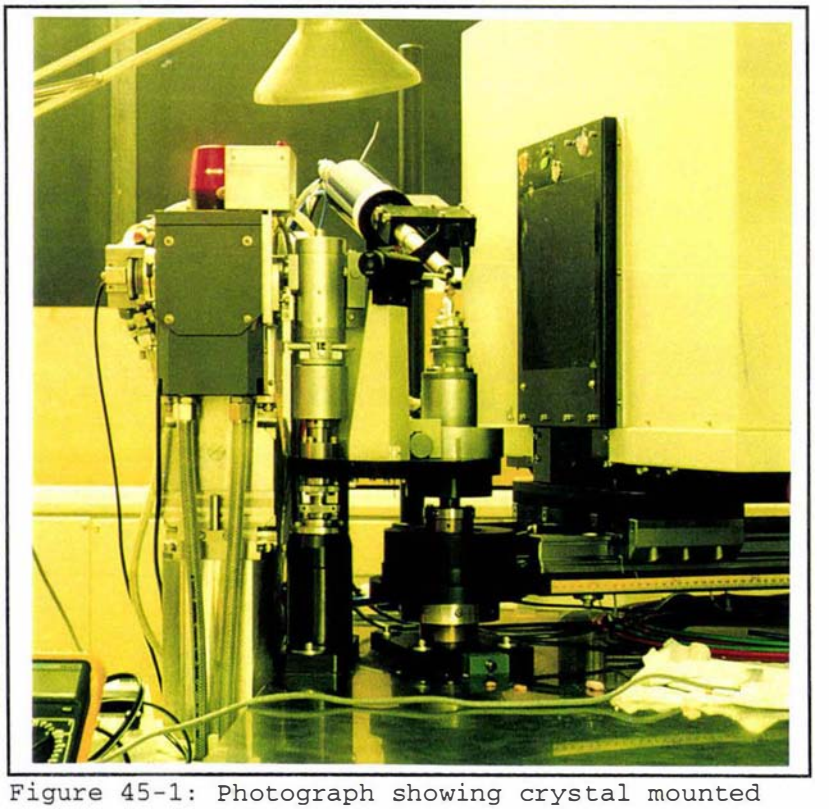


Figure 45-1: Photograph showing crystal mounted on the goniometer

4.5.2 Crystal mounting and data collection

Bovine ß-lactoglobulin crystals were mounted into thin glass capillaries (ϕ 0.7 mm, Mark-Rohrchen, Germany), under a 30 × stereoscopic microscope.

After mounting, both ends of the capillary were filled with the mother liquor, then sealed with bees' wax. In general, the finished capillary with mounted crystal was 2-2.5 cm in length. Such capillaries were vertically attached to the top of the goniometer with plasticine. The diffraction patterns were collected at room temperature with generator voltage 50 kV and current 100 mA by the oscillation method with the image plate distance ~100 mm and oscillation angle 1.2 or 1.5 degree.

The first image of the diffracted crystal was collected in 12 minutes. With PREDICT [Jeffery, 1997] or STRATEGY [Ravelli, 1997] the crystal's orientation was determined and the optimized start rotation angle was calculated for the collection of subsequent images. The diffraction images were saved to magnetic tapes after finishing the diffraction experiments. A representative image, the second frame of data set blgza10, is illustrated in Figure 45-2. In general, the diffraction patterns extend beyond the water circle.

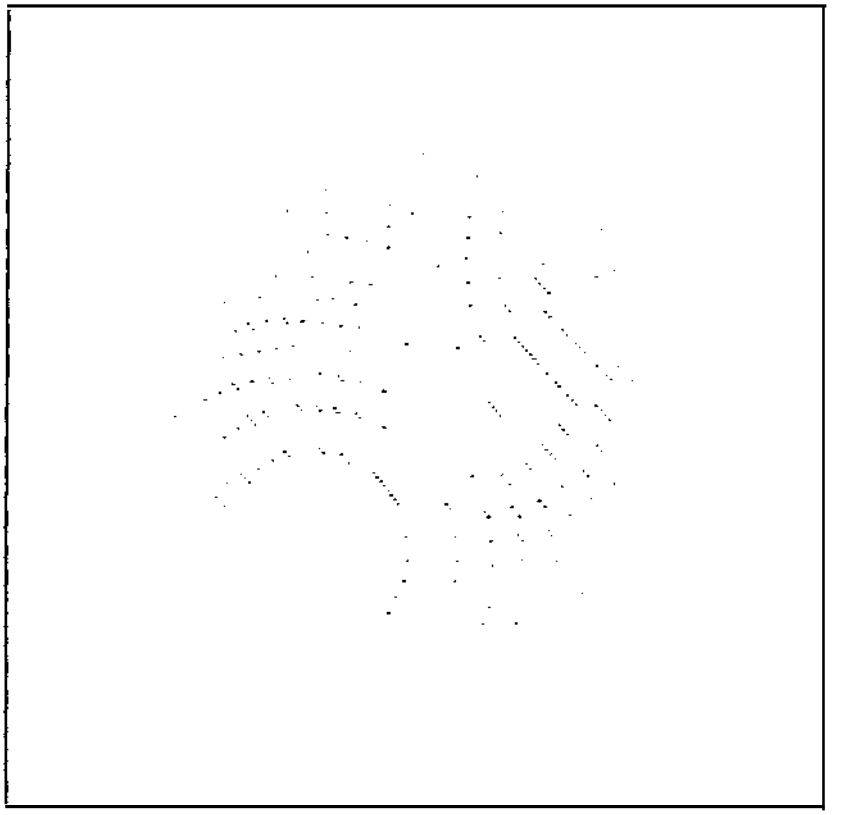


Figure 45-2: Representative X-ray diffraction image

4.6 Processing of diffraction data

4.6.1 Indexing and merging of reflections

Each spot on a frame of data from the RAxis IIC must be related to the reflection plane (hkl) which generates this diffraction. The oscillation method records the diffraction pattern in a manner that distorts reciprocal space, so that the task of indexing the diffraction pattern is very difficult to do manually. DENZO [Otwinowski, 1996] is used to index automatically the reflections collected. With each frame of the X-ray diffraction pattern, DENZO extracts the diffraction spots and reduces them to a text file, which contains the hkl index and intensity for each reflection along with the background and an estimation of the error. In processing data, the coarse crystal parameters and the detector orientation parameters are required in order to index, first a single frame of the data set (usually, the first image). To obtain the proper crystal parameters, DENZO tests all 14 Bravais lattices and outputs the best unit cell for each lattice and the corresponding index of distortion as a percentage. A suitable cell is chosen according to a criterion of the highest symmetry lattice but with a low index of distortion (usually less than 0.1%). Based on the indexing of the first frame, all the following frames of data are indexed in batch mode. The crystal parameters and the reciprocal lattice orientation with respect to the spindle, beam, and vertical axes of the data collection system are refined for each frame.

Some reflections may be recorded more than once. The redundant or duplicate measurements for an individual reflection need to be merged or averaged together. Besides this way, redundancy can originate in another way. First, for lattice types other than triclinic, the symmetry operations introduce redundancy to the indexed reflection file. Second, the difference in intensity of reflections hkl and -h-k-l can be ignored usually. Program SCALEPACK [Otwinowski, 1996] merges the duplicated reflections (in BLG cases, the symmetry-related reflections and Friedel pairs) with appropriate scaling for crystal decomposition and absorption, and produces a file containing only the unique reflections. The error model was adjusted so that $\chi^2 \approx 1$, independent of the magnitude or resolution of the intensity data, in order to improve the uniformity

of the reflection data set. The statistics describing the merging process in the later sections are defined as follows:

$$R = \frac{\sum |I - I_{ave}|}{\sum |I|} \qquad (461\text{-}1)$$

$$R^2 = \frac{\sum (I - I_{ave})^2}{\sum I^2} \qquad (461\text{-}2)$$

$$\chi^2 = \frac{\sum (I - I_{ave})^2}{Error^2 \cdot \frac{N}{N-1}} \qquad (461\text{-}3)$$

Data were processed and those retained for subsequent calculations had a mean value for $I/\sigma(I)$ in the highest resolution shell of ~2.0. Single measurements are excluded from the summation. I: intensity of reflection, I_{ave}: average of duplicate and redundant measurements, N: number of duplicate and redundant measurements, $Error$: estimated error for a given reflection shell, which is subjected to adjustment in each cycle to give $\chi^2 \approx 1$.

4.6.2 Data processing for BLGA at pH 8.2

The output from DENZO for indexing a frame of data into each of the 14 Bravais lattices is given in Table 462-1. The primitive hexagonal lattice gives a very low distortion index of only 0.09%, and has the highest symmetry compared to other lattices with low indices of distortion, such as C-centred orthorhombic and primitive monoclinic. As a result, the primitive hexagonal is the suggested lattice type for BLGA at pH 8.2. The coarse crystal parameters and crystal orientation parameters are provided simultaneously. Through batch mode, all image frames were then reduced to corresponding indexed diffraction data files. The indexed diffraction data were merged by program SCALEPACK [Otwinowski, 1996]. After ten cycles of post-refinement of the crystal parameters, the redundant reflections were scaled and merged together to provide a unique diffraction data file. The completeness is given in Table 462-2, and the data merging statistics are shown in Table 462-3. The overall completeness to 2.46 Å is 97.1% with an overall value of R of 0.080 for averaging the data.

Table 462-1: Auto-indexing by DENZO for BLGA at pH 8.2

Lattice	Metric tensor distortion index	Best cell (symmetrized) Best cell (without symmetry restraints)					
primitive cubic	52.57%	94.08	54.39	113.32	90.07	89.98	30.11
		90.65	90.65	90.65	90.00	90.00	90.00
I centred cubic	40.90%	125.61	94.43	113.32	90.14	154.35	89.97
		111.86	111.86	111.86	90.00	90.00	90.00
F centred cubic	34.42%	147.27	147.31	125.63	46.11	46.15	79.40
		140.44	140.44	140.44	90.00	90.00	90.00
primitive rhombohedral	8.81%	125.61	125.64	125.63	24.99	25.01	43.98
		125.63	125.63	125.63	32.43	32.43	32.43
		76.84	76.84	356.75	90.00	90.00	120.00
primitive hexagonal	0.09%	54.39	54.37	113.32	90.10	90.07	119.77
		54.38	54.38	113.32	90.00	90.00	120.00
primitive tetragonal	7.83%	54.37	54.39	113.32	90.07	89.90	60.23
		54.38	54.38	113.32	90.00	90.00	90.00
I centred tetragonal	4.23%	54.37	54.39	232.97	83.32	96.65	60.23
		54.38	54.38	232.97	90.00	90.00	90.00
primitive orthorhombic	7.83%	54.37	54.39	113.32	90.07	89.90	60.23
		54.37	54.39	113.32	90.00	90.00	90.00
C centred orthorhombic	0.08%	54.57	94.08	113.32	89.98	89.83	90.03
		54.57	94.08	113.32	90.00	90.00	90.00
I centred orthorhombic	4.23%	54.37	54.39	232.97	83.32	96.65	60.23
		54.37	54.39	232.97	90.00	90.00	90.00
F centred orthorhombic	4.44%	54.37	94.43	232.99	90.09	103.40	90.21
		54.37	94.43	232.99	90.00	90.00	90.00
primitive monoclinic	0.07%	54.37	113.32	54.39	90.07	119.77	90.10
		54.37	113.32	54.39	90.00	119.77	90.00
C centred monoclinic	0.02%	54.57	94.08	113.32	90.02	90.17	90.03
		54.57	94.08	113.32	90.00	90.17	90.00
primitive triclinic	0.00%	54.37	54.39	113.32	90.07	90.10	119.77

Volume of the primitive cell 290890.
autoindex unit cell 54.52 54.52 113.32 90.00 90.00 120.00
crystal rotx, roty, rotz -75.677 127.201 32.191
Autoindex Xbeam, Ybeam 100.24 97.02

Table 462-2 : Completeness of data for BLGA at pH 8.2

Resolution shell (Å)		I/Sigma in resolution shells: % of reflections with I / σI less than 0	1	2	3	5	10	20	>20	total
99.00	6.45	0.4	1.2	1.8	2.4	3.0	4.0	66.3	24.2	90.5
6.45	5.12	1.8	2.5	3.2	4.3	6.9	34.8	92.7	5.0	97.7
5.12	4.47	1.7	3.1	3.3	5.0	6.7	34.8	98.3	0.0	98.3
4.47	4.06	1.6	3.7	5.1	6.2	10.0	36.8	99.3	0.0	99.3
4.06	3.77	1.4	3.6	4.8	7.5	9.6	43.0	98.8	0.0	98.8
3.77	3.55	1.9	4.7	8.0	10.1	15.3	48.4	99.8	0.0	99.8
3.55	3.37	2.7	4.0	5.7	9.4	16.5	57.3	98.3	0.0	98.3
3.37	3.22	3.1	7.0	10.3	15.3	25.7	77.2	99.0	0.0	99.0
3.22	3.10	4.2	9.9	15.7	19.0	32.2	83.1	98.8	0.0	98.8
3.10	2.99	4.0	10.6	16.3	22.4	38.2	86.9	99.0	0.0	99.0
2.99	2.90	7.4	17.0	26.1	34.1	58.0	97.1	98.6	0.0	98.6
2.90	2.82	4.6	15.6	26.9	39.7	65.1	99.0	99.3	0.0	99.3

(continued)

Table 462-2 : Completeness of data for BLGA at pH 8.2

Resolution shell (Å)	I/Sigma in resolution shells: % of reflections with I / σI less than 0	1	2	3	5	10	20	>20	total
2.82 2.74	12.6	22.1	36.4	48.2	80.2	98.2	98.2	0.0	98.2
2.74 2.67	12.5	27.7	41.9	57.1	81.4	99.0	99.0	0.0	99.0
2.67 2.61	12.5	27.7	44.3	56.5	77.6	96.9	97.2	0.0	97.2
2.61 2.56	11.0	27.6	48.3	70.7	92.4	95.7	95.7	0.0	95.7
2.56 2.51	13.0	28.2	51.1	73.3	89.8	91.6	91.6	0.0	91.6
2.51 2.46	12.7	29.8	52.1	66.7	86.6	89.6	89.6	0.0	89.6
All hkl	5.9	13.3	21.8	29.7	43.1	69.6	95.2	1.9	97.1

Table 462-3: Data merging statistics for BLGA at pH 8.2

Resolution shell (Å)	I	error	stat.	χ^2	R	R^2
99.00 6.45	8994.2	483.9	107.1	1.076	0.041	0.046
6.45 5.12	3810.2	303.7	58.9	1.094	0.064	0.073
5.12 4.47	7802.3	708.7	87.8	1.056	0.070	0.073
4.47 4.06	4742.8	423.3	72.6	1.063	0.076	0.092
4.06 3.77	3169.3	299.7	67.5	0.930	0.086	0.114
3.77 3.55	2264.6	223.8	60.5	0.972	0.086	0.106
3.55 3.37	1672.6	168.6	56.5	1.045	0.093	0.107
3.37 3.22	1243.1	151.5	56.0	0.989	0.101	0.116
3.22 3.10	946.3	123.1	54.1	0.933	0.109	0.118
3.10 2.99	662.6	99.4	56.2	1.025	0.128	0.128
2.99 2.90	524.4	98.1	54.6	0.934	0.143	0.111
2.90 2.82	443.6	96.8	55.4	0.861	0.158	0.135
2.82 2.74	281.0	81.7	56.4	1.071	0.265	0.246
2.74 2.67	227.0	74.1	55.9	1.169	0.297	0.269
2.67 2.61	195.3	68.0	55.7	1.130	0.317	0.296
2.61 2.56	161.1	73.4	56.0	1.071	0.385	0.337
2.56 2.51	143.3	73.1	55.7	1.119	0.437	0.388
2.51 2.46	135.2	67.6	56.0	1.001	0.390	0.319
All reflections	2170.5	206.4	62.8	1.028	0.080	0.072

4.6.3 Data processing for BLGA at pH 7.1

As for BLGA at pH 8.2, the primitive hexagonal lattice is the suggested lattice type for BLGA at pH 7.1, as shown in Table 463-1. The redundancy of this data set is relatively high compared to others (see Table 511-1). The crystal parameters of BLGA at pH 7.1 were refined by the post-refinement procedure provided by the SCALEPACK. The redundant reflections were then merged together to provide a unique diffraction data file. The overall completeness is 96.4 % to 2.24 Å, as detailed in Table 463-2, with an overall value of R of 0.038 for averaging the data (Table 463-3).

Table 463-1: Auto-indexing by DENZO for BLGA at pH 7.1

Lattice	Metric tensor distortion index	Best cell (symmetrized) Best cell (without symmetry restraints)					
primitive cubic	52.59%	53.89	93.25	112.31	89.86	89.89	30.04
		89.84	89.84	89.84	90.00	90.00	90.00
I centred cubic	40.98%	93.44	124.45	112.31	25.64	90.06	90.09
		110.80	110.80	110.80	90.00	90.00	90.00
F centred cubic	34.42%	145.81	124.59	146.15	46.03	79.40	46.16
		139.21	139.21	139.21	90.00	90.00	90.00
primitive rhombohedral	8.82%	124.61	124.48	124.45	25.04	44.03	24.97
		124.51	124.51	124.51	32.45	32.45	32.45
		76.21	76.21	353.56	90.00	90.00	120.00
primitive hexagonal	0.08%	53.89	53.96	112.31	90.01	90.11	120.09
		53.92	53.92	112.31	90.00	90.00	120.00
primitive tetragonal	7.85%	53.84	53.89	112.31	89.89	90.12	119.89
		53.87	53.87	112.31	90.00	90.00	90.00
I centred tetragonal	4.23%	53.96	53.89	230.88	96.57	83.31	59.91
		53.92	53.92	230.88	90.00	90.00	90.00
primitive orthorhombic	7.85%	53.84	53.89	112.31	89.89	90.12	119.89
		53.84	53.89	112.31	90.00	90.00	90.00
C centred orthorhombic	0.08%	53.84	93.44	112.31	89.94	90.12	89.92
		53.84	93.44	112.31	90.00	90.00	90.00
I centred orthorhombic	4.23%	53.89	53.96	230.88	83.31	83.43	120.09
		53.89	53.96	230.88	90.00	90.00	90.00
F centred orthorhombic	4.42%	53.84	93.44	230.88	89.93	76.64	89.92
		53.84	93.44	230.88	90.00	90.00	90.00
primitive monoclinic	0.08%	53.84	112.31	53.96	90.01	120.02	89.88
		53.84	112.31	53.96	90.00	120.02	90.00
C centred monoclinic	0.02%	93.25	53.96	112.31	90.01	90.14	90.05
		93.25	53.96	112.31	90.00	90.14	90.00
primitive triclinic	0.00%	53.84	53.89	112.31	89.89	89.88	60.11

Volume of the primitive cell	282533.						
autoindex unit cell		53.84	53.84	112.31	90.00	90.00	120.00
crystal rotx, roty, rotz	98.764	-106.274	-56.816				
Autoindex Xbeam, Ybeam	100.16	96.80					

Table 463-2: Completeness of data for BLGA at pH 7.1

	I/Sigma in resolution shells: % of reflections with I / σI less than								
Resolution shell (Å)	0	1	2	3	5	10	20	>20	total
99.00 6.08	0.0	0.5	1.1	1.2	2.0	3.2	4.8	84.3	89.1
6.08 4.83	0.4	0.6	1.7	2.1	2.9	4.9	9.9	84.3	94.2
4.83 4.22	0.8	1.0	1.2	2.1	2.7	4.6	10.1	84.9	95.0
4.22 3.83	0.2	1.4	3.0	4.2	5.4	8.7	24.3	72.6	96.8
3.83 3.56	0.6	1.5	2.9	4.4	7.1	13.6	31.3	65.6	96.9
3.56 3.35	2.1	3.5	4.4	5.8	8.6	15.9	48.0	49.3	97.3
3.35 3.18	1.9	3.4	7.2	10.8	14.6	27.1	64.4	33.1	97.5
3.18 3.04	2.3	4.1	7.1	8.7	14.5	32.4	81.3	17.0	98.3
3.04 2.92	2.5	5.9	9.8	13.3	25.0	54.5	97.3	0.0	97.3
2.92 2.82	5.0	8.8	16.2	22.1	32.6	62.6	95.8	0.2	96.0

(continued)

Table 463-2: Completeness of data for BLGA at pH 7.1

		I/Sigma in resolution shells: % of reflections with I / σI less than								
Resolution shell (Å)		0	1	2	3	5	10	20	>20	total
2.82	2.73	5.2	11.7	19.7	25.2	39.4	68.8	97.5	0.4	97.9
2.73	2.66	6.5	15.3	26.8	36.9	52.0	88.1	97.1	0.0	97.1
2.66	2.59	8.2	17.4	30.2	41.7	65.3	97.8	97.8	0.0	97.8
2.59	2.52	9.1	19.5	33.3	48.7	77.7	98.0	98.0	0.0	98.0
2.52	2.47	8.1	17.8	32.6	50.4	80.2	97.6	97.6	0.0	97.6
2.47	2.41	10.7	23.7	41.6	62.2	94.6	97.8	97.8	0.0	97.8
2.41	2.36	12.4	24.5	42.6	65.6	92.1	96.9	96.9	0.0	96.9
2.36	2.32	13.3	26.3	49.4	76.1	94.9	95.3	95.3	0.0	95.3
2.32	2.28	14.0	27.5	50.3	73.6	95.4	96.2	96.2	0.0	96.2
2.28	2.24	15.0	28.5	51.4	71.5	92.3	95.6	95.6	0.0	95.6
All hkl		5.8	11.9	21.2	30.6	44.0	56.8	70.7	25.6	96.4

Table 463-3: Data merging statistics for BLGA at pH 7.1

Resolution shell (Å)		I	error	stat.	χ^2	R	R^2
99.00	6.08	6597.8	138.2	70.1	1.105	0.017	0.019
6.08	4.83	3811.4	87.4	40.4	1.135	0.023	0.024
4.83	4.22	5372.8	136.9	64.6	1.020	0.022	0.025
4.22	3.83	2861.2	92.0	46.3	1.033	0.030	0.031
3.83	3.56	1947.6	73.3	38.7	1.114	0.037	0.037
3.56	3.35	1285.9	59.9	31.1	1.029	0.042	0.041
3.35	3.18	910.9	51.5	27.8	1.026	0.051	0.050
3.18	3.04	651.1	43.5	25.7	1.031	0.064	0.063
3.04	2.92	424.6	42.4	22.9	1.008	0.089	0.078
2.92	2.82	309.0	33.8	22.3	1.099	0.114	0.095
2.82	2.73	224.8	27.4	21.7	1.007	0.127	0.130
2.73	2.66	157.1	28.2	21.3	1.060	0.186	0.168
2.66	2.59	129.2	30.6	21.1	0.918	0.225	0.201
2.59	2.52	109.7	30.9	21.8	1.136	0.305	0.271
2.52	2.47	97.6	30.5	22.2	1.052	0.326	0.311
2.47	2.41	85.8	32.8	22.0	0.977	0.360	0.328
2.41	2.36	72.5	29.0	21.9	1.122	0.434	0.404
2.36	2.32	70.5	33.3	22.6	1.024	0.451	0.378
2.32	2.28	67.8	32.2	22.6	1.150	0.497	0.459
2.28	2.24	58.2	28.4	22.9	1.193	0.540	0.495
All reflections		1299.2	53.8	30.7	1.061	0.038	0.026

4.6.4 Data processing for BLGA at pH 6.2

As for BLGA at pH 8.2, the primitive hexagonal lattice is the suggested lattice by DENZO. The index result of 14 Bravais lattices for the first image of this crystal is shown in Table 464-1. For BLGA at pH 6.2, the distortion index (0.39%) is somewhat higher than those for BLGA at pH 8.2 and 7.1 (Table 464-1). The overall completeness of data is 99.1% to 2.56 Å (Table 464-2), with an overall value for R of 0.066 for averaging

the data (Table 464-3).

Table 464-1: Auto-indexing by DENZO for BLGA crystal at pH 6.2

Lattice	Metric tensor distortion index	Best cell (symmetrized) Best cell (without symmetry restraints)					
primitive cubic	51.94%	93.52	54.41	111.84	90.35	90.15	30.11
		89.84	89.84	89.84	90.00	90.00	90.00
I centred cubic	40.05%	111.84	94.74	124.06	90.32	25.74	90.46
		110.87	110.87	110.87	90.00	90.00	90.00
F centred cubic	34.26%	145.98	145.61	124.07	46.25	46.66	79.80
		138.93	138.93	138.93	90.00	90.00	90.00
primitive rhombohedral	8.90%	124.06	124.07	124.07	25.08	25.33	44.28
		124.07	124.07	124.07	32.67	32.67	32.67
		76.51	76.51	352.03	90.00	90.00	120.00
primitive hexagonal	0.39%	54.59	54.41	111.84	90.35	89.55	120.75
		54.50	54.50	111.84	90.00	90.00	120.00
primitive tetragonal	7.86%	53.88	54.41	111.84	90.35	89.90	60.54
		54.14	54.14	111.84	90.00	90.00	90.00
I centred tetragonal	4.26%	53.88	54.41	229.84	83.37	96.68	60.54
		54.14	54.14	229.84	90.00	90.00	90.00
primitive orthorhombic	7.85%	53.88	54.41	111.84	90.35	89.90	60.54
		53.88	54.41	111.84	90.00	90.00	90.00
C centred orthorhombic	0.33%	54.59	93.52	111.84	90.15	89.55	90.64
		54.59	93.52	111.84	90.00	90.00	90.00
I centred orthorhombic	4.25%	53.88	54.41	229.84	83.37	96.68	60.54
		53.88	54.41	229.84	90.00	90.00	90.00
F centred orthorhombic	4.47%	54.59	93.52	229.84	89.99	103.29	90.64
		54.59	93.52	229.84	90.00	90.00	90.00
primitive monoclinic	0.21%	53.88	111.84	54.41	90.35	119.46	90.10
		53.88	111.84	54.41	90.00	119.46	90.00
C centred monoclinic	0.09%	94.74	53.88	111.84	90.10	90.46	89.78
		94.74	53.88	111.84	90.00	90.46	90.00
primitive triclinic	0.00%	53.88	54.41	111.84	90.35	90.10	119.46

Volume of the primitive cell 285450.
autoindex unit cell 54.17 54.17 111.84 90.00 90.00 120.00
crystal rotx, roty, rotz 227.615 5.480 28.105
Autoindex Xbeam, Ybeam 100.16 96.94

Table 464-2: Completeness of data for BLGA at pH 6.2

Resolution shell (Å)		I/Sigma in resolution shells: % of reflections with I / σI less than 0	1	2	3	5	10	20	>20	total
99.00	6.83	1.7	2.5	3.0	3.5	3.7	7.2	14.7	77.1	91.8
6.83	5.42	1.1	2.2	3.9	5.3	8.4	15.7	32.5	66.7	99.2
5.42	4.74	1.1	1.7	3.4	4.2	6.5	14.1	33.3	66.4	99.7
4.74	4.30	0.9	2.6	2.6	4.0	8.3	13.2	34.4	65.6	100.0
4.30	3.99	0.6	2.3	4.0	7.5	11.6	24.3	48.6	51.4	100.0
3.99	3.76	1.8	5.3	6.2	7.9	12.6	25.6	57.9	41.8	99.7
3.76	3.57	0.9	2.4	4.8	7.8	15.3	30.3	69.1	30.6	99.7
3.57	3.42	2.3	6.2	9.4	13.1	23.0	42.6	82.7	17.0	99.7
3.42	3.28	3.0	8.4	13.5	20.7	28.4	47.3	85.0	14.7	99.7
3.28	3.17	3.8	9.3	15.7	20.1	34.2	61.3	94.6	5.4	100.0

(continued)

Table 464-2: Completeness of data for BLGA at pH 6.2

Resolution shell (Å)	I/Sigma in resolution shells: % of reflections with I / σI less than 0	1	2	3	5	10	20	>20	total
3.17 3.07	2.8	7.5	17.4	25.7	39.2	70.4	99.4	0.6	100.0
3.07 2.98	5.4	12.7	21.8	31.0	46.8	90.5	99.7	0.0	99.7
2.98 2.91	6.9	17.1	26.6	38.3	53.9	87.4	99.7	0.0	99.7
2.91 2.83	6.9	17.4	29.7	40.2	58.6	94.0	99.7	0.0	99.7
2.83 2.77	7.6	18.7	32.9	46.1	73.2	99.7	99.7	0.0	99.7
2.77 2.71	9.7	21.5	36.4	49.2	80.7	98.8	98.8	0.0	98.8
2.71 2.66	11.0	28.0	46.7	62.8	91.7	99.4	99.4	0.0	99.4
2.66 2.61	13.9	31.5	46.3	70.6	93.8	100.0	100.0	0.0	100.0
2.61 2.56	12.8	30.6	52.2	74.7	94.7	97.2	97.2	0.0	97.2
All hkl	4.8	11.7	19.4	27.4	40.4	57.7	75.0	24.1	99.1

Table 464-3: Data merging statistics for BLGA at pH 6.2

Resolution shell (Å)		I	error	stat.	χ^2	R	R^2
99.00	6.83	1526.5	38.3	21.7	1.048	0.020	0.023
6.83	5.42	677.2	24.4	14.4	1.086	0.035	0.034
5.42	4.74	1136.6	40.3	20.0	1.087	0.037	0.051
4.74	4.30	1439.8	51.3	24.3	1.078	0.036	0.037
4.30	3.99	769.6	33.2	20.2	1.098	0.047	0.044
3.99	3.76	601.4	29.7	19.8	1.086	0.056	0.053
3.76	3.57	541.1	30.4	20.0	1.056	0.062	0.061
3.57	3.42	372.7	26.6	18.3	1.087	0.081	0.076
3.42	3.28	299.3	23.7	18.3	1.076	0.091	0.080
3.28	3.17	215.1	21.4	17.1	1.032	0.123	0.108
3.17	3.07	191.8	23.0	17.2	1.032	0.143	0.128
3.07	2.98	139.7	23.0	16.7	1.131	0.200	0.209
2.98	2.91	126.8	21.9	17.1	1.036	0.211	0.187
2.91	2.83	106.4	21.2	17.0	1.088	0.256	0.223
2.83	2.77	87.1	22.9	17.1	1.017	0.309	0.294
2.77	2.71	72.1	22.4	17.8	1.097	0.391	0.374
2.71	2.66	57.6	22.4	17.7	1.046	0.439	0.416
2.66	2.61	53.2	22.8	18.2	1.040	0.466	0.422
2.61	2.56	47.4	23.7	18.7	1.055	0.530	0.495
All reflections		459.3	27.7	18.5	1.067	0.066	0.048

4.6.5 Data processing for BLGB at pH 7.1

As for BLGA at pH 8.2, DENZO provides the index result of 14 Bravais lattices for the first image of this crystal in Table 465-1. The primitive hexagonal lattice with a low distortion index of 0.09% is the suggested lattice type. The overall completeness of data is 91.2% Å (Table 465-2), with an overall value of R of 0.086 for averaging the data (Table 465-3).

Table 465-1: Auto-indexing by DENZO for BLGB at pH 7.1

Lattice	Metric tensor distortion index	Best cell (symmetrized) Best cell (without symmetry restraints)					
primitive cubic	52.57%	54.19	93.75	112.89	89.80	89.88	30.00
		90.31	90.31	90.31	90.00	90.00	90.00
I centred cubic	40.95%	93.91	125.00	112.89	25.64	90.01	90.01
		111.34	111.34	111.34	90.00	90.00	90.00
F centred cubic	34.44%	125.14	146.99	146.50	79.41	46.20	45.99
		139.91	139.91	139.91	90.00	90.00	90.00
primitive rhombohedral	8.81%	125.12	125.00	125.14	25.02	44.08	25.02
		125.09	125.09	125.09	32.49	32.49	32.49
		76.67	76.67	355.14	90.00	90.00	120.00
primitive hexagonal	0.09%	54.19	54.19	112.89	90.10	90.12	120.11
		54.19	54.19	112.89	90.00	90.00	120.00
primitive tetragonal	7.86%	54.10	54.19	112.89	89.88	90.22	119.94
		54.14	54.14	112.89	90.00	90.00	90.00
I centred tetragonal	4.21%	54.19	54.19	231.97	96.57	83.42	59.89
		54.19	54.19	231.97	90.00	90.00	90.00
primitive orthorhombic	7.86%	54.10	54.19	112.89	89.88	90.22	119.94
		54.10	54.19	112.89	90.00	90.00	90.00
C centred orthorhombic	0.10%	54.10	93.91	112.89	89.99	90.22	90.00
		54.10	93.91	112.89	90.00	90.00	90.00
I centred orthorhombic	4.21%	54.19	54.19	231.97	83.42	83.43	120.11
		54.19	54.19	231.97	90.00	90.00	90.00
F centred orthorhombic	4.39%	54.10	93.91	231.97	89.99	76.73	90.00
		54.10	93.91	231.97	90.00	90.00	90.00
primitive monoclinic	0.12%	54.19	112.89	54.19	89.90	120.11	89.88
		54.19	112.89	54.19	90.00	120.11	90.00
C centred monoclinic	0.01%	54.10	93.91	112.89	89.99	90.22	90.00
		54.10	93.91	112.89	90.00	90.22	90.00
primitive triclinic	0.00%	54.10	54.19	112.89	89.88	89.78	60.06

Volume of the primitive cell 286749.
autoindex unit cell 54.13 54.13 112.89 90.00 90.00 120.00
crystal rotx, roty, rotz 64.685 -68.331 90.954
Autoindex Xbeam, Ybeam 100.26 97.06

Table 465-2: Completeness of data for BLGB at pH 7.1

Resolution shell (Å)	I/Sigma in resolution shells: % of reflections with I / σI less than 0	1	2	3	5	10	20	>20	total
99.00 5.35	0.0	0.2	0.2	0.5	1.0	17.0	77.3	3.5	80.8
5.35 4.25	0.4	0.8	1.1	1.6	2.4	19.3	88.1	0.0	88.1
4.25 3.71	1.1	2.0	2.6	3.4	4.3	33.6	93.4	0.0	93.4
3.71 3.37	1.5	2.5	3.9	5.2	9.3	45.9	92.8	0.0	92.8
3.37 3.13	1.9	4.4	6.4	8.4	13.0	53.0	94.9	0.0	94.9
3.13 2.94	3.3	6.1	9.2	13.5	23.4	85.3	96.7	0.0	96.7
2.94 2.80	5.3	9.6	16.3	23.3	41.7	95.1	97.2	0.0	97.2
2.80 2.68	6.2	13.6	21.1	31.4	54.6	97.5	97.7	0.0	97.7
2.68 2.57	6.6	15.3	27.9	42.4	82.5	97.9	97.9	0.0	97.9
2.57 2.48	8.5	18.2	32.3	54.0	89.5	96.9	96.9	0.0	96.9
2.48 2.41	9.5	23.7	42.3	65.4	94.8	94.8	94.8	0.0	94.8
2.41 2.34	12.3	26.8	49.2	70.0	90.4	90.9	90.9	0.0	90.9
2.34 2.28	12.1	29.7	57.7	77.2	84.9	84.9	84.9	0.0	84.9
2.28 2.22	11.7	26.8	41.3	57.0	70.8	71.8	71.8	0.0	71.8
All hkl	5.6	12.6	21.8	31.7	46.4	69.3	90.9	0.3	91.2

Table 465-3: Data merging statistics for BLGB at pH 7.1

Resolution shell (Å)		I	error	stat.	χ^2	R	R^2
99.00	5.35	12857.6	1000.8	113.6	1.039	0.066	0.072
5.35	4.25	10666.7	914.6	88.8	1.064	0.080	0.094
4.25	3.71	5812.6	522.3	67.9	1.101	0.084	0.100
3.71	3.37	3193.4	308.7	54.1	1.122	0.091	0.102
3.37	3.13	1893.3	190.9	47.3	1.147	0.097	0.108
3.13	2.94	1072.8	143.7	44.4	1.081	0.117	0.118
2.94	2.80	686.2	110.0	43.8	1.004	0.138	0.121
2.80	2.68	401.1	80.3	42.4	0.979	0.176	0.157
2.68	2.57	307.4	83.5	42.7	1.050	0.228	0.189
2.57	2.48	235.7	76.4	41.3	0.950	0.279	0.248
2.48	2.41	178.3	71.0	39.5	1.059	0.341	0.291
2.41	2.34	163.0	69.6	40.8	0.951	0.315	0.240
2.34	2.28	133.9	72.0	40.6	0.992	0.363	0.281
2.28	2.22	118.2	55.4	39.6	1.058	0.351	0.263
All reflections		2725.7	266.8	53.5	1.048	0.086	0.085

4.6.6 Data processing for complex of BLGA with BrC12 at pH 7.3

As for BLGA at pH 8.2, the primitive hexagonal lattice is the suggested lattice type for the BLGA-BrC12 complex at pH 7.3 (Table 466-1). The overall completeness of data for BLGA-BrC12 is impressive, at 99.5% to 2.23 Å, with an overall value for R of 0.048 for averaging the data (Table 466-3).

Table 466-1: Auto-indexing by DENZO for BLGA-BrC12 at pH 7.3

Lattice	Metric tensor distortion index	Best cell (symmetrized) Best cell (without symmetry restraints)					
primitive cubic	52.27%	93.47	54.06	112.02	90.13	90.07	30.10
		89.83	89.83	89.83	90.00	90.00	90.00
I centred cubic	40.69%	93.75	112.02	124.28	25.79	89.95	90.07
		110.73	110.73	110.73	90.00	90.00	90.00
F centred cubic	34.32%	145.99	145.81	124.35	46.29	46.30	79.68
		139.08	139.08	139.08	90.00	90.00	90.00
primitive rhombohedral	8.88%	124.36	124.28	124.35	25.09	25.11	44.17
		124.33	124.33	124.33	32.56	32.56	32.56
		76.35	76.35	352.90	90.00	90.00	120.00
primitive hexagonal	0.10%	54.06	54.00	112.02	90.00	90.13	119.77
		54.03	54.03	112.02	90.00	90.00	120.00
primitive tetragonal	7.88%	54.00	54.06	112.02	90.13	90.00	60.23
		54.03	54.03	112.02	90.00	90.00	90.00
I centred tetragonal	4.27%	54.00	54.06	230.39	83.33	96.77	60.23
		54.03	54.03	230.39	90.00	90.00	90.00
primitive orthorhombic	7.88%	54.00	54.06	112.02	90.13	90.00	60.23
		54.00	54.06	112.02	90.00	90.00	90.00
C centred orthorhombic	0.08%	54.22	93.47	112.02	90.07	89.87	90.08
		54.22	93.47	112.02	90.00	90.00	90.00
I centred orthorhombic	4.27%	54.00	54.06	230.39	83.33	96.77	60.23
		54.00	54.06	230.39	90.00	90.00	90.00

(continued)

Table 466-1: Auto-indexing by DENZO for BLGA-BrC12 at pH 7.3

Lattice	Metric tensor distortion index	Best cell (symmetrized) Best cell (without symmetry restraints)					
F centred orthorhombic	4.47%	54.06	93.75	230.36	90.01	76.55	89.73
		54.06	93.75	230.36	90.00	90.00	90.00
primitive monoclinic	0.07%	54.00	112.02	54.06	90.13	119.77	90.00
		54.00	112.02	54.06	90.00	119.77	90.00
C centred monoclinic	0.06%	54.22	93.47	112.02	89.93	90.13	90.08
		54.22	93.47	112.02	90.00	90.13	90.00
primitive triclinic	0.00%	54.00	54.06	112.02	90.13	90.00	119.77

Volume of the primitive cell 283870.
autoindex unit cell 54.19 54.19 112.02 90.00 90.00 120.00
crystal rotx, roty, rotz -70.555 140.272 0.337
Autoindex Xbeam, Ybeam 100.26 97.14

Table 466-2: Completeness of data for BLGA-BrC12 at pH 7.3

Resolution shell (Å)	I/Sigma in resolution shells: % of reflections with I / σI less than 0	1	2	3	5	10	20	>20	total
99.00 5.50	0.5	1.6	1.9	2.4	4.3	6.8	15.8	79.5	95.2
5.50 4.37	1.3	1.6	2.5	3.2	3.8	7.7	18.9	80.6	99.4
4.37 3.81	2.9	4.2	5.5	6.1	8.0	12.6	31.9	68.1	100.0
3.81 3.46	1.4	3.4	5.7	6.6	9.7	15.2	40.3	59.7	100.0
3.46 3.22	2.6	4.4	6.9	9.1	16.2	28.0	76.2	23.8	100.0
3.22 3.03	2.7	6.0	9.2	12.7	21.1	40.0	87.9	12.1	100.0
3.03 2.87	4.4	9.0	14.7	20.9	31.0	60.0	98.0	2.0	100.0
2.87 2.75	5.2	12.1	20.9	28.8	43.4	71.2	98.7	1.1	99.8
2.75 2.64	8.2	16.1	28.0	37.2	51.7	82.0	99.4	0.6	100.0
2.64 2.55	8.2	18.6	31.8	42.7	58.6	84.8	99.5	0.5	100.0
2.55 2.47	8.1	22.7	37.7	51.3	76.8	99.7	100.0	0.0	100.0
2.47 2.40	12.4	28.5	42.8	55.8	77.9	95.6	99.7	0.3	100.0
2.40 2.34	14.0	29.8	49.2	62.8	82.2	94.8	99.7	0.3	100.0
2.34 2.28	15.7	34.7	55.0	68.0	84.7	96.0	99.5	0.5	100.0
2.28 2.23	16.0	37.4	60.3	71.2	86.2	95.2	98.2	0.3	98.6
All hkl	6.8	15.1	24.3	31.3	42.8	58.2	76.5	23.0	99.5

Table 466-3: Data merging statistics for BLGA-BrC12 at pH 7.3

Resolution shell (Å)	I	error	stat.	χ^2	R	R^2
99.00 5.50	2711.4	79.2	60.4	1.043	0.018	0.018
5.50 4.37	3222.0	103.6	76.7	1.111	0.022	0.023
4.37 3.81	1839.4	68.2	45.5	1.110	0.029	0.030
3.81 3.46	1214.5	50.5	33.4	1.098	0.036	0.038
3.46 3.22	766.0	46.3	27.0	0.999	0.048	0.046
3.22 3.03	462.3	34.2	23.9	1.094	0.070	0.063
3.03 2.87	311.4	30.6	22.4	1.078	0.096	0.083
2.87 2.75	211.9	26.1	21.7	1.159	0.133	0.120
2.75 2.64	149.3	24.5	22.0	1.154	0.178	0.162
2.64 2.55	125.6	24.2	22.7	1.085	0.201	0.190
2.55 2.47	97.6	28.1	23.5	1.085	0.303	0.277
2.47 2.40	75.9	24.7	24.2	1.112	0.355	0.342
2.40 2.34	67.8	25.8	25.8	1.167	0.433	0.402
2.34 2.28	63.0	26.9	26.9	1.073	0.460	0.408
2.28 2.23	56.2	27.6	27.6	1.117	0.548	0.498
All reflections	787.0	42.0	32.7	1.099	0.048	0.029

4.6.7 Summary of data processing

X-ray diffraction data have been collected from a number of BLG crystals including BLGA at three different pH, BLGB, BLGA-BrC12 complex and a possible BLGA-SDS complex. The BLGA-SDS crystal diffracted to only 5.5 Å with space group I4. While several data sets were collected and processed, only data from the best diffracting crystals were used in subsequent structure determinations and refinements. The diffraction data files, which were used in later structure determinations, are listed in Table 467-1. The parameters of these crystals are listed in Table 467-2.

Table 467-1 : Data sets and backup tape name

File name	blgza2	blgza9	blgza10	blgzb1	Br1
pH	pH 8.2	pH 6.2	pH 7.1	pH 7.1	7.3
variant	A	A	A	B	A
operators	Maria/Bin	Bin	Bin	Maria/Bin	Bin
tape	BLG6	BLG7	BLG7	BLG6	BLG8
date	2/10/96	06/05/97	08/05/97	4/10/96	26/03/98
oscillation start	40	-90	-110	-40	170
oscillation end	80.5	-30	-50	8	215
oscillation range	1.5	1.5	1.5	1.2	1.5

Table 467-2 : Crystal parameters

File name		blgza9	blgza10	blgza2	blgzb1	Br1
Space group		$P3_221$	$P3_221$	$P3_221$	$P3_221$	$P3_221$
Unit cell	a Å	53.75	53.96	54.29	54.29	54.03
	b Å	53.75	53.96	54.29	54.29	54.03
	c Å	111.56	112.41	113.18	113.16	112.18
	α °	90.	90.	90.	90.	90.
	ß °	90.	90.	90.	90.	90.
	γ °	120.	120.	120.	120.	120.
pH		6.2	7.1	8.2	7.1	7.3
Unit cell volume	$Å^3$	279111	283429	288901	288854	283564
Lattice type code		Z	Z	Z	Z	Z

4.7 Initial structures for bovine BLG

4.7.1 Structure determinations of BLG by molecular replacement

In the first chapter, several BLG structures were reviewed. The low- and medium-resolution BLG structures have unavoidable uncertainty. While the high-resolution BLG structures are reliable in the calyx fold, the core part of the BLG structure, there are problems in several loops, which are not well defined in the high-resolution structures determined in the X and Y lattices. In many cases, the loops of protein molecules are functionally important. Therefore, a reasonable understanding about BLG requires clarity of the loops. Based on these facts, the crystals of bovine BLG in lattice Z provide another possibility to obtain the best understanding of the loops. Molecular replacement methods are the best choice to solve efficiently the structures of bovine BLG in lattice Z, since high resolution, reliable structures now exist for at least the core part of bovine BLG. The principles of molecular replacement methods were reviewed in Chapter 3. From a number of programs for MR methods, program AMoRe [Navaza, 1994], which offers an integrated molecular replacement methods package including rotation, translation and fitting procedures, was chosen.

4.7.2 Molecular replacement for BLGA at pH 8.2

With the program AMoRe, the structure of BLGA at pH 8.2 in lattice Z was solved by molecular replacement methods, where the search model (the known molecule) was the redetermined structure of BLGA in the orthorhombic crystal form (lattice Y) [Bewley *et al.*, 1998]. This structure model (in its early stage of development) included only the residues 1-61, 66-110 and 114-154. The reflections used in calculating the Patterson map were in the resolution range 10 - 3 Å. The ten best solutions are listed in Table 472-1. The top solution is unambiguously the best, and the corresponding rotation and translation matrices are given in Table 472-2.

Table 472-1: Molecular replacement result for BLGA at pH 8.2

Label	α	β	γ	x	y	z	coefficients	R
#11	103.9	84.7	12.1	0.0719	0.3482	0.1921	65.2	39.5
#11	21.6	87.6	13.2	0.7239	0.6644	0.4418	36.6	52.7
#11	19.9	86.3	13.8	0.7549	0.9923	0.1123	36.1	52.1
#11	54.1	65.0	299.1	0.6077	0.1246	0.1565	36.0	52.6
#11	82.4	82.5	331.8	0.0698	0.3345	0.2660	35.3	52.0
#11	38.3	82.6	14.9	0.0449	0.6489	0.1914	35.1	53.0
#11	13.7	89.9	356.5	0.4334	0.6535	0.1824	34.9	52.6
#11	94.3	62.9	78.2	0.2896	0.1029	0.3201	34.9	52.6
#11	94.2	62.3	79.2	0.3009	0.1043	0.0910	35.4	52.7
#11	39.4	67.4	78.4	0.2445	0.1331	0.0538	34.9	52.7

Table 472-2: Translation and rotation matrices for BLGA at pH 8.2

	-0.54175	-0.51951	-0.66078		-10.82
initial za2 =	-0.42589	0.84740	-0.31706	* blga in Y +	41.00
	0.72466	0.10965	-0.68033		52.05

4.7.3 Molecular replacement for BLGA at pH 7.1

With the programme AMoRe, the structure of BLGA at pH 7.1 in lattice Z was solved by molecular replacement methods, where the search model (the known molecule) was the fully determined structure of BLGA in lattice Z at pH 8.2 (za2, see previous section). This structure included all 162 residues of bovine BLGA. The reflections used in calculating the Patterson map were in the resolution range 10 - 3 Å. The ten best solutions are listed in Table 473-1. Again the first solution is unambiguously the best, and the corresponding rotation and translation matrices to move the search model (za2) are given in Table 473-2.

Table 473-1: Molecular replacement results for BLGA at pH 7.1

Label	α	β	γ	x	y	z	coefficients	R
#11	94.3	49.8	2.2	0.0865	0.3282	0.1994	86.2	25.0
#11	92.3	6.0	103.2	0.9707	0.3950	0.0993	39.7	51.8
#11	90.7	5.6	104.8	0.9815	0.0283	0.4287	39.1	51.9
#11	115.6	36.6	62.7	0.6931	0.9479	0.0894	39.7	51.9
#11	116.3	36.6	63.3	0.3421	0.8745	0.1693	39.2	52.1
#11	47.9	29.8	114.8	0.9212	0.8716	0.0461	39.3	52.2
#11	46.4	30.5	115.5	0.9862	0.5143	0.2769	38.7	52.4
#11	100.5	5.7	92.0	0.9887	0.0418	0.0942	39.9	51.7
#11	93.1	5.5	99.4	0.9713	0.3914	0.0978	38.8	51.9
#11	78.6	58.4	331.5	0.9759	0.2576	0.2086	39.2	52.5

Table 473-2: Translation and rotation matrices for BLGA at pH 7.1

	1.00000	0.00061	-0.00243		0.33
initial za10 =	-0.00063	0.99996	-0.00861 * za2	+	-0.03
	0.00243	0.00861	0.99996		-0.35

4.7.4 Molecular replacement for BLGA at pH 6.2

With the programme AMoRe, the structure of BLGA at pH 6.2 in lattice Z was solved by molecular replacement methods, where the search model (the known molecule) was the fully determined structure of BLGA in lattice Z at pH 8.2 (za2). This structure included all 162 residues of bovine BLGA. The reflections used in calculating the Patterson map were in the resolution range 10 - 3 Å. The nine best solutions are listed in Table 474-1. The first solution is unambiguously the best, although the second solution is not unattractive in terms of R factor and the correlation coefficient. However, after applying the corresponding rotation and translation matrices, the BLG molecules overlap each other, thus excluding the second choice. The rotation and translation matrices to move the search model (za2) are given in Table 474-2.

Table 474-1: Molecular replacement results for BLGA at pH 6.2

Label	α	β	γ	x	y	z	co-efficients	R
#11	77.4	47.1	6.2	0.0911	0.3257	0.1994	88.9	21.0
#11	77.3	47.3	6.2	0.7579	0.6591	0.0567	57.9	40.5
#11	54.5	48.4	6.9	0.7007	0.6658	0.4506	37.1	50.6
#11	10.3	48.0	7.2	0.7069	0.6078	0.4728	36.3	51.2
#11	42.0	54.0	7.8	0.7117	0.6499	0.0367	34.9	52.5
#11	103.7	14.8	194.6	0.7846	0.6643	0.4550	34.2	51.7
#11	116.8	75.1	232.0	0.6480	0.8383	0.5016	33.3	51.2
#11	89.8	64.1	279.8	0.1832	0.7402	0.3815	32.6	52.5
#11	97.9	14.9	186.8	0.8010	0.6966	0.0423	32.7	52.0

Table 474-2: Translation and rotation matrices for BLGA at pH 6.2

	1.00000	0.00025	-0.00095		0.02
initial za9 =	-0.00025	1.00000	0.00000 * za2	+	0.01
	0.00095	0.00000	1.00000		0.00

4.7.5 Molecular replacement for BLGB at pH 7.1

With the programme AMoRe, the structure of BLGB at pH 7.1 in lattice Z was solved by molecular replacement methods, where the search model (the known molecule) was the fully determined structure of BLGA in lattice Z at pH 7.1. This structure included all 162 residues of bovine BLGA. The reflections used in calculating the Patterson map were in the range 8 - 3 Å. The best ten solutions are listed in Table 475-1. The first one is unambiguously the best solution. Similar to BLGA at pH 6.2, the second solution is quite alluring, but again leads to overlapping molecules. The rotation and translation matrices to move the search model (za10) are given in Table 475-2.

Table 475-1 : Molecular replacement results for BLGB at pH 7.1

Label	α	β	γ	x	y	z	coefficients	R
#11	37.5	49.6	1.5	0.3303	0.2462	0.3662	84.2	26.4
#11	37.4	49.4	1.8	0.9973	0.5793	0.4749	54.9	42.9
#11	88.2	72.5	219.1	0.1698	0.4841	0.4006	37.6	51.6
#11	88.7	72.7	218.9	0.5044	0.1518	0.1313	36.3	52.3
#11	75.0	21.4	8.4	0.5670	0.7207	0.4881	37.5	52.2
#11	77.7	20.8	5.8	0.4837	0.8720	0.2994	37.4	51.9
#11	106.0	29.5	117.9	0.6895	0.6075	0.4401	38.2	51.2
#11	5.9	34.3	10.3	0.9756	0.7497	0.1102	36.9	51.8
#11	43.1	52.3	312.1	0.1039	0.3518	0.3161	37.2	52.0
#11	43.8	52.6	311.4	0.7610	0.6878	0.1191	37.4	51.7

Table 475-2 : Rotation and translation matrices for BLGB at pH 7.1

	0.50098	0.86542	0.00796			-0.06
initial zb1 =	-0.86546	0.50098	0.00160	* za10	+	0.24
	-0.00260	-0.00770	0.99997			19.11

4.7.6 Molecular replacement for BLGA-BrC12 at pH 7.3

With the programme AMoRe, the structure of BLGA co-crystallized with 12-bromododecanoic acid (BrC12) at pH 7.3 in lattice Z was solved by molecular replacement methods, where the search model (the known molecule) was the water-excluded, fully determined structure of BLGA in lattice Z at pH 7.1 (za10). This structure included all 162 residues of bovine BLGA and with the loop EF in the open conformation. The reflections used in calculating the Patterson map were in the range 10

- 3 Å. The top ten solutions are listed in Table 476-1. The first one is the unambiguous solution. The rotation and translation matrices, to move the search model are given in Table 476-2. The electron density map (difference map) calculated immediately after molecular replacement showed a strong positive peak of more than five times the rms level of the map inside the BLG calyx; there was the continuous density at more than 2σ level for a length of around 15 Å. This suggested that the ligand $BrCH_2$-$(CH_2)_{10}COOH$ was present in the crystal and was bound inside the calyx, as shown in Figure 476-1. Further refinement of the initial structure retained this characteristic. The density was modelled well by assuming a standard geometry for $BrCH_2(CH_2)_{10}COOH$. The ligand refined well providing unequivocal evidence for its presence.

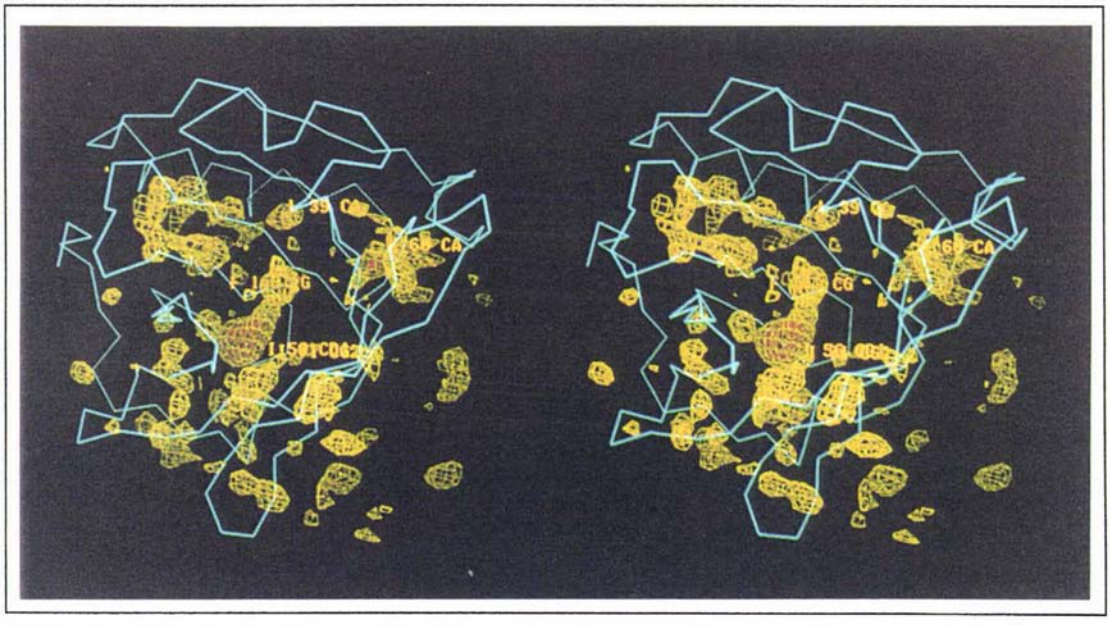


Figure 476-1: Difference electron density map of BLGA-BrC12, red contoured at +5 σ, yellow +2 σ

Table 476-1 : Molecular replacement results for BLGA-BrC12

Label	α	β	γ	x	y	z	coefficients	R
#11	34.9	49.6	3.4	0.3266	0.2423	0.3653	85.5	25.6
#11	88.9	48.9	3.6	0.0719	0.7486	0.3669	37.7	52.3
#11	108.0	21.1	246.1	0.1168	0.7430	0.4395	35.1	53.1
#11	106.0	16.6	226.2	0.1020	0.7769	0.3158	35.3	53.3
#11	105.2	16.7	227.0	0.1635	0.8037	0.4834	34.9	53.6
#11	3.0	22.0	242.8	0.3500	0.2302	0.4186	34.9	52.8
#11	5.8	21.2	241.1	0.6330	0.8853	0.1681	34.4	54.1
#11	103.1	51.2	267.7	0.8485	0.7817	0.4270	35.0	53.3
#11	102.3	51.6	267.3	0.1952	0.4271	0.4694	34.9	53.2
#11	59.8	32.3	57.6	0.7599	0.2111	0.4263	34.5	53.2

Table 476-2 : Rotation and translation matrices for BLGA-BrC12

	0.50383	0.86380	-0.00190			0.06
initial br1 =	-0.86379	0.50382	-0.00594	* zal0	+	0.13
	-0.00417	0.00464	0.99998			18.52

4.8 Structure refinement

4.8.1 X-PLOR and TURBO-FRODO

After the molecular replacement, the resulting coarse solution for the protein structure must be refined to get the better congruence between the model and the diffraction data. A number of programs can do the structure refinement, such as SHELXL [Sheldrick, 1997], TNT [Tronrud, 1995] and X-PLOR [Brünger, 1988], etc. We chose X-PLOR mainly because this program suite contains the simulated annealing protocol, which can move groups of atoms more than 3 Å [Brünger, 1988]. This may save time in the refinement and avoid false minima. However, the simulated annealing refinement protocol was unsuccessfully applied to the structures of bovine BLG. The discrepancy between the R_f and the R factor tends to run out of control, with differences, R_f - R, as large as 15 ~ 20 %. To avoid such catastrophes, the positional refinement protocol, following the principle of least-squares was used in the refinement. As mentioned in Chapter 3, with the method of least-squares, it is easy to become trapped in local minima, at which stage manual intervention is required. Programs TURBO-FRODO [Cambillau *et al.*, 1996], O [Jones and Kjeldgaard, 1993] or TOM [Oatley, 1994] allow rebuilding and adjustment of the model to incorporate new structural features (e.g. waters, etc.), and then to escape from local minima. The set of reflections for R_f was chosen randomly for each structure. While a common set of reflections would have been preferable, the poor merge between pairs of data sets (R ≈ 0.25) indicates that these sets of reflections for R_f were relatively independent of each other.

4.8.2 Structure refinement of BLGA at pH 8.2

The initial R factor was 0.39 (10 ~ 3.0 Å) for a model comprising residues 1-61, 66-110, 116-152. In the initial stages of refinement of BLGA at pH

8.2, only the reflections between 6 ~ 2.6 Å were used, as X-PLOR [Brünger, 1988] does not handle well disordered solvent in its early version. With a new version of X-PLOR, with improved function in handling the solvent correction, the reflections used in refinement were extended to a lower resolution of 15 Å. Several hundreds of reflections (576) were set aside for monitoring R_f, after the solvent correction was applied to the original reflection file. The simulated annealing refinement protocol was tested but, the discrepancy between R_f and R ran out of control; see Figure 482-1 for details.The positional refinement protocol (least-squares method) was used to refine BLGA at pH 8.2. Refinement were monitored closely to keep R_f - R smaller than 8%.

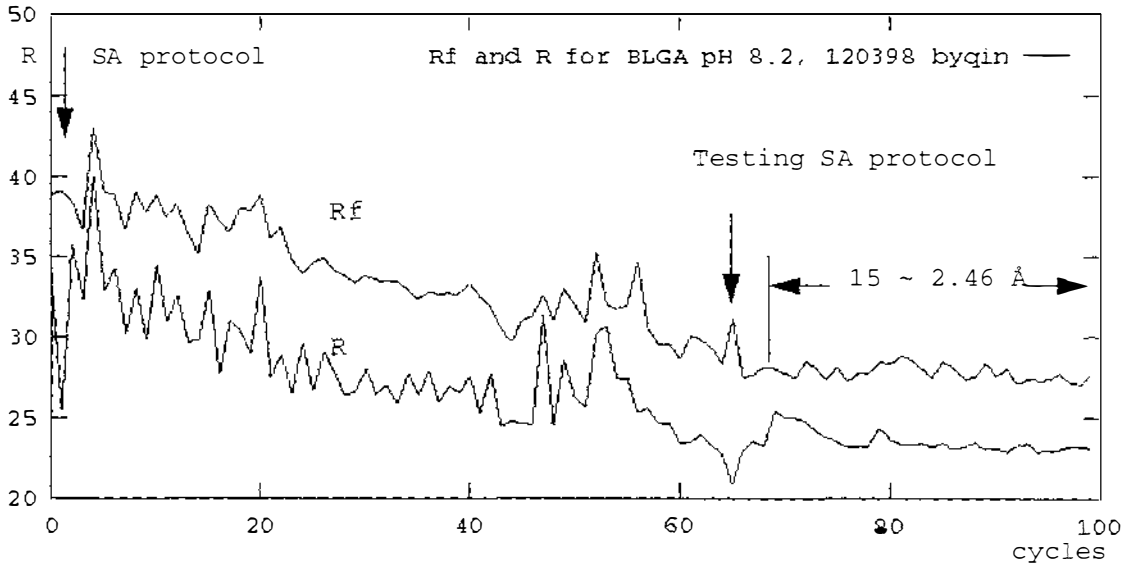


Figure 482-1: Progress of refinement for BLGA at pH 8.2

After positional refinement, electron density maps (2Fo-Fc and Fo-Fc) were calculated and the molecular graphics program TURBO [Cambillau *et al.*, 1996] used to rebuild the partial model. The rebuilding first began from the disulfide bond between Cys66-Cys160. The amino acid residues were included one by one in a very careful manner, but ignored other quality criteria such as geometry and preferred φ/ψ values. After including several residues around Cys160, it became clear that the residues after 150 in lattice Z extended in a different direction compared to the search model. Successful rebuilding from 150 - 162 then allowed loop CD to be defined. The loop GH was included later after the conformations of existing residues in this region were changed. With these changes and additions, values for R_f and R improved dramatically.

For the first 69 cycles of refinement, the reflection data file was indexed only to 2.6 Å. Some uninterpretable electron density in the map raised suspicions about the quality of data processing. The re-indexed and reprocessed diffraction data set extended to a slightly higher resolution of 2.46 Å. The uninterpretable electron density disappeared on further refinement and rebuilding. Figure 482-1 shows the progress of R and R_f factors during different refinement and rebuilding cycles.

4.8.3 Structure refinement of BLGA at pH 7.1

Based on the structure for BLGA at pH 8.2, the initial R factor was 0.25 (10 ~ 3.0 Å) with a correlation coefficient of 0.86 for BLGA at pH 7.1. This starting model had all 162 residues. The reflections used in the refinement lay in the range of 15 ~ 2.24 Å. A total of 557 reflections were set aside for monitoring R_f, after the solvent correction was applied to the original reflection file. The positional refinement protocol (least-squares method) was used to refine the structure. The steps of refinement were chosen to keep the discrepancy between R_f and R to smaller than 8%. For the first 48 cycles, the reflection data file was indexed to only 2.30 Å. The reprocessed diffraction data set extended to a slightly higher resolution of 2.24 Å. At first, the refinement included data only to 2.28 Å, Later, after the incorrectly included residue Q63 was corrected to Asn, refinement included data to 2.24 Å. Figure 483-1 shows the progress of refinement.

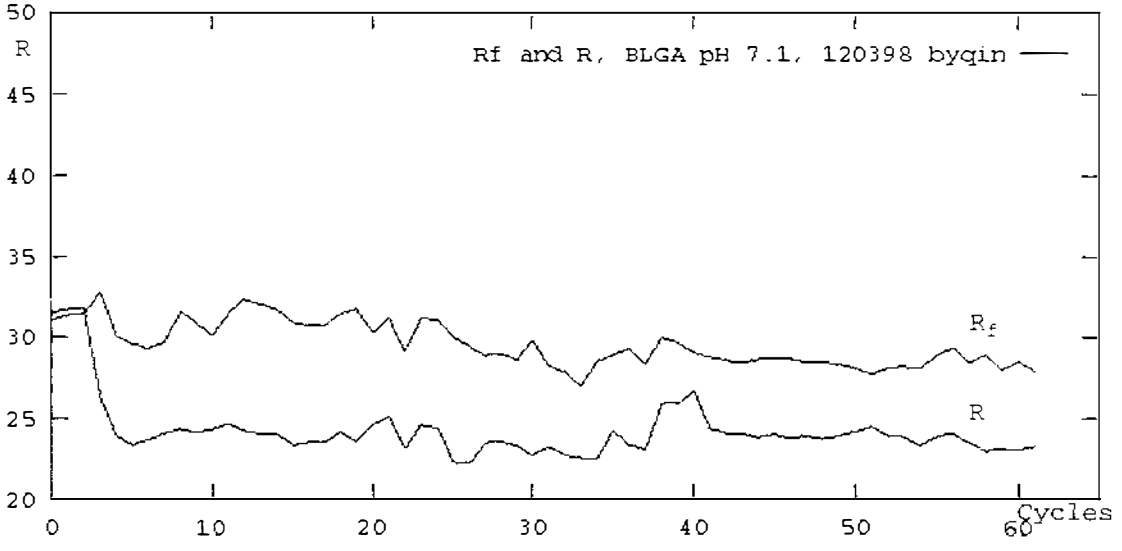


Figure 483-1: Progress of the refinement for BLGA at pH 7.1

4.8.4 Structure refinement of BLGA at pH 6.2

Based on the structure for BLGA at pH 8.2, the initial R factor was 0.21 (10 ~ 3.0 Å) with a correlation coefficient of 0.89 for BLGA at pH 6.2. This original model contains 162 residues. The reflections used in the refinement lay in the range of 15 ~ 2.56 Å. About 5% of data (545 reflections) were set aside for monitoring R_f, after the solvent correction was applied to the original reflection file. The positional refinement protocol (least-squares method) was used. The initial electron density map suggested that the loop EF adopted a different conformation from the original model. Rebuilding of loop EF and further cycles of refinement and rebuilding have provided good values for R_f and R. Initially, only a few waters were included in this model. After correcting residue Gln63 to Asn63 and carefully rebuilding loop EF, refinement proceeded much better with many water molecules being found. In the last stages, R_f dropped to 0.24, and R dropped to 0.19. Taking the resolution into account (2.56 Å), we are surprised but pleased with this result. Inspection of the structures of BLGA at pH 7.1 and 8.2 with this structure as a reference, did not identify regions of the former two structures that needed alteration. Figure 483-1 shows the progress of R and R_f during refinement and rebuilding cycles.

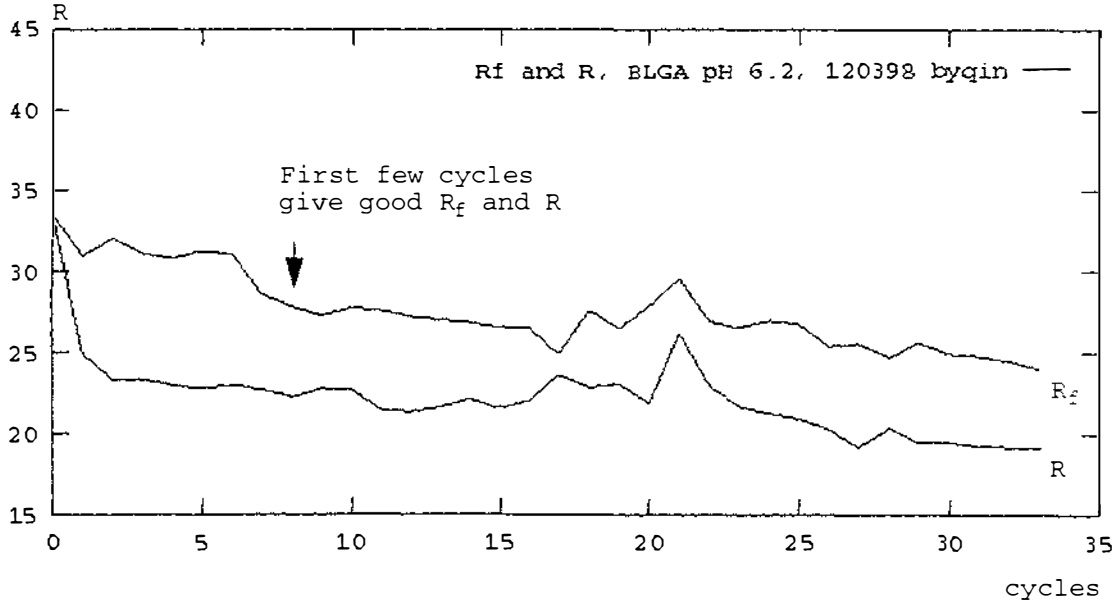


Figure 484-1: Progress of refinement for BLGA at pH 6.2

4.8.5 Structure refinement of BLGB at pH 7.1

Based on the structure for BLGA at pH 7.1, the initial R factor was 0.26 (10 ~ 3.0 Å) with a correlation coefficient of 0.84 for BLGB at pH 7.1. The residues Asp64 and Val118 were changed to Gly64 and Ala118, respectively. The reflections used in the refinement lay in the range of 15 ~ 2.4 Å, at the beginning stage. Initially loop EF was modelled with two conformations. Even though values for R_f and R were not bad (see Figure 485-1), the electron density map in this region was not good. This fact caused reindexing and reprocessing of the data set. The reprocessed diffraction data file extended to 2.22 Å resolution. A total of 576 reflections were set aside for monitoring R_f, after the solvent correction was applied to the original reflection file. The positional refinement protocol (least-squares method) was used to refine the structure. The initial electron density map suggested that the loop EF was different in conformation from the structure of BLGA at pH 7.1. In the last stage, R_f dropped to 0.28, and R dropped to 0.23. Figure 485-1 shows the progress of refinement.

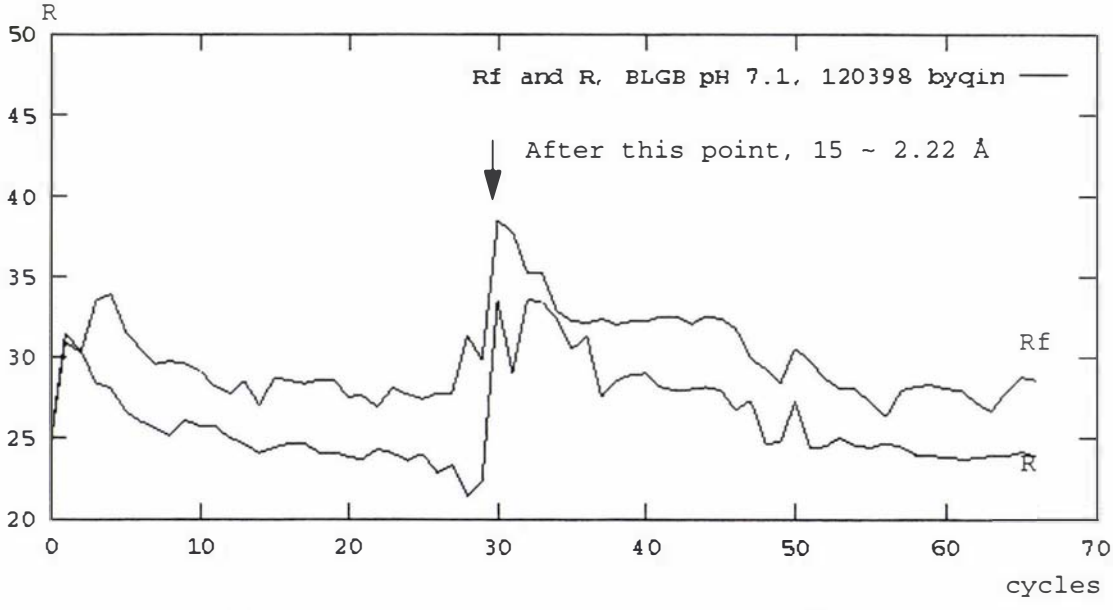


Figure 485-1: Progress of refinement for BLGB at pH 7.1

4.8.6 Structure refinement of BLGA-BrC12

Based on the structure for BLGA at pH 7.1, the initial R factor was 0.256 (10 ~ 3.0 Å) with a correlation coefficient of 0.855 for BLGA-BrC12 at pH 7.3. This original model contained 162 residues and no water molecules. The reflections used in the refinement were in the resolution range of 15 to 2.23 Å. A total of 511 reflections were set aside for monitoring R_f, after the solvent correction was applied to the original reflection file. The positional refinement protocol (least-squares method) was used to refine the structure. The steps of refinement were monitored to keep the discrepancy between R_f and R smaller than 8%. At the starting point, R = 30.93% and R_f = 30.41%. (Calculation of R factor involves more reflections than that of R_f, as a result, R may be slightly larger than R_f initially). After including the ligand 12-Br$(CH_2)_{11}$COOH in the structure, the R and R_f factors have shifted to 32.77% and 34.16%, respectively. The normal refinement and rebuilding cycles were followed through the whole refinement process. The final values for R_f and R factors were 27.93% and 23.23%, respectively. Figure 486-1 shows the progress of refinement.

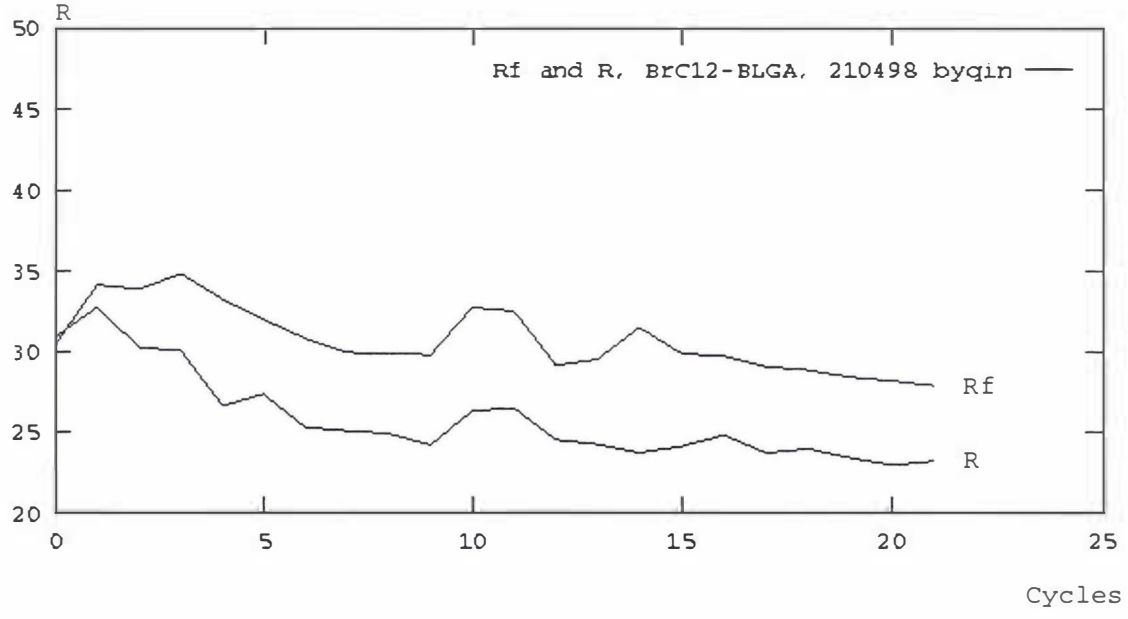


Figure 486-1:Progress of the refinement for BLGA-BrC12

Chapter 5 Results

5.1 Validation of models

5.1.1 Quality of reflection data sets

Although X-ray diffraction data sets were collected from a number of crystals, only five of them were used for refinement and model building. The diffraction pattern from the crystal of bovine BLGA at pH 8.2 was the first data set collected, followed by data from the crystal of bovine BLGB at pH 7.1. Initially, it was supposed that the BLG crystals in lattice Z were very fragile, and data were collected very rapidly, as the crystals were expected to decompose. This generated two data sets (BLGA at pH 8.2 and BLGB at pH 7.1) with limited completeness and redundancy, especially in the highest resolution shell (Table 511-1). The average redundancy was less than three and the completeness for last shell was lower than 90%. Together with a poorly chosen error model in SCALEPACK, these two data sets were initially processed to only 2.6 and 2.4 Å. However, the lattice Z crystals are not as fragile as first supposed. The exposure time for each frame in subsequent data sets was increased from ~15 mins to 24 mins, the total data collection time taking 50% longer. The oscillation range of each frame was also increased from 1.2° to 1.5°. Improved capability in use of the data processing software also helped to obtain better quality data sets for BLGA at pH 6.2, BLGA at pH 7.1 and BLGA-BrC12 complex at pH 7.3; See Table 511-1 for details.

Table 511-1 : Data collection statistics for bovine BLG crystals

File	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Variant	A	A	A	B	A
pH	6.2	7.1	8.2	7.1	7.3
Unique	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
last shell Å	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Redundancy	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
R-merge	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Completeness	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
<I/sigmaI>	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Chi**2	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

The values in parentheses pertain to the last shell

5.1.2 Quality of models

A small uncontaminated reflection set was set for each data set to monitor the progress of R_f during the refinement process. In general, the discrepancy of R_f and R was kept to less than ~ 8% to avoid overfitting the model to the data. The final models for BLGA at different pH conditions all have the R_f < 28% and R < 24%. Especially, for the structure of BLGA at pH 6.2, the value for R_f is only 24.03% while that for R is 19.19%. The R_f for BLGB is somewhat higher at 28.62%, but well below the 30% value above which serious inadequacies in the model or data are indicated. In all cases, only one residue is located in a disallowed region of the Ramachandran plot. This residue, Tyr99, sits at the centre of a γ-turn, which is stabilized by a hydrogen-bond network with neighbouring residues, as described in detail in section 5.7. The RMS for bond distances (the root-mean square deviation from the target values, which come from the very precisely determined structures of small molecules) in all structures is lower than 0.009 Å. There are only a limited number of outliers for each geometric criterion. Therefore, confidence may be placed on the reliability of the structures presented in this thesis. For more details, see Table 512-1.

Table 512-1 : Refinement statistics for structures of BLG in lattice Z

Content	BLGA	BLGA	BLGA	BLGB	BLGA-BrC12
pH condition	pH 6.2	pH 7.1	pH 8.2	pH 7.1	pH 7.3
Resolution limit (Å)	15.0-2.56	15.0-2.24	15.0-2.46	15.0-2.22	15.0-2.23
R_f (#reflections)	24.03%(545)	27.93%(557)	27.69%(576)	28.62%(576)	27.93%(511)
R	19.19%	23.35%	23.16%	23.93%	23.23%
Ramachandran plot[a]					
core	82.6%	83.9%	81.2%	83.1%	82.6%
allowed	15.4%	14.8%	18.1%	14.9%	16.8%
generously allowed	1.3%	0.7%	0.0%	1.4%	0.0%
disallowed	0.7%	0.7%	0.7%	0.7%	0.7%
Geometrics RMS					
for bonds Å	0.009	0.008	0.005	0.009	0.008
for angles °	1.60	1.52	1.41	1.69	1.62
for dihedrals	25.32	25.51	25.13	25.11	25.49
for impropers	1.20	1.15	1.11	1.21	1.18
Outlier residues					
bonds >0.06 Å	1	0	0	0	2
angles >10 °	0	3	2	1	3
dihedrals >60 °	0	2	2	0	2
impropers >5 °	1	3	1	1	3

[a] PROCHECK [Laskowski, 1993]

5.2 General description of bovine BLGA and BLGB in lattice Z

The model for bovine BLGA in lattice Z at pH 7.1, which has been refined with data extending to a resolution of 2.24 Å, see Table 512-1, will be used as the reference structure in the following structure descriptions. BLG will be disassembled into sub-regions. Focussing on the BLGA pH 7.1 model, bovine BLG will be described in detail in the following categories: the differences in the structures due to pH and variant type, van der Waals interactions, solvent accessibility of residues, hydrogen-bonds (main chain and side chain), ψ and ϕ torsion angles, other torsion angles, surface charges, etc.

5.2.1 Core structure of bovine ß-lactoglobulin

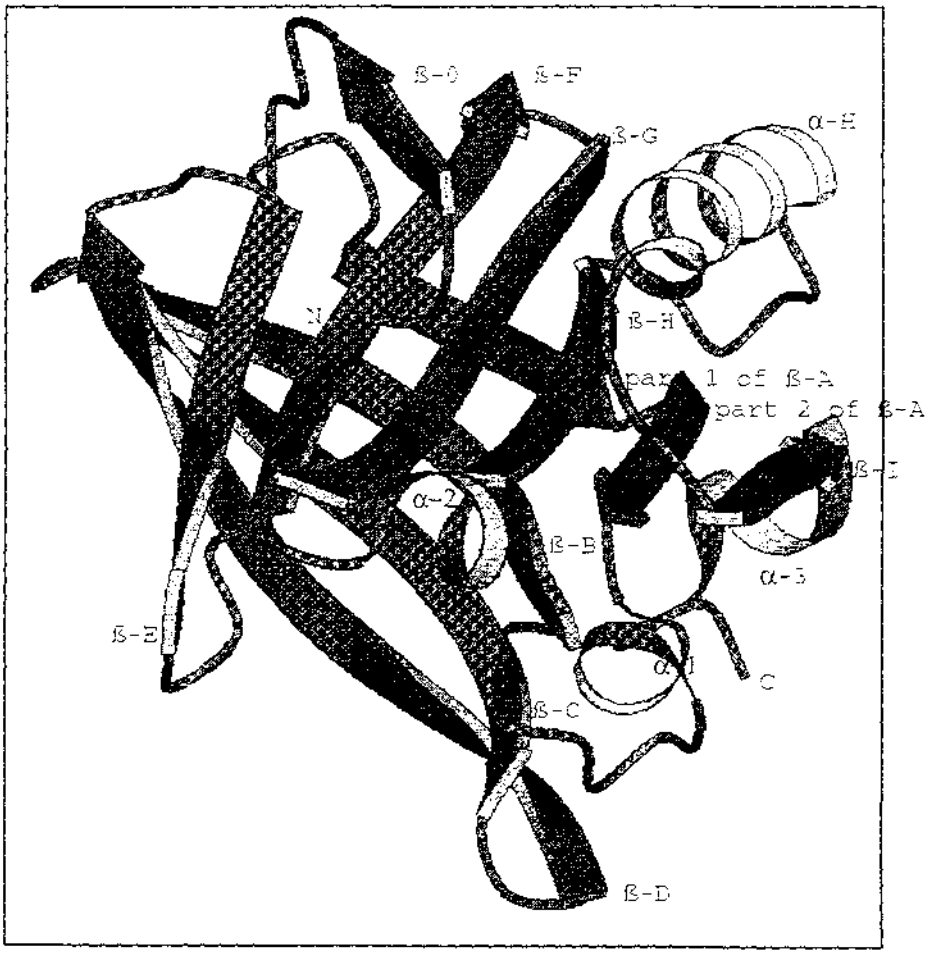


Figure 521-1: Cartoon picture of the bovine BLG calyx, prepared by MOLSCRIPT [Kraulis, 1991].

BLGA is a small globular protein molecule with 162 residues, the majority of which belong to ß-strands. Ten ß-strands (labelled as ß-0, ß-A to ß-

I) and four α-helices (labelled as α-H, α-1 to α-3) can be identified for bovine BLGA in lattice Z at pH 7.1. Eight anti-parallel ß-strands, ß-A to ß-H, fold back and forth to form the calyx (Figure 521-1), a typical feature of the lipocalin super-family. The strands are connected by loop regions, labelled as loops 0A, AB, BC, etc. The strand ß-A involves residues 16-27 and can be viewed as two consecutive ß-strands with a kink between residue 21 and residue 22. Thus, strand ß-A is divided into part 1 (residues 16-21) and part 2 (residues 22-27). Attached to the calyx are two extra ß-strands. One of them, ß-0, is close to the N-terminus and involves residues 4-8; the second, ß-I, is near the C-terminus end of the molecule. This strand serves as the dimer interface for BLG. Through ß-I, an extended ß-sheet, which includes all the ß-strands of the two subunits of the BLG dimer, is created, as illustrated in Figure 521-2. For BLG in the trigonal lattice Z and orthorhombic lattice Y, a crystallographic dyad relates the two monomers.

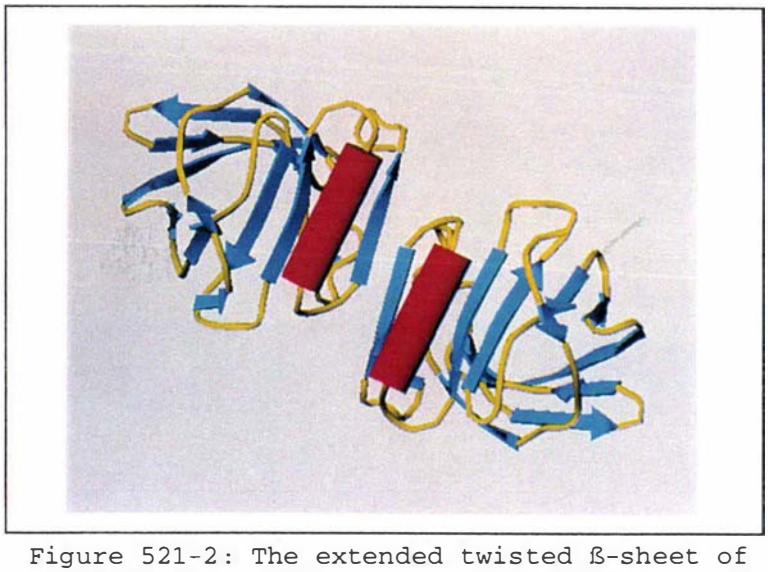


Figure 521-2: The extended twisted ß-sheet of the BLG dimer. Figure prepared with TURBO [Cambillau *et al.*, 1996].

The core structure of BLG, the ten ß-strands and helix α-H, is unaffected by changes in pH in the range 6.2 to 8.2. The structure of BLGA at pH 6.2 *versus* the structure of BLGA at pH 8.2 has an rms displacement of 0.23 Å for 145 Cα atoms. The structure of BLGA at pH 7.1 *versus* the structure of BLGA at pH 8.2 has an rms displacement of 0.21 Å for 145 Cα atoms. The superimposed Cα atom traces of BLGA at different pH are displayed in Figure 521-3. Except for loop EF, changes in other parts of the BLGA structures are virtually unidentifiable.

The point mutations characterizing variants A and B of BLG have limited influence on the calyx. The structures of BLGA at pH 7.1 and BLGB at pH 7.1 are very similar; 155 Cα atoms with a maximum displacement less than 0.7 Å have an rms displacement of 0.233 Å. The superimposed Cα atom traces of BLGA and BLGB at pH 7.1 are displayed in Figure 521-4.

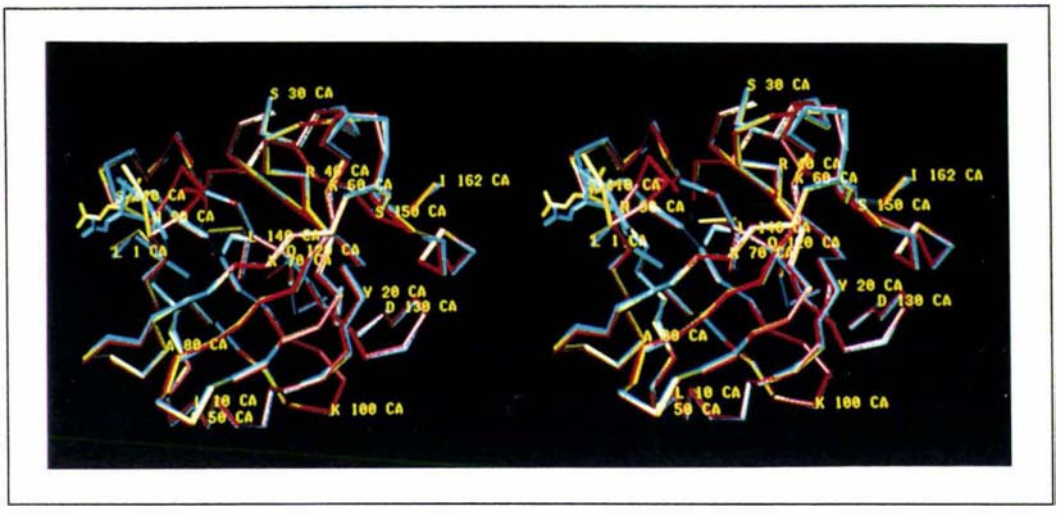


Figure 521-3: Superposition of Cα traces of BLGA in lattice Z at pH 6.2 (red), pH 7.1 (yellow) and pH 8.2 (blue) in lattice Z. The side chain of Glu89 and disulfide bonds 66-160 and 106-119 are shown. Figure prepared with TURBO [Cambillau *et al.*, 1996]

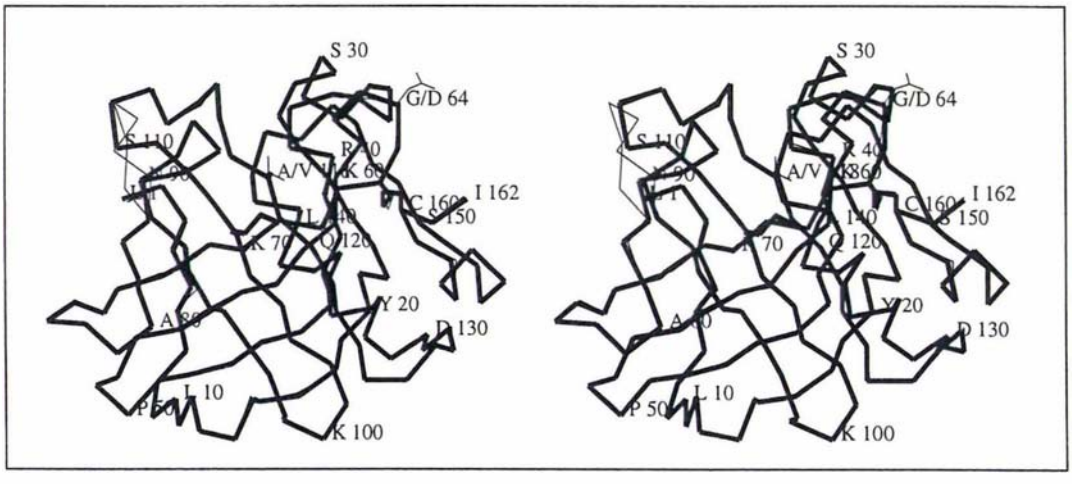


Figure 521-4: Superposition of the Cα traces of BLGA (thin line), BLGB (thick line) at pH 7.1 in lattice Z. Figure prepared with TURBO [Cambillau *et al.*, 1996]. The side chains at the point mutation sites (D64G and V118A) and the disulfide bridges are shown.

5.2.2 Topology of the main chain hydrogen bonds

The details of the main chain hydrogen-bond patterns of bovine BLGA and BLGB are shown in the topological diagrams of Figures 522-1 to 522-4. Secondary structure was defined according to Kabsch and Sanders, 1983, except that residues at the ends of secondary structure elements that were involved in main chain-main chain hydrogen bonds were added to the

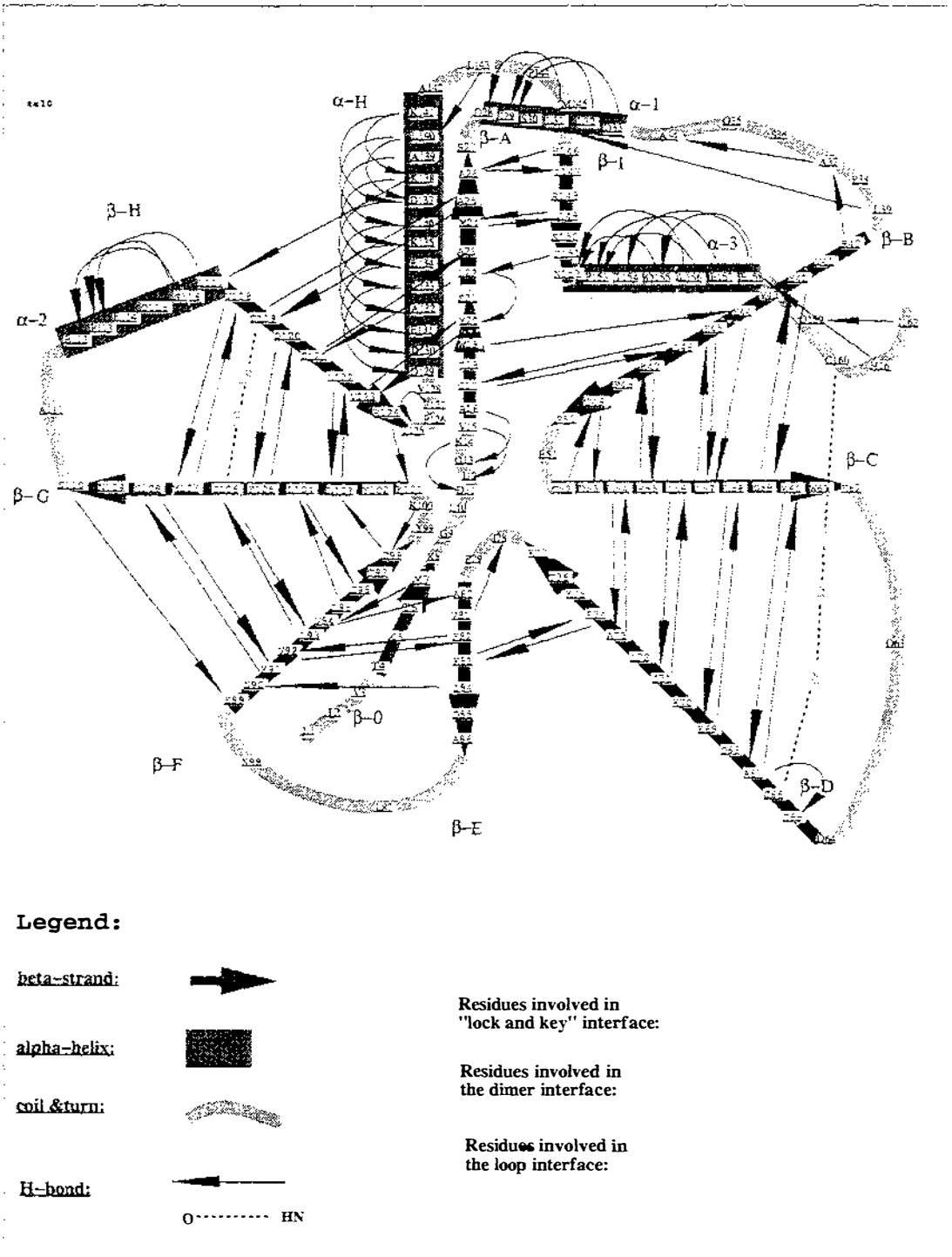


Figure 522-1: Topology of main-chain hydrogen bonds of BLGA at pH 7.1.

element and residues with very high B values and no main chain-main chain hydrogen bonds were assigned to loops. Hydrogen bonds satisfied the criteria X-(H)...Y 2.9 ± 0.4 Å, the angle X-H...Y > 120 ° (X or Y are N or O). The length of the ß-F and ß-G strands varies according to experimental conditions, and the short α-2 helix is not conserved in all structures. The ß-I strand is attached to the calyx by part 2 of the ß-A strand. The ß-0 strand is attached to the calyx by only one hydrogen

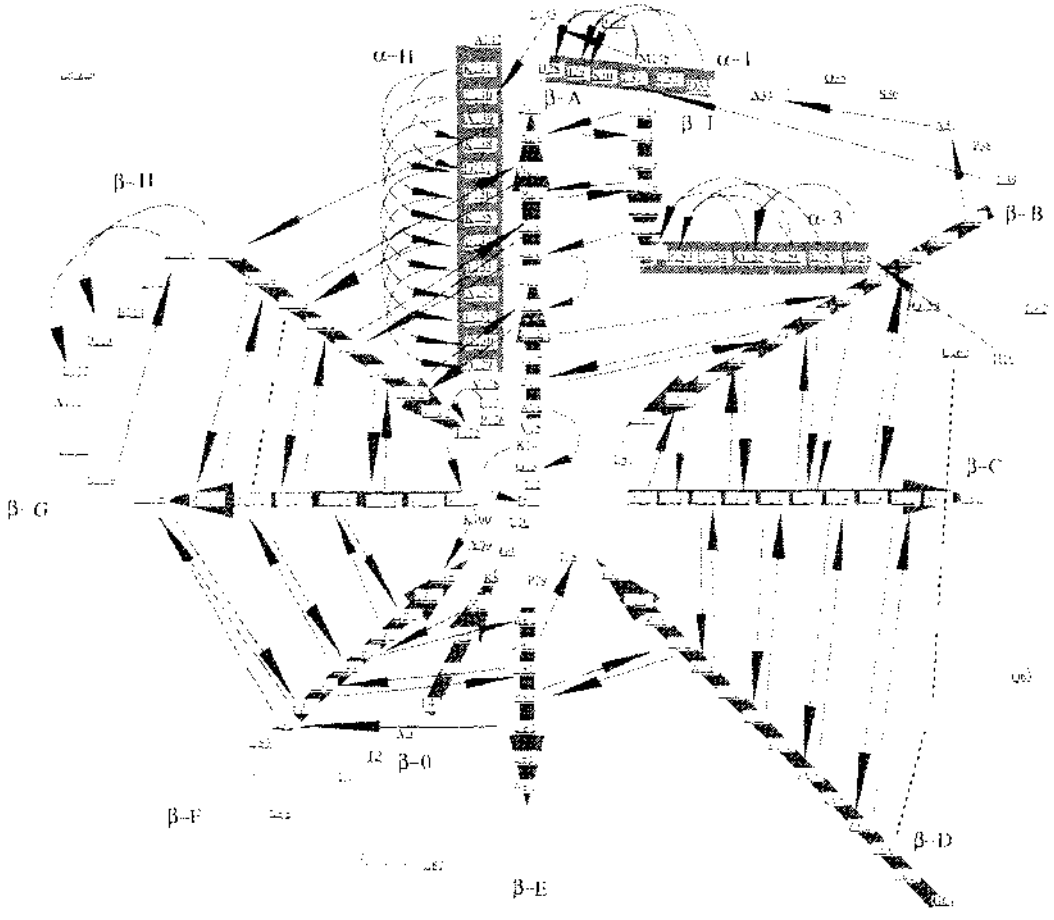


Figure 522-2: Topology of main-chain hydrogen bonds of BLGA at pH 6.2.

bond. The full hydrogen bond potentiality between the ß-D and ß-E strands is not used up from these patterns, only one pair of main chain hydrogen bonds can be identified between the ß-D (14 residues) and ß-E (7 residues) strands. Between the ß-A and ß-B strands is along loop. From this point of view, the eight ß-strands which form the calyx can be grouped into two sets: ß-B, ß-C and ß-D strands as one set, ß-E, ß-F, ß-G, ß-H and ß-A (part 2) strands as another. The shortest ß-strand of the

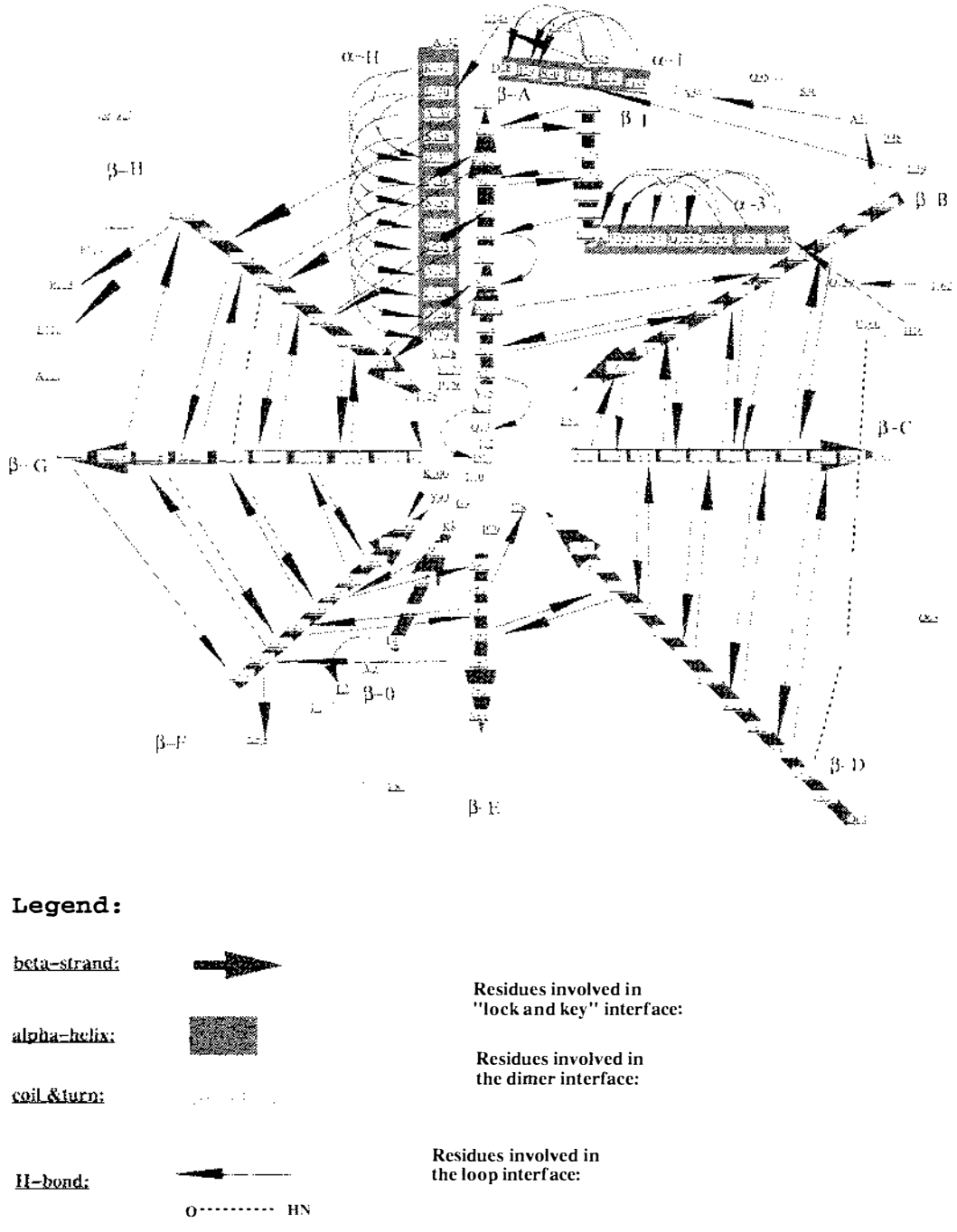


Figure 522-3: Topology of main-chain hydrogen bonds of BLGA at pH 8.2.

first group is longer than the longest ß-strand of the second group (11 *vs.* 10 residues).

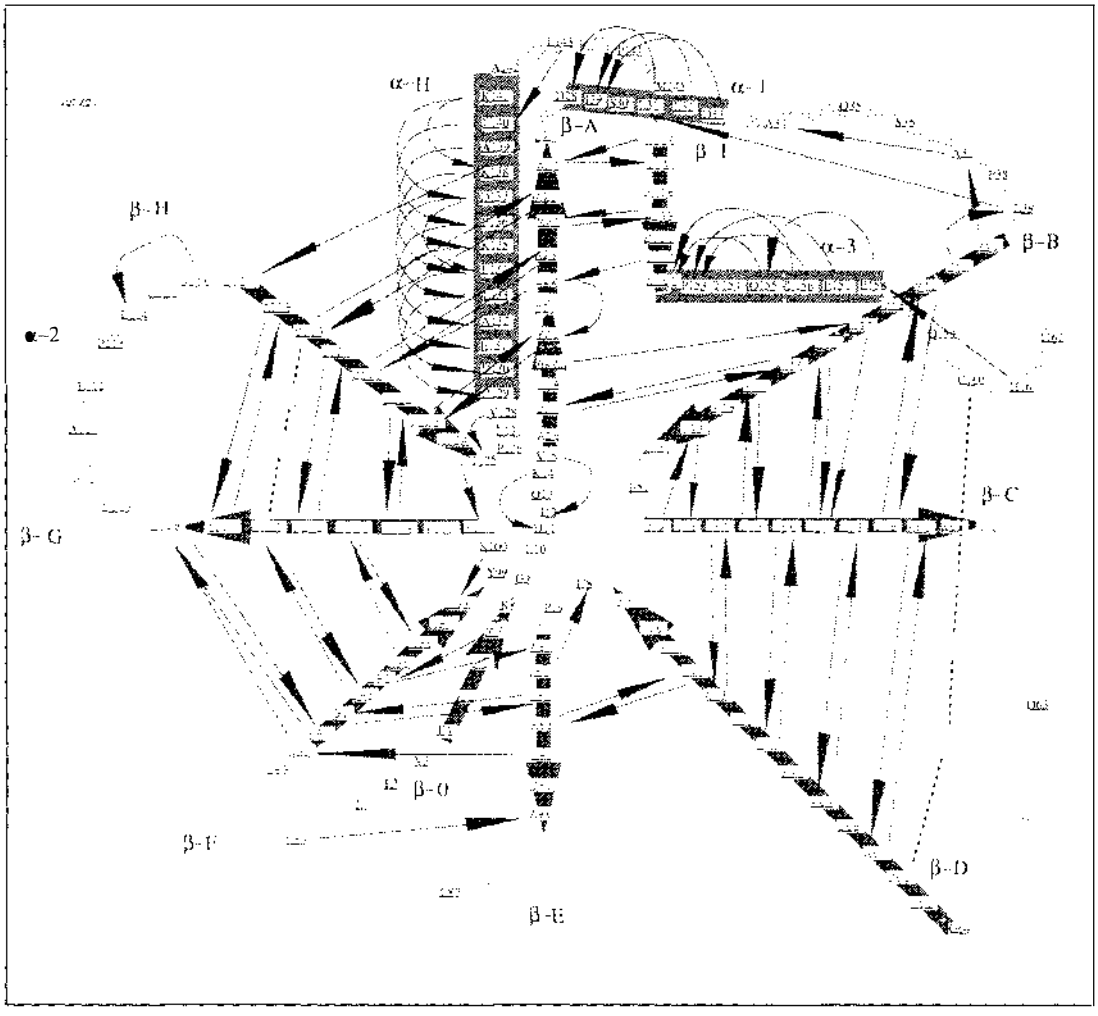


Figure 522-4: Topology of main-chain hydrogen bonds of BLGB at pH 7.1; legend is the same as Figure 522-3.

The topology diagram (Figure 522-4) of BLGB at pH 7.1 reveals the similarity with that of BLGA at pH 6.2 in the region of ß-F and ß-G. The hydrogen bond S110N...O89E in BLGA at pH 7.1 and pH 8.2 is absent in both BLGB at pH 7.1 and BLGA at pH 6.2.

5.2.3 Sub-regions of ß-lactoglobulin

BLGA and BLGB are dimers in solution at physiological conditions in bovine milk. Each monomer can be disassembled into several structurally distinct regions: top region (green), bottom region (white), handle region (yellow), ß-sheet I (red) and ß-sheet II (blue), as illustrated in Figure 523-1, where water molecules are displayed in orange. The BLG calyx is partially surrounded by the handle region, which consists of

the C and N termini, helix α-H and strand ß-I, as described in more detail in section 5.5. At the top of calyx are several loops, labelled as AB, CD, EF and GH. They look like the doors to the entrance to the calyx, as detailed in section 5.3. The bottom of the calyx is formed by loop 0A, residues 124, 125, loop FG and part 1 of ß-A; see section 5.4 for more details. The calyx can be viewed as actually being comprised of two ß-sheets, labelled as ß-sheet I (section 5.6) and ß-sheet II (section 5.7). The strands ß-B, ß-C and ß-D produce a distinct curved ß-sheet I. The strands ß-E, ß-F, ß-G, ß-H and ß-A (part 2) produce a relatively flat ß-sheet, ß-sheet II, which is oriented perpendicular to ß-sheet I; see Figure 521-1, 521-3, 521-4 and 523-1.

Figure 523-1: Disassembly of BLGA molecule into sub-regions: top region (green), bottom region (white), handle (yellow), ß-sheet I (red), ß-sheet II (blue) and water (orange). Figure prepared with TURBO [Cambillau *et al.*, 1996]. Different regions deliberately isolated by omission of peptide bond of neighbor residues.

5.2.4 Average B factors of bovine BLG

The B factors for BLGA/B are summarized in Table 524-1. The average B factor of BLGA at pH 6.2 is 41 $Å^2$, relatively lower than that of BLGA at pH 7.1, 8.2 and BLGB at pH 7.1. This may reflect the fact that BLGA at pH 6.2 has lower values for R_f and R. Within an individual molecule, the average B factors of different regions reflect the mobility of that region. Waters, with average B factors from 68 ~ 75 $Å^2$, are relatively more mobile than the amino acid residues, with average B factors from 39

~ 51 Å². The atoms in the side chains have higher average B factors than those in the polypeptide backbone. The top region of bovine BLG is the most mobile part of the BLG molecule, especially the loop EF. Opposite to the top region, the bottom region is relatively rigid. For each structure, the average B factor of ß-sheet I (41 ~ 57 Å²) is higher than that of the corresponding ß-sheet II (32 ~ 42 Å²).

Table 524-1: Average B factors in different regions of BLG (Å²)

Resolution limits (Å)	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Region/Molecule	BLGA pH [illegible]	BLGA pH 7.1	BLGA pH 8.2	BLGB pH 7.1	[illegible]-BLGA
All (including waters)	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Residues (protein atoms)	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Backbone	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Side chains	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Top region	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Loop EF ([illegible])	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Loop GH ([illegible])	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Loop AB ([illegible])	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Loop CD ([illegible])	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Bottom region	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Handle region	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
N-terminus ([illegible])	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
α-H ([illegible])	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
ß-I ([illegible])	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
C-terminus ([illegible])	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
ß-Sheet I ([illegible])	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
ß-Sheet II	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Waters	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

5.2.5 Characteristics of the surface of BLGA and BLGB

From the top of the BLG calyx, the surfaces of bovine BLGA at pH 7.1 and 8.2 are similar; see Figures 525-1 and 525-2. In these two cases, the loop EF adopts an open conformation, so that the inside of the BLG calyx is visible and accessible. The depth of the cavity (in the middle of molecule) is beyond the center of the BLG molecule (the cross point of

three axes defines the molecular center); see Figure 525-2. In Figure 525-1, a retinol molecule is modelled in the calyx. The little pillar which protrudes to the left and down is Lys77. The loop EF of BLGA in lattice Z at pH 6.2 and that of BLGB in lattice Z at pH 7.1 are different from the cases above, and adopt a closed conformation, so that the inside of the BLG calyx is not observable from the top of the calyx; see Figures 525-3 and 525-4. The orientation of molecules in Figures 525-3 and 525-4 is the same as in Figures 525-1 and 525-2.

Whereas the top region of BLG is strongly negatively charged, the bottom region is more hydrophobic and surrounded by a set of mostly positive residues (Figure 525-5). At the center, however, is Lys8. This residue is involved in an important intermolecular contact in lattice Z, whereby the side chain of Lys8, functioning as a key, inserts into a lock, the small cavity near Glu44, illustrated in Figure 525-6. This "lock and key" interface will be discussed in detail in section 5.9.2. The area buried in this interface is summarised in Table 525-1. The positively charged regions are colored blue in Figure 525-1 to 525-6; the negatively charged regions are colored in red.

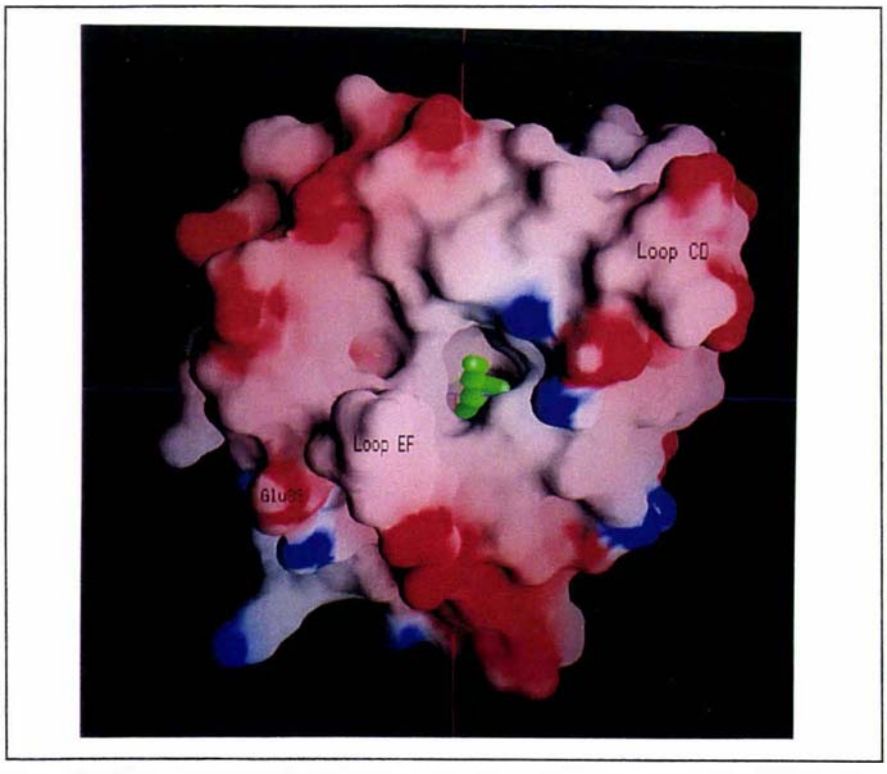


Figure 525-1: Surface charge on the top region of BLGA in lattice Z at pH 7.1. A retinol molecule has been fit into the ligand binding site Figure prepared with GRASP [Nicholls *et al.*, 1991].

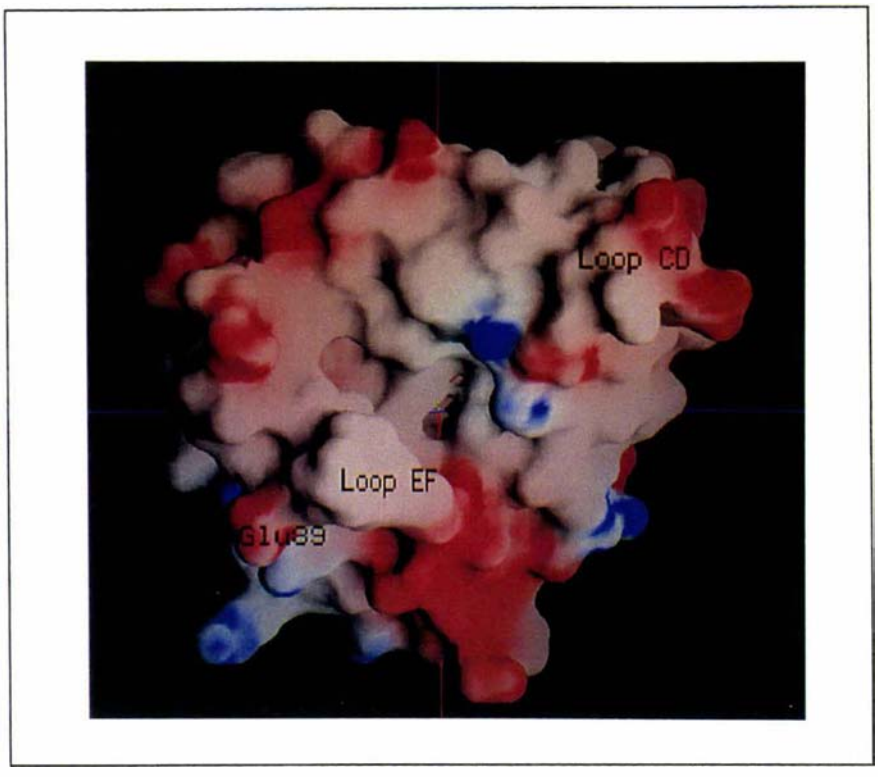


Figure 525-2: Surface charge on the top region of BLGA in lattice Z at pH 8.2. Figure prepared with GRASP [Nicholls *et al.*, 1991].

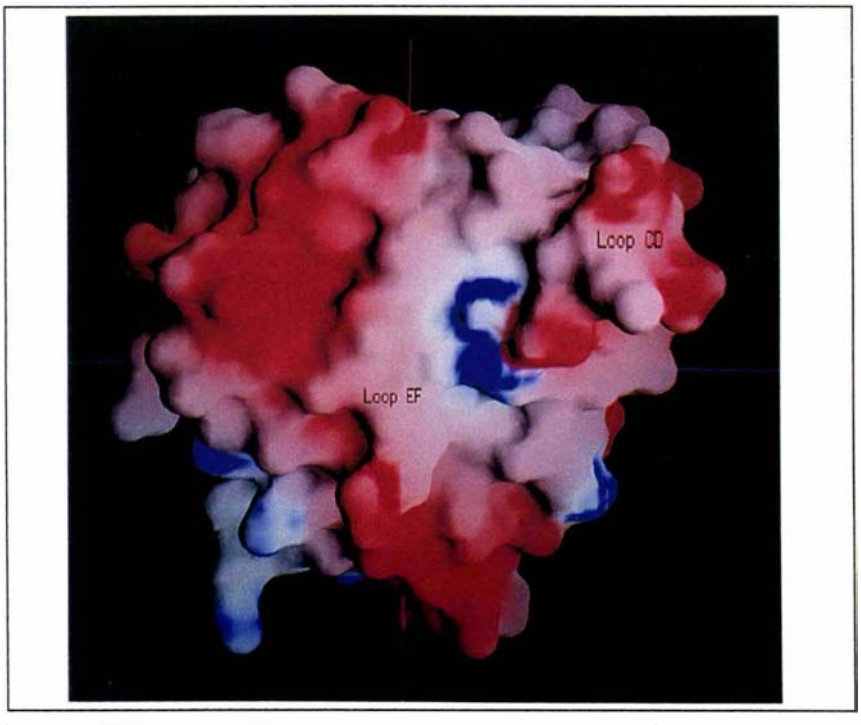


Figure 525-3: Surface charge on the top region of BLGA in lattice Z at pH 6.2. Figure prepared with GRASP [Nicholls *et al.*, 1991].

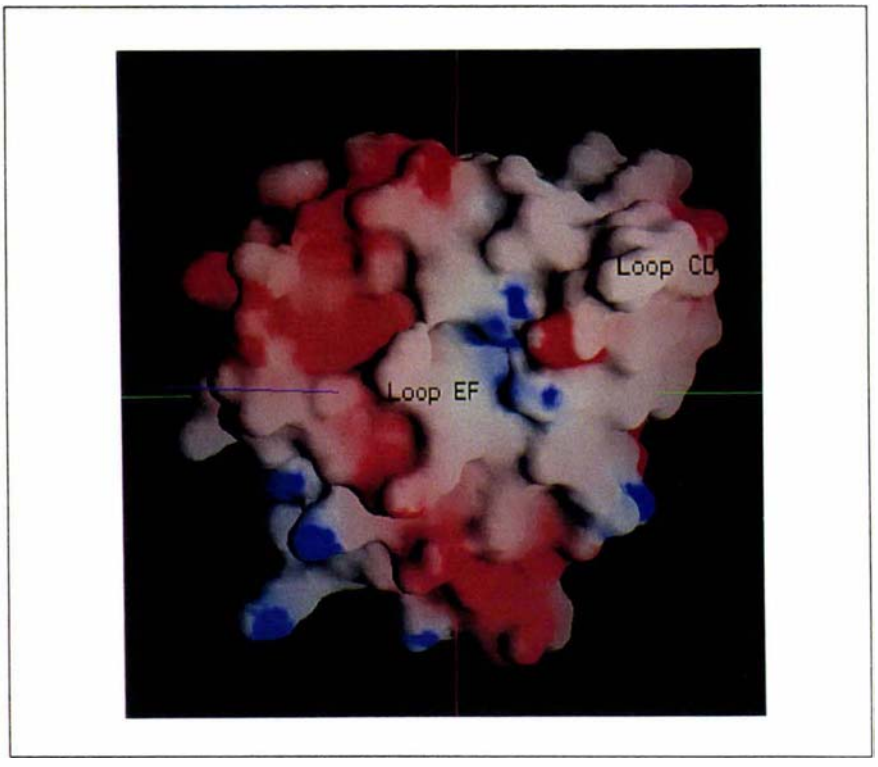


Figure 525-4: Surface charge on the top region of BLGB in lattice Z at pH 7.1. Compare with Figure 525-3 to see the loss of charge in loop CD arising from the D64G point mutation. Figure prepared with GRASP [Nicholls *et al.*, 1991].

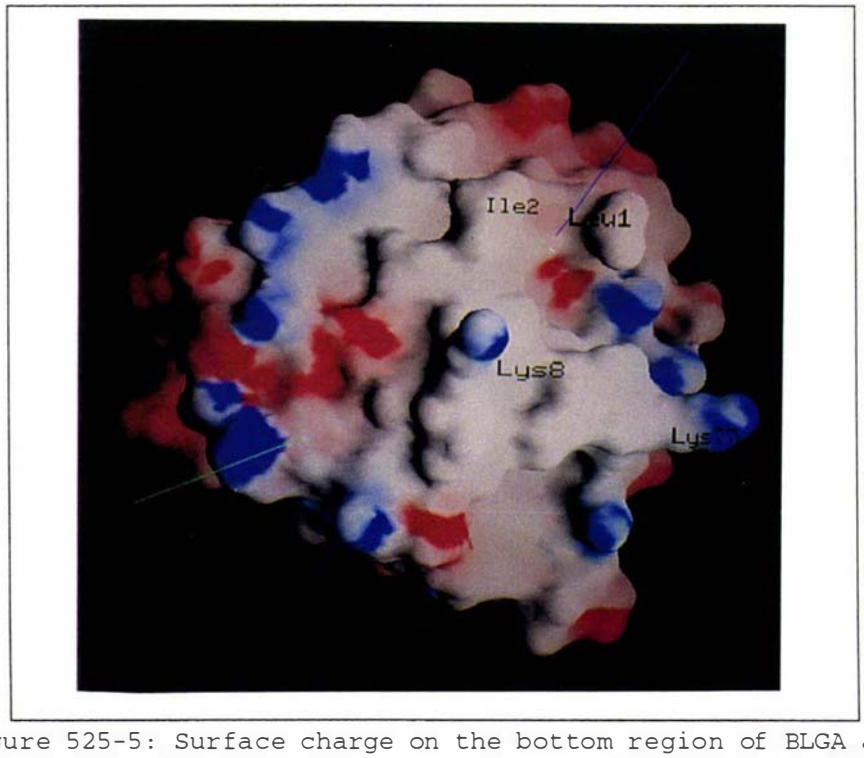


Figure 525-5: Surface charge on the bottom region of BLGA at pH 7.1. Lys8 is involved in an important intermolecular contact, in which it functions as a key. Figure prepared with GRASP [Nicholls *et al.*, 1991].

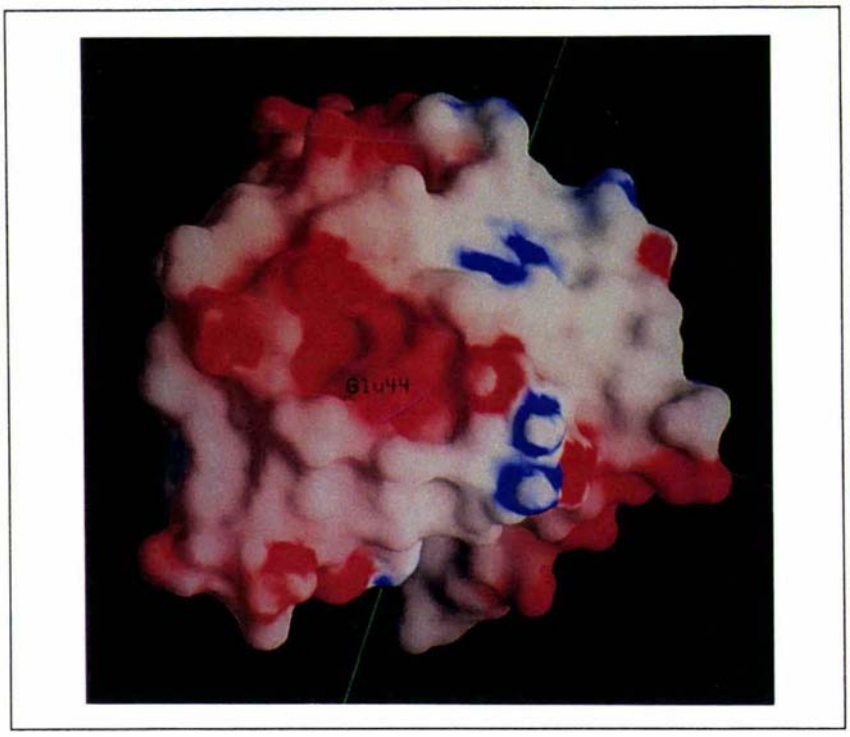


Figure 525-6: Surface charge in the vicinity of the cavity by Glu44 of BLGA at pH 7.1. This cavity accepts Lys8 to form an important intermolecular interaction. Figure prepared with GRASP [Nicholls *et al.*, 1991].

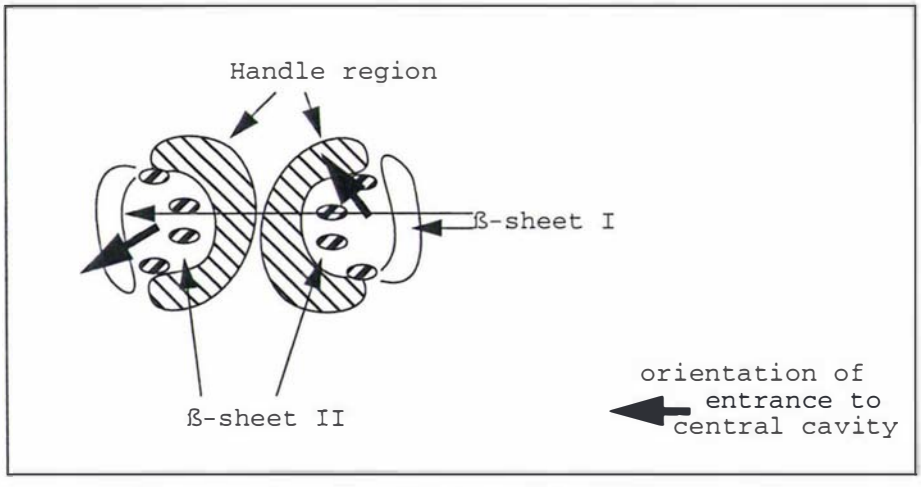


Figure 525-7: Relative position of monomers

The two monomers of BLGA dimer are symmetrically related in lattice Z: a crystallographic two-fold axis exists linking the ß-I strands to give an extended ß-sheet spanning two monomers. The calyx handle regions approach together to form the dimer interface. This interface is sandwiched between two BLG calices; see Figure 525-7. A 60° angle is found

between the orientations of the two calyx entrances of the dimer. The surface in the vicinity of the dimer interface is displayed in Figure 554-4. The surface of this region is intertwined with positively and negatively charged residues. At the center of the hydrophobic region (ß-I) is Arg148, a positively charged residue.

The accessible surface areas of the bovine BLGA and BLGB have been calculated with programs AREAIMOL and RESAREA of the CCP4 suite [Collaborative, 1994]; see Table 525-1. From pH 6.2 to 8.2, the surface area of the monomer and dimer of BLGA increase ~430 Å^2 and ~800 Å^2, respectively. The formation of BLGA and BLGB dimers as well as the "lock and key" interface buries part of the monomer surface. For BLGA the surface area buried in the dimer interface increases slightly with pH, while that associated with the "lock and key" decreases slightly. As a percentage of the total monomer surface area, the area buried in the dimer interface stays constant at 5.74%, 5.75% and 5.84%. BLGB at pH 7.1 resembles BLGA at pH 6.2 in terms of monomer and dimer surface areas and area buried in the dimer interface (5.69%). These percentages lie at the low end of the range typically found for strongly associated dimers [Jones and Thorton, 1995]. For the "lock and key" interface the percentage area buried per pair of molecules decreases from 6.64% at pH 6.2 to 6.18% at pH 7.1 to 5.69% at pH 8.2. BLGB at pH 7.1 with a percentage at 7.09% resembles, again, BLGA at pH 6.2. Note that each molecule has both a lock and a key which in the crystal structure effectively doubles the percentage of surface areas buried.

Table 525-1 : Accessible surface areas for bovine BLGA and BLGB

pH	BLGA 6.2	BLGA 7.1	BLGA 8.2	BLGB 7.1
Monomer Å^2	8235	8458	8666	8269
Dimer Å^2	15524	15944	16319	15597
Lock and key' Å^2	15377	15870	16346	15365
Buried for dimer Å	946	972	1013	941
Buried for lock and key Å	1093	1046	986	1173

' This interface leads to chains of molecules in the solid state; the surface area is calculated from a pair of molecules

5.3 Top region of ß-lactoglobulin

5.3.1 General description of the top region

The top region in BLGA at pH 7.1 consists of residues 85-90 (loop EF), 109-116 (loop GH), 28-39 (loop AB), and 61-67 (loop CD). The residue composition of top region may vary slightly according to the conformation of the loops EF and GH. These four loops are not structurally related to each other, but are distinctive compared with other regions of bovine BLG, especially, with respect to their high average thermal motion. In addition, these loops may be functionally cooperative with each other, as the door or lid to the BLG calyx. The top region has one disulfide bond link with the handle region (Cys66-Cys160). Helix α-1, which belongs to the top region, has extensive hydrophobic interactions with the residues 142-144, which are immediately after the α-H helix. In general, the top region is isolated from other regions and is rather mobile.

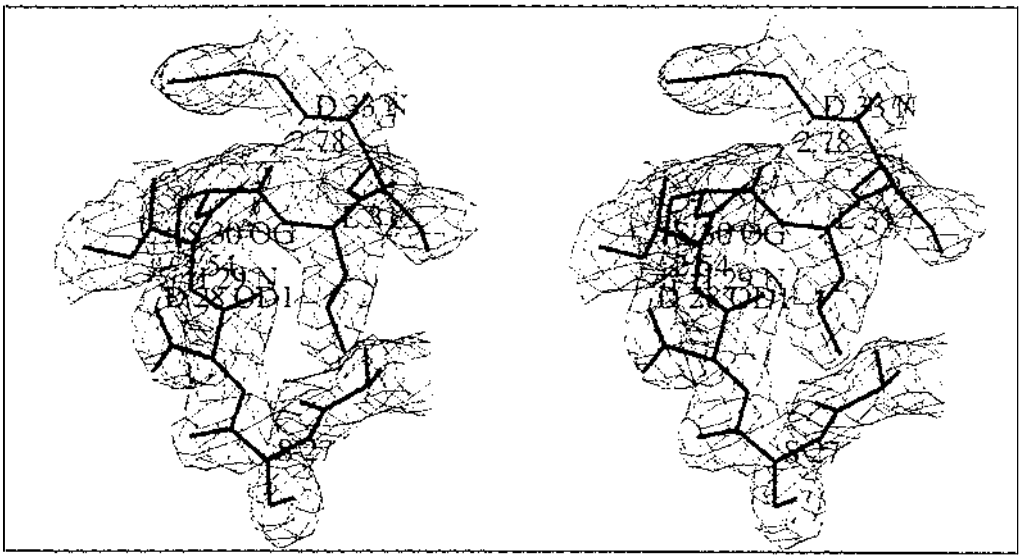


Figure 532-1: 2Fo-Fc electron density map, showing two hydrogen bonds of D28OD1...OG30S (2.54 Å) and D33N...O29I (2.78 Å) in the helix α-1 within loop AB of BLGA at pH 7.1.

5.3.2 Loop AB

Loop AB is made up of a short helix, α-1 (28 ~ 33), with rest of the residues forming a random coil. As the pH increases from 6.2 to 8.2, Gln35 and Leu39 become more exposed; see Table 532-1. This region provides hydrophobic interactions for the BLG dimer interface, especially α-1, and this region is isolated from the rest of the top region, loops CD, EF and GH. These three loops form the "loop interface"

in the lattice Z, which provides a cushion between the chains of BLG molecules; for more details see section 5.9.3. In the formation of the dimer interface, Ile29 and Asp33 are the most affected residues. Their accessible surfaces decrease from 85 Å^2 to 17 Å^2 and from 95 Å^2 to 41 Å^2, respectively. This is exemplified by the case of BLGA at pH 7.1; see Table 532-1.

The residues after α-1 in loop AB are mostly hydrophobic residues with the exception of Gln35 and Ser36. The side chain of Gln35 appears to form a hydrogen bond with O atom of Cys160, although the side chain of Gln35 proved difficult to fit into the density map at all stages of rebuilding and refinement. The OD1 atom of Asp28 and OG atom of Ser30 form a strong hydrogen bond, with an OD1...OG separation of 2.54 Å; see Figure 532-1. This hydrogen bond is helpful in stabilizing the 3$_{10}$ helix of α-1.

Table 532-1: Accessible surface area of residues in loop AB (Å^2)

variant residue		A pH 8.2	A pH 6.2	B pH 7.1	A pH 7.1	A(in dimer) pH 7.1
ASP	28	64	57	53	63	57
ILE	29	85	90	70	85	17
SER	30	69	60	60	65	67
LEU	31	22	18	21	25	24
LEU	32	0	0	2	0	0
ASP	33	99	85	89	95	41
ALA	34	47	48	56	46	37
GLN	35	115	48	41	36	37
SER	36	72	77	78	68	73
ALA	37	5	9	6	8	6
PRO	38	87	59	52	68	66
LEU	39	50	1	8	43	45
Total change between the monomer alone and the monomer in dimer exemplified by BLGA at pH 7.1						132

5.3.3 Loop CD

The average B factor of loop CD, residues 61-67, is very high, about 70% higher than the overall average B factor of all residues; see Table 533-1. Such high B factors suggest that loop CD is very mobile. Residues of loop CD appear at relatively low density (0.8σ) in 2Fo-Fc maps, although an omit map clearly shows their existence (see Figure 533-1). This highly

mobile region is anchored to the C-terminus of BLG by the disulfide bond Cys66-Cys160. All residues of loop CD are highly exposed, irrespective of pH in the range of 6.2 to 8.2 and variant A or B, as detailed in Table 533-1. Cys66 in BLGB is less exposed compared to that in BLGA. Loop CD contains a rather exposed hydrophobic residue, Trp61 (solvent accessible surface area, 54 $Å^2$) compared to Trp19 (only 3 $Å^2$).

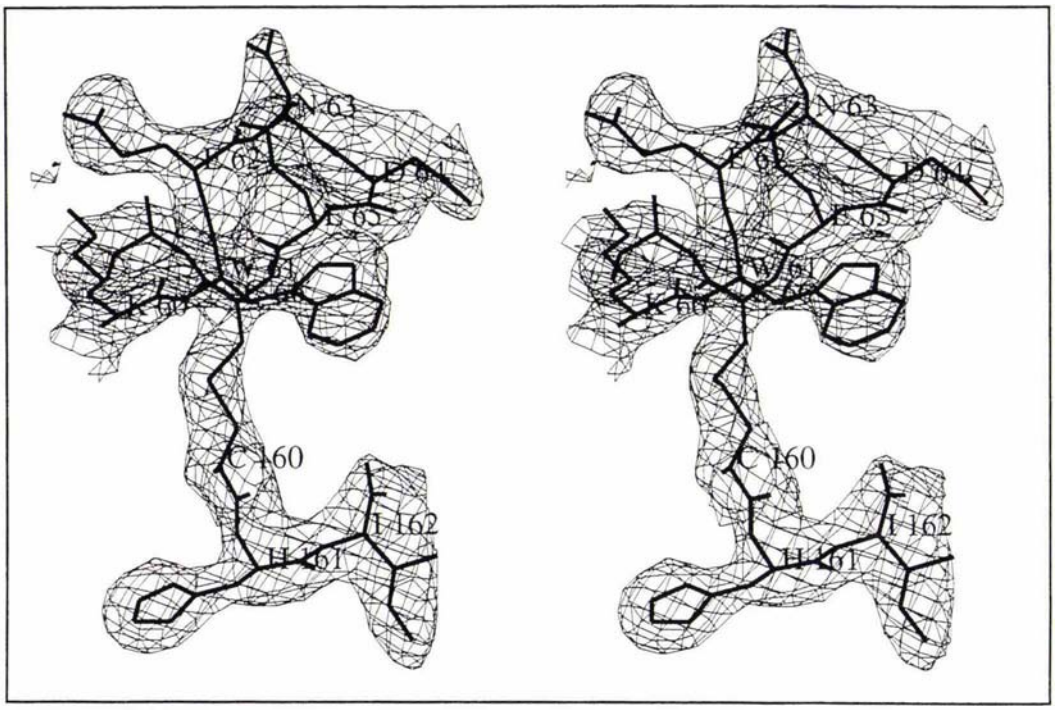


Figure 533-1: 2Fo-Fc electron density map, showing loop CD of BLGA in lattice Z at pH 7.1

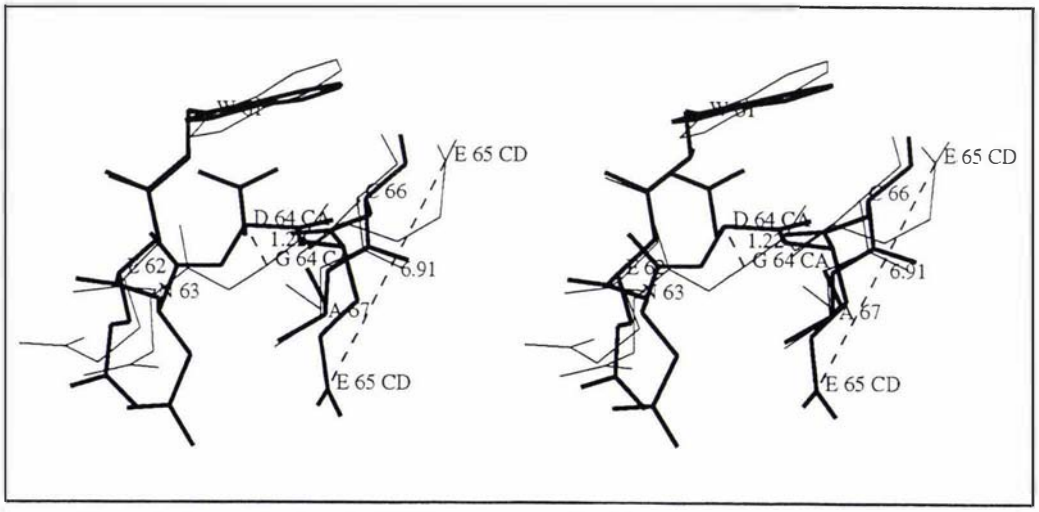


Figure 533-2: Orientation of Glu65 in BLGA (thick lines) a and BLGB (thin lines) at pH 7.1 in lattice Z.

Asp64 in BLGA has been replaced by Gly64 in BLGB. The phi/psi torsion angles of Gly64 and Asp64 are possibly different: -101°/-23°, -143°/37°, respectively. This difference influences the orientation of the side chains of Glu 62, Asn63 and Glu65, especially the side chain of Glu65, which points explicitly in a different direction in BLGB compared to BLGA at pH 7.1 (see Figure 533-2). The orientation of Glu65 in BLGA at pH 6.2 bears similarity to BLGA at pH 7.1, but not to BLGB at pH 7.1. Such differences may be introduced by the point mutation D64G.

Table 533-1: Accessible surface area of residues in loop CD (Å^2)

variant residue		A pH 8.2	A pH 6.2	B pH 7.1	A pH 7.1	A(in dimer) pH 7.1
TRP	61	59	60	56	59	63
GLU	62	91	87	110	98	96
ASN	63	139	133	143	116	109
ASP/GLY	64	123	129	56	133	137
GLU	65	140	132	105	103	97
CYS	66	38	34	19	36	34
ALA	67	34	22	39	16	16

Total change between the monomer alone and the monomer in dimer exemplified by BLGA at pH 7.1 9

The small increases in surface area for the dimer compared to the monomer species is due to coodinate round off in TURBO-FRODO and difference in the grid used by RESAREA of CCP4. Surface area numbers are considered precice to ± 3 Å^2.

5.3.4 Loop EF

While the BLGA calyx experiences only minor changes in lattice Z from pH 6.2 to 8.2, a major conformational change occurs for the loop EF, residues 85-90, as shown in Figure 521-3. At low pH, loop EF flips towards the calyx and serves as a door/lid to block access into the calyx. At high pH, loop EF flips away from the calyx, and the calyx becomes accessible for substrates. In addition, this conformational change leads to Glu89 changing from a buried situation at low pH to exposed at high pH; see Figure 534-1a,b. The solvent-accessible area of Glu89 is 4 Å^2 at pH 6.2, 116 Å^2 at pH 7.1 and 136 Å^2 at pH 8.2, as detailed in Table 534-1. Indeed, the 4 Å^2 accessible surface area of Glu89 at pH 6.2 is contributed by its carbonyl oxygen atom, and the COOH group of Glu89 is totally buried. The side chain of Glu89 at low pH is mostly inserted into

the interior of the BLG molecule. Within 5 Å of the CD atom of Glu89 at pH 6.2, the following residues can be identified: Leu39, Leu87, Asn88, Met107, Glu108, Asn109, Ser116 and Val118. On the other hand, only Ser110 and two water molecules lie within 5 Å of the CD of Glu89 at pH 7.1. At pH 6.2, the protonated carboxylate atom OE1 of the deeply inserted Glu89 forms a hydrogen bond with the carbonyl oxygen atom of Ser116 (2.89 Å); see Figure 534-2. The conformation of BLGA at pH 8.2 is similar to that of BLGA at pH 7.1; see Figure 534-1c.

Loop EF of BLGB at pH 7.1 in lattice Z adopts a conformation different from that for loop EF of BLGA at the same pH, as shown in Figure 634-1. The Cα trace of BLGB at pH 7.1 bears similarity to that for BLGA at pH 6.2, as detailed in Figure 534-3. The φ/ψ values of individual residues are shown in Table 534-2. The accessible surface areas of residues are shown in Table 534-1.

Table 534-1: Accessible surface area of residues in loop EF ($Å^2$)

variant / residue		A / pH 8.2	A / pH 6.2	B / pH 7.1	A / pH 7.1	A(in dimer) / pH 7.1
ASP	85	72	152	141	131	13[illegible]
ALA	86	56	54	52	62	60
LEU	87	155	52	15	165	165
ASN	88	61	92	121	56	52
GLU	89	136	4	7	116	114
ASN	90	28	53	46	16	17
Total change between the monomer alone and the monomer in dimer exemplified by BLGA at pH 7.1						8

Table 534-2 : Comparison of φ/ψ torsion angle of residues in loop EF

Residues		BLGA_pH8.2	BLGB_pH7.1	BLG_in_X	BLGA_pH7.1	BLGA_pH6.2	BLGA-Br[illegible]_pH7.1
[illegible]	φ/ψ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	φ/ψ	[illegible] 125	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	φ/ψ	[illegible]	-134 [illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	φ/ψ	-130 [illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	φ/ψ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	φ/ψ	[illegible]	[illegible] 19	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	φ/ψ	[illegible]	-118 [illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	φ/ψ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
91	φ/ψ	[illegible] 139	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	φ/ψ	-124 [illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

BLG_in_X: BLG in lattice X, Brownlow et al.,1997

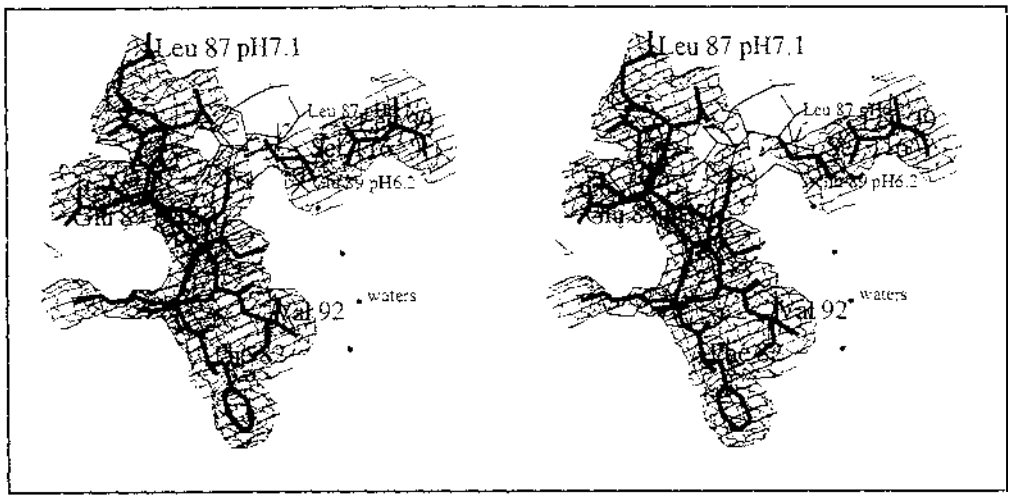


Figure 534-1a: 2Fo-Fc electron density map (pH 7.1), showing conformation of loop EF of BLGA at pH 7.1 (thick line), and at pH 6.2 (thin line).

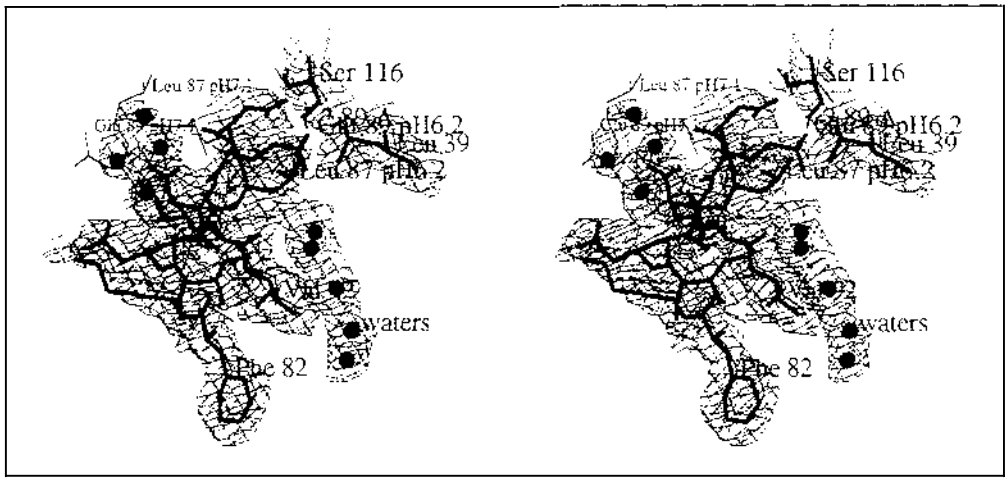


Figure 534-1b: 2Fo-Fc electron density map (pH 6.2), showing conformation of loop EF of BLGA at pH 6.2 (thick line), and at pH 7.1 (thin line).

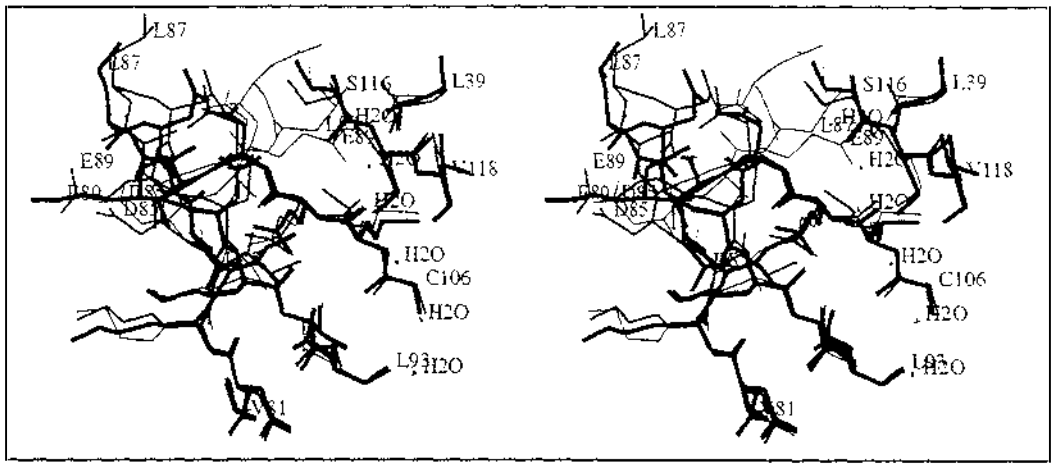


Figure 534-1c: Superposition of loop EF of BLGA at pH 6.2 (thin lines), pH 7.1 (medium lines) and pH 8.2 (thick lines). The water molecules inside the calyx BLGA at pH 6.2 are labelled as H2O.

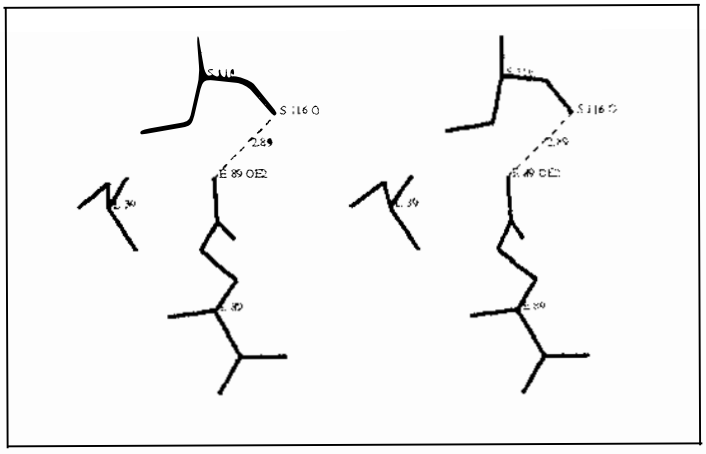


Figure 534-2: Hydrogen bond between Glu89 and Ser116 in BLGA at pH 6.2

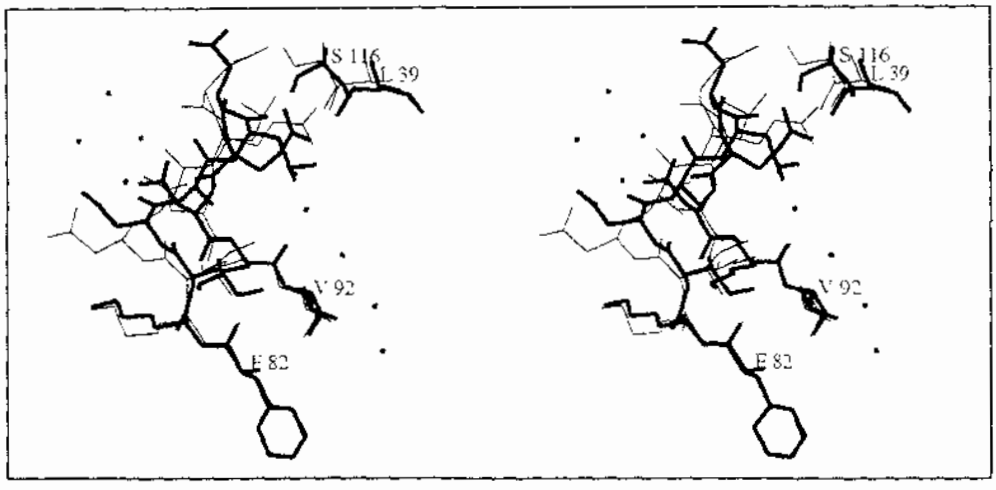


Figure 534-3: Superposition of loop EF of BLGA at pH 6.2 (thick lines) and BLGB at pH 7.1 (thin lines). Water molecules for BLGA at pH 6.2 are shown as dots.

5.3.5 Loop GH

Similar to loop AB, loop GH, as shown in Figure 535-1, contains a short 3_{10} helix, α-2, which runs from residues 110-117. Residue 117 serves both as the end of α-2 and the beginning of ß-H. This 3_{10} helix is not conserved with pH, as at pH 6.2 and pH 8.2 Ser116 falls outside the helical region of the Ramachandran plot. BLGB at pH 7.1 in this instance is conformationally identical to BLGA at pH 7.1. The detailed φ/ψ values of residues in loop GH are given in Table 535-1. Ser116 plays a major role to define this helical hydrogen bond pattern. The residues in loop GH are exposed, except for Ser116 in BLGA at pH 6.2 and BLGB at pH 7.1; see Table 535-2.

Table 535-1: φ/ψ torsion angles of residues in loop GH

pH (φ/ψ)	Asn109	Ser110	Ala111	Glu112	Pro113	Glu114	Gln115	Ser116	Leu117
8.2 A	-91 108	-64 -88	-78 -41	-118 85	-68 -49	-63/-39	-88 1	-103 [illegible]	-104 117
7.1 A	-81 [illegible]	-88 -37	-74 -48	-114 84	-80 -74	-83 -28	-94 4	-107 -41	-43 108
6.2 A	-178 101	-83 -24	-81 -87	-104 89	-66 -67	-82 -27	-81 -28	-148/71	-101 102
7.1 B	-117 [illegible]	-87 -18	-80/-43	-130 [illegible]	-57 -87	-61 -85	-101 11	-174 -81	-46 103

Table 535-2: Accessible surface area of residues in loop GH (Å²)

variant residure		A pH 8.2	A pH 6.2	B pH 7.1	A pH 7.1	A(in dimer) pH 7.1
ASN	109	27	15	34	22	21
SER	110	48	72	71	49	44
ALA	111	84	99	86	74	76
GLU	112	91	72	99	75	77
PRO	113	46	19	37	46	47
GLU	114	124	87	124	126	125
GLN	115	108	102	120	107	114
SER	116	27	3	8	22	21
LEU	117	0	0	0	0	0
Total change between the monomer alone and the monomer in dimer exemplified by BLGA at pH 7.1						-4

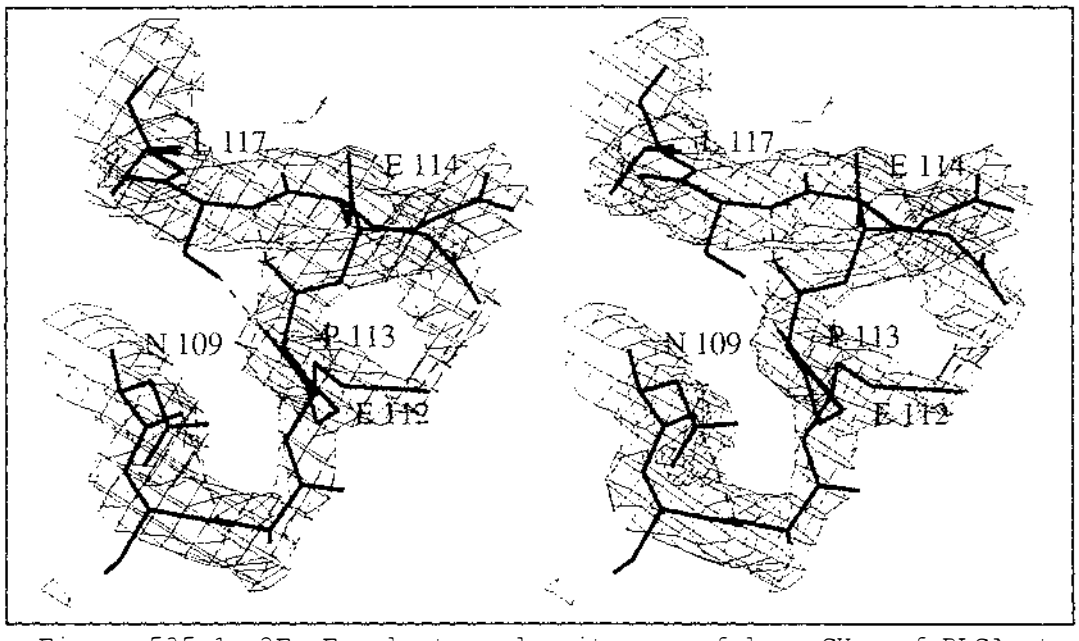


Figure 535-1: 2Fo-Fc electron density map of loop GH of BLGA at pH 7.1. Figure prepared with TURBO [Cambillau *et al.*, 1996].

5.4 Bottom region of ß-lactoglobulin

The bottom region of BLGA/B has extensive main-chain hydrogen bond links with the ß-sheets I and II, as shown in the topological diagrams of the BLG main chain in Figures 522-1 to 522-4. ß-sheets I and II have several turns at the bottom end of the bovine BLG ß-barrel, which are the short loops BC, DE and FG. These loops are far apart from each other, and result in an open bottom to the calyx. To convert the ß-barrel of BLGA/B into a calyx, the bottom portion acts like a plug. As a result, extensive van der Waals interactions exist between these loops and the bottom portion of bovine BLG.

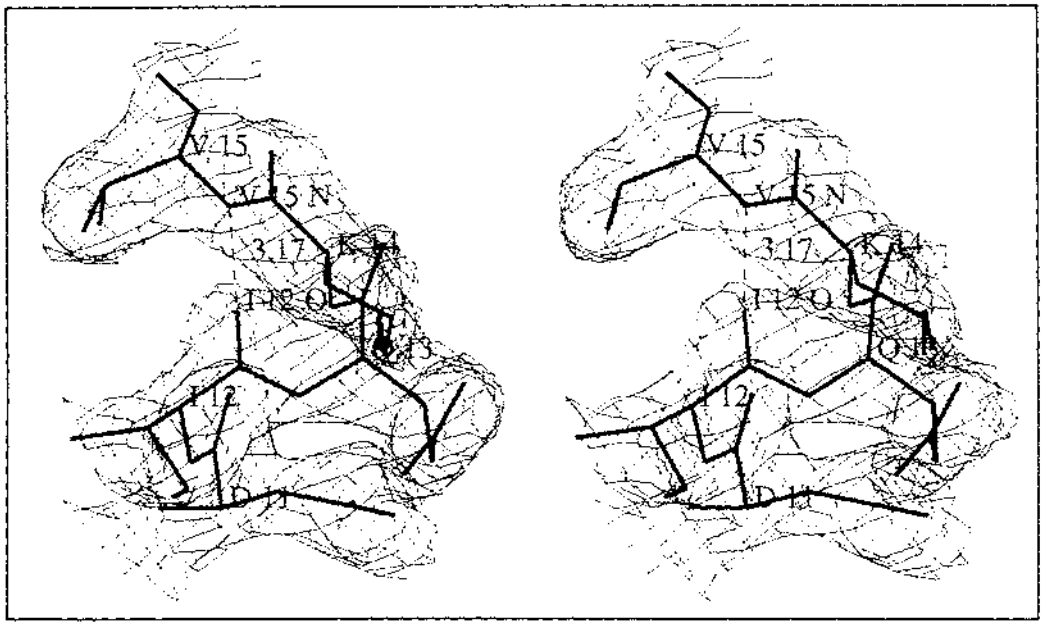


Figure 54-1: 2Fo-Fc electron density map, showing conformation of loop 0A (residues 9-15).

The bottom region includes a short 3_{10} helix, which runs through residues 9 ~ 15 and is conserved in all structures. Immediately after this helix is the beginning of part 1 of strand ß-A. For BLG, residues Arg124 and Thr125 should be included as part of the bottom region, as well Thr97, Asp98 and Tyr99. The bottom 3_{10} helix is somewhat amphipathic, the hydrophilic residues generally pointing to the outside of the BLG molecule and exposed to the solvent (Table 54-1), while the hydrophobic residues generally point to the inside of the calyx; see Figure 54-1. There are no significant changes in surface exposure with pH or variant for this region (Table 54-1). Three hydrophobic residues (Leu10, Ile12, Val15) may be associated with the ß-ionone moiety of retinol, as discussed in section 5.10.1.

As noted previously, the ß-A strand can be viewed as two consecutive ß-strands with a large kink between residue 21 and residue 22. Part 1 of the ß-A strand is mostly comprised of hydrophilic residues, except for Ala16 and Trp19. Part 1 of ß-A serves to link one of the interfaces of ß-sheets I and II between ß-B and ß-H, to form the ß-barrel. Tyr20 is a well exposed tyrosine which forms part of the lock in the "lock and key" interface of BLG in lattice Z, as detailed in section 5.9.2. A clear cavity can be identified in this region; see Figure 525-6.

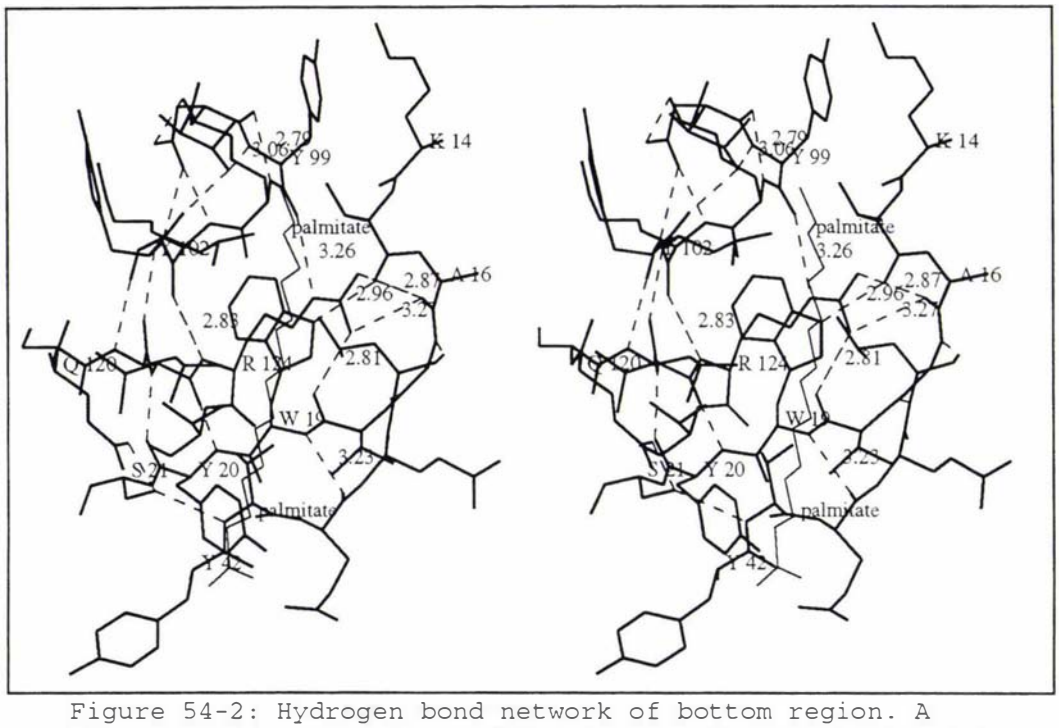


Figure 54-2: Hydrogen bond network of bottom region. A palmitate molecule, modelled inside the calyx, indicates the direction of the cavity.

The bottom region includes several absolutely conserved residues of ß-lactoglobulin: Leu10, Gly17, Trp19, Thr97, Asp98, Tyr99 and Arg124 (Table 234-2). These residues and their neighbours form a complicated hydrogen bond network (Table 54-2). The side chains of Trp19 and Arg124 are parallel, as detailed in Figure 54-2. The hydrophobic Trp19 is conserved throughout all lipocalins [Katakura *et al.*, 1994] and is shielded from the solvent environment. Almost all atoms of Trp19 are buried, except for the carbonyl oxygen atom W19O, which has 3 $Å^2$ exposed area. The guanidinium plane of Arg124 is fully anchored by hydrogen bonds: R124NH1-T18O, R124NH2-A16O and another between R124NE and Y99O, which helps to stabilise the peptide conformation of Tyr99. The parallel alignment of the guanidinium plane and the Trp side chain allow the stacking of π electron clouds, and thus further stabilize the bottom region. The

glycine at position 17 is absolutely essential to allow the side chain of Arg124 to hydrogen bond to A160 and T180 —— any substituent at position 17 would sterically prevent these hydrogen bonds. In addition W19NE1 hydrogen bonds to V150.

Table 54-1: Accessible surface area of residues in the bottom region (Å²)

variant residue		A pH 8.2	A pH 6.2	B pH 7.1	A pH 7.1	A(in dimer) pH 7.1
GLY	9	66	61	71	71	63
LEU	10	26	23	24	24	21
ASP	11	66	69	70	65	66
ILE	12	23	24	20	23	22
GLN	13	109	105	105	105	101
LYS	14	118	115	116	110	112
VAL	15	4	2	2	4	2
ALA	16	42	39	38	42	39
GLY	17	22	26	20	21	20
THR	18	74	79	87	75	74
TRP	19	3	5	7	2	2
TYR	20	51	44	51	40	38
SER	21	0	0	0	0	0
THR	97	4	7	4	6	6
ASP	98	40	44	42	35	35
TYR	99	38	37	37	30	30
ARG	124	49	47	55	53	51
THR	125	69	69	72	71	67
Total change between the monomer alone and the monomer in dimer exemplified by BLGA at pH 7.1						28

Table 54-2: Hydrogen-bond network involving conserved residues of the bottom region (M: Main chain; S: Side chain)

Donor	Length (Å)	Acceptor	Type	Donor	Length (Å)	Acceptor	Type
GLY17N	2.94	LEU46O	MM	ARG124NE	3.26	TYR99O	SM
LEU46N	2.89	GLY17O	MM	ARG124NH2	2.87	ALA16O	SM
TRP19NE1	2.96	VAL15O	SM	ARG124NH1	2.81	THR18O	SM
TRP19N	3.23	GLU44O	MM	ARG124NE	3.26	TYR99O	SM
THR97N	2.89	ASP96OD2	MS	ARG124N	2.83	LYS101O	MM
TYR99N	3.06	THR97OG1	MS	TYR99OH	3.23	GLY9O	MS
THR97OG1	2.75	TYR102●	SM	TYR99N	3.06	THR97OG1	MS
LYS101N	2.87	ASP98OD1	MS	LYS100N	2.79	ASP98O	MM
TYR102N	3.09	ASP98OD1	MS	LYS100NZ	3.13	ASP98OD2	SS
ASP11N	2.84	TYR99OH	MS				

5.5 Handle of the BLG calyx

5.5.1 General description of the handle region of BLG calyx

The N- and C-termini, α-H helix and the ß-0, ß-I strands are grouped into the handle region. The handle region partially covers the ß-barrel of BLGA and BLGB, especially the ß-sheet II. In another point of view, ß-sheet II is sandwiched between ß-sheet I and the handle region of bovine BLG. The van der Waals interactions between the handle region and ß-sheet II are extensive. Without the handle region, the hydrophobic area of ß-sheet II would be seriously exposed. The strand ß-0 has a main chain hydrogen bond with ß-sheet II through ß-F. The ß-I strand has several main chain hydrogen bonds with part 2 of ß-A strand of ß-sheet II.

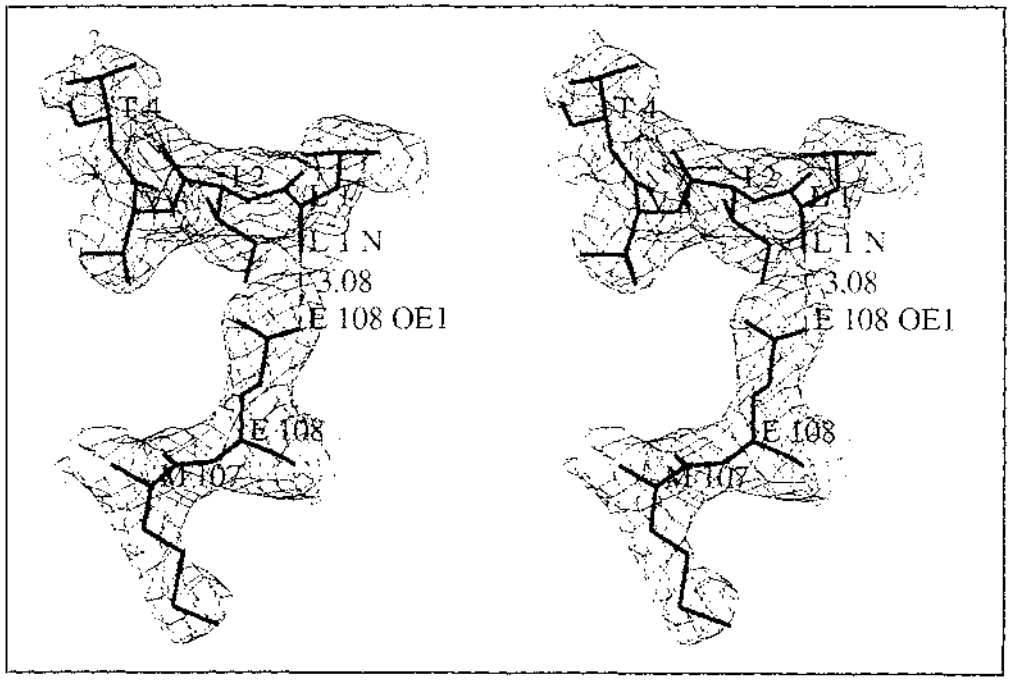


Figure 552-1: 2Fo-Fc electron density map, showing the salt bridge between the N-terminus and Glu89

5.5.2 N-terminus and ß-0

The N-terminus, residues 1-4 which were not observed in the structure of BLG in lattice X [Brownlow *et al.*, 1997], is well defined in lattice Z, as illustrated in Figure 552-1. A very short strand, ß-0, which appears to have been overlooked previously, runs for five residues: Thr4, Gln5, Thr6, Met7 and Lys8. The φ/ψ angles of those residues are listed in Table 552-1; The ψ value of Thr4 falls outside the range for ß-strands, as

shown in Table 552-1. Strand ß-0 is attached to Val94 of ß-F by a hydrogen bond M8N--V94O. Strand ß-0 sits on the top of the ß-lactoglobulin monomer, covering part of strand ß-F between α-H and loop DE. The residues of ß-0 are in direct contact with the solvent. Their accessible surfaces are listed in Table 552-2. The N-terminal α-amino group of Leu1 is right on the top of Glu108 and buries this residue; see Table 552-2 and 57-1.

Table 552-1 : φ/ψ torsion angles of residues in strand ß-0 for BLGA at pH 7.1

	Thr4	Gln5	Thr6	Met7	Lys8
ψ	3	128	143	130	129
φ	-87	-111	-138	-86	-86

Table 552-2: Accessible surface area of residues at the N-terminus and in strand ß-0 (Å²). Glu108 is included for completeness.

variant residue		A pH 8.2	A pH 6.2	B pH 7.1	A pH 7.1	A(in dimer) pH 7.1
LEU	1	144	181	194	186	182
ILE	2	50	42	44	46	43
VAL	3	8	18	13	19	16
THR	4	130	130	113	128	126
GLN	5	68	71	82	67	68
THR	6	60	67	70	73	71
MET	7	41	38	38	37	35
LYS	8	187	203	193	190	187
GLU	108	5	9	3	4	4
Total change between the monomer alone and the monomer in dimer exemplified by BLGA at pH 7.1						18

5.5.3 C-terminus and α-3

In lattice Z of BLGA, the ß-strand I is followed by the helix α-3, which runs from residues 152 to 158. Residue 152 serves both as the end of dimer interface strand ß-I and the beginning of α-3. Because the loop CD is connected to the C-terminus via the disulfide bond Cys160-Cys66, α-3 may serve as a spring to restrain the flexible loop CD (average B factor of residues 61-67 for BLGA at pH 7.1, 80.7 Å²). The C-terminus Ile162 is similar to the N-terminus, Leu1, in terms of the solvent accessibility;

see Table 552-2 and Table 553-1. The side chains of His161, Phe151, and Tyr42 are grouped together. No potential π electron cloud stacking of aromatic rings is seen, but the more commonly observed perpendicular arrangement is found, where a C-H moiety of one ring is directed into the plane of the second ring. This is exemplified by the side chain of His161, which is perpendicular to the benzene plane of Phe151; see Figure 553-1. The differences in the acccessible surface areas of Asn152 and Gln155 in the monomer and the dimer suggest that helix α-3 is somewhat involved in the dimer interface.

Table 553-1: Accessible surface area of residues at the C-terminus and in helix α-3 (Å²)

variant residue		A pH 8.2	A pH 6.2	B pH 7.1	A pH 7.1	A(in dimer) pH 7.1
ASN	152	70	63	64	67	50
PRO	153	59	55	58	55	55
THR	154	86	98	98	91	91
GLN	155	54	49	42	42	30
LEU	156	8	9	9	9	9
GLU	157	108	118	114	114	114
GLU	158	88	80	91	83	83
GLN	159	119	117	112	109	109
CYS	160	9	0	0	0	0
HIS	161	13	10	11	12	12
ILE	162	110	121	121	130	130
Total change between the monomer alone and the monomer in dimer exemplified by BLGA at pH 7.1						29

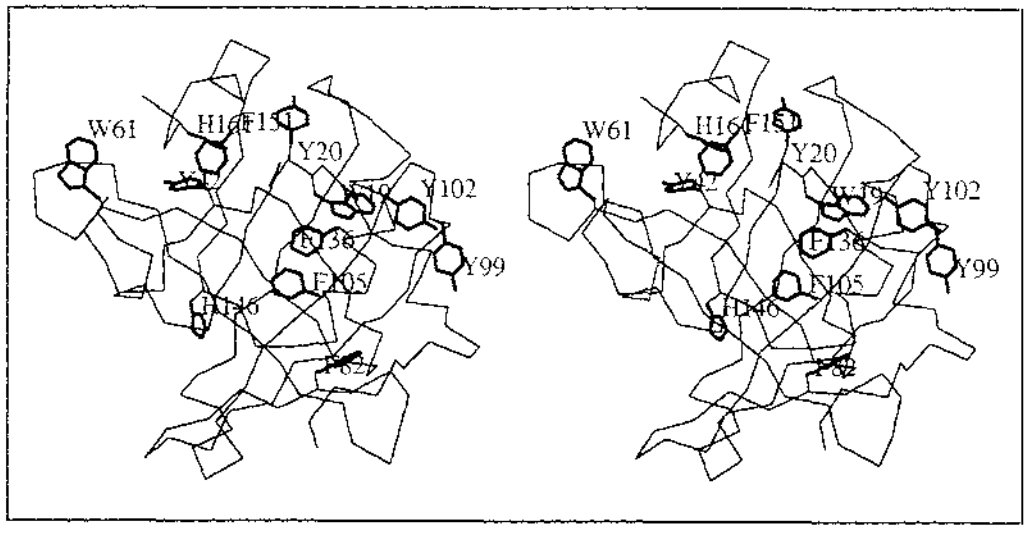


Figure 553-1: Location of aromatic side chains of BLG

5.5.4 α-H and ß-I region

The ß-I strand is an extension of ß-sheet II, and is not involved in the ß-barrel of BLG. This strand provides the hydrogen-bond interface for the formation of BLG dimers. The dimer interface will be discussed later, in section 5.9.1. The three-turn α-helix, which runs from residue 129 to residue 141, labelled as α-H, ends with a 3_{10} turn, as shown in Figure 554-1. The amphipathic nature of the helix is clearly seen in Figure 554-2. The side chains of the hydrophilic residues are on the top of BLG molecule and point to solvent. The side chains of hydrophobic residues Leu132, Ala133, Phe136 and Leu140, face towards the BLG molecule and interact with the side chains of strands ß-H and part 2 of ß-A. In addition, they form a small cavity around residue Cys121; see Figure 554-3. This cavity can hold an Hg^{2+} atom without serious disturbance to the surrounding secondary structures, as detailed in section 5.10.2. The differences in the accessible surface area of hydrophilic and hydrophobic residues of helix α-H are distinctive; see details in Table 554-1. The decrease in accessible surface area for Gln131 for BLGA at pH 7.1 is probably mostly an artefact of very high thermal motions of this side chain. The surface of the BLG molecule around α-H is displayed in Figure 554-4, where a patch of carboxylates (red) is followed by a line of lysines (blue). The positive and negative charges are mixed up around this area.

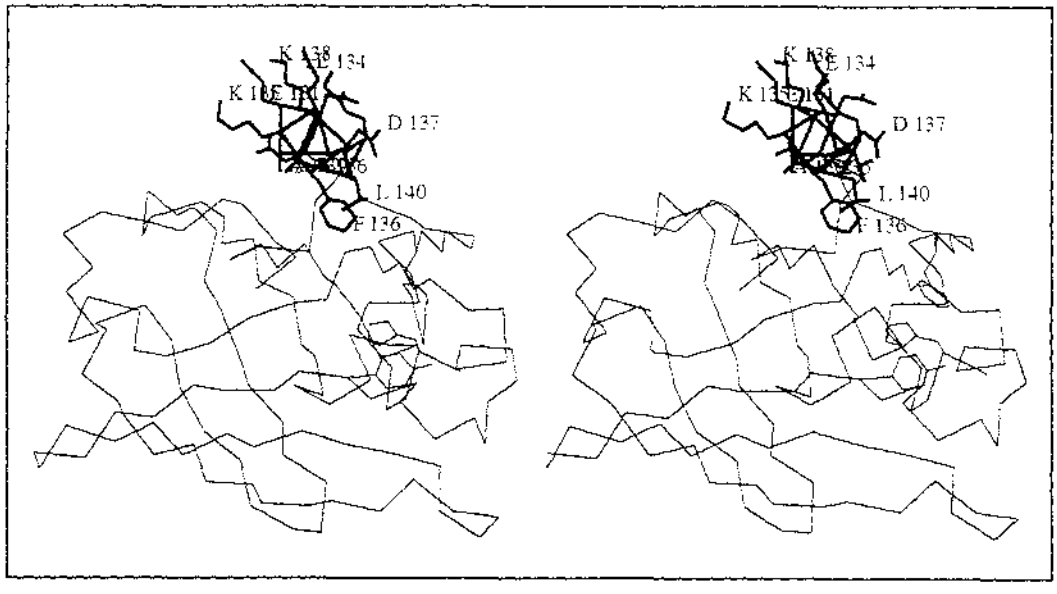


Figure 554-1: Orientation of α-H helix with respect to calyx.

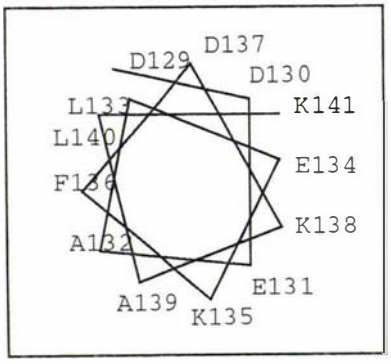


Figure 554-2: Amphipathic nature of helix α-H.

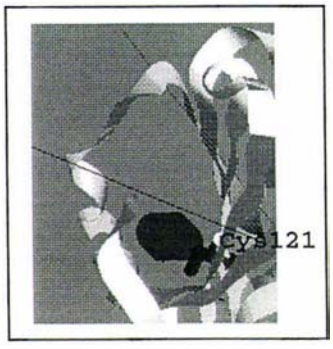


Figure 554-3 : Small cavity formed by the α-H helix.

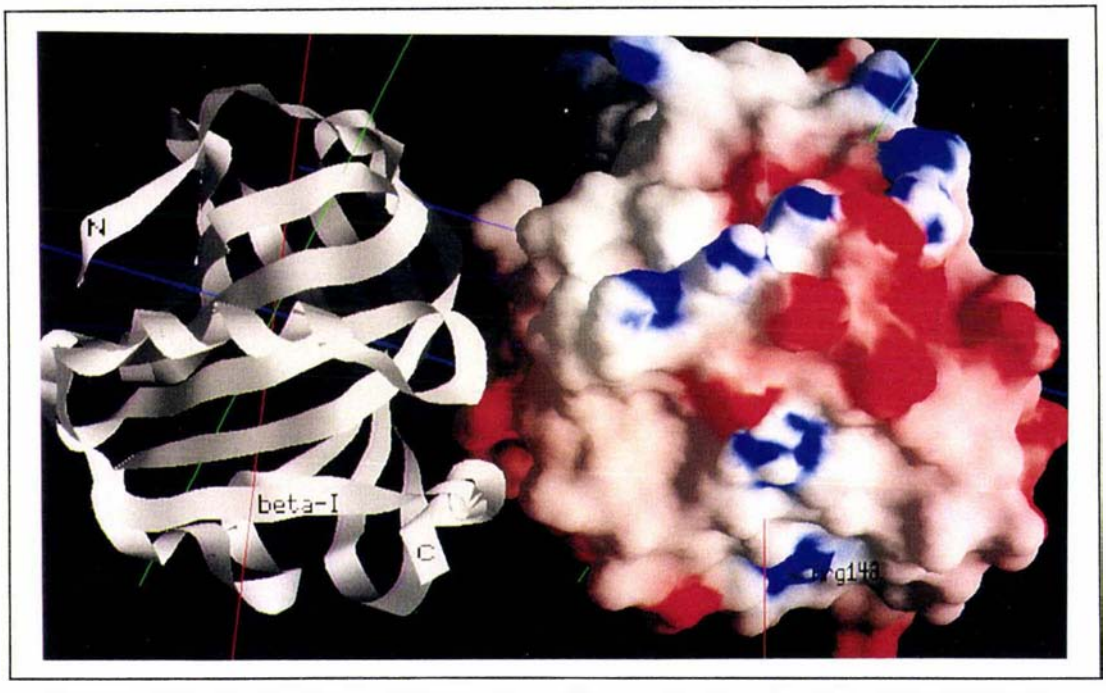


Figure 554-4: Surface charge of BLGA at pH 7.1 in the vicinity of α-H helix, which runs across the middle of the figure from right to left. Figure prepared by GRASP [Nicholls *et al.*, 1991].

Table 554-1: Accessible surface area of residues in the α-H region (Å^2)

variant		A pH 8.2	A pH 6.2	B pH 7.1	A pH 7.1	A(in dimer) pH 7.1
PRO	126	44	47	42	43	43
GLU	127	118	108	116	115	115
VAL	128	70	74	65	73	72
ASP	129	33	26	41	41	41
ASP	130	110	115	113	105	105
GLU	131	105	107	110	68	68

(continued)

Table 554-1: Accessible surface area of residues in the α-H region (Å²)

variant		A pH 8.2	A pH 6.2	B pH 7.1	A pH 7.1	A(in dimer) pH 7.1
ALA	132	0	0	2	0	0
LEU	133	20	20	22	20	20
GLU	134	106	103	94	97	97
LYS	135	85	74	82	64	64
PHE	136	2	1	5	3	3
ASP	137	63	66	63	64	64
LYS	138	138	104	105	116	116
ALA	139	16	11	11	10	10
LEU	140	17	15	26	17	17
LYS	141	167	167	160	160	160
ALA	142	87	89	77	80	80
LEU	143	15	16	24	18	18
PRO	144	38	46	34	39	39
MET	145	36	38	30	34	23
Total change between the monomer alone and the monomer in dimer exemplified by BLGA at pH 7.1						12

5.6 ß-Sheet I of the calyx

ß-sheet I is comprised of three ß strands, ß-B, ß-C and ß-D. The shortest strand in the ß-sheet I is longer than the longest one in the ß-sheet II. This may be one of the reasons why ß-sheet I is a twisted ß-sheet, see Figure 56-1. In general, the side chains of the hydrophobic and hydrophilic residues extend in opposite directions in the ß-sheet I. The hydrophobic side of ß-sheet I forms part of interior of the BLG calyx, while the hydrophilic side interacts directly with the solvent environment, as exemplified in Table 56-1, since the calyx handle region only partially covers the BLG calyx.

ß-B runs through Arg40, Val41, Tyr42, Val43, Glu44, Glu45, Leu46, Lys47, Pro48, Thr49 and Pro50. Three hydrophilic residues, Glu44, Glu45 and Lys47, extend to the molecular surface. The peptide φ/ψ angles for Val43, Glu44, Glu45 are -79/128, -101/-50 and -153/148, respectively. Thus, there is a kink in this strand, which causes Glu44 to be exposed.

Loop BC is a type I ß-turn [Wilmot, 1990]. This ß-turn is close to the bottom region, and contains four residues: Thr49, Pro50, Glu51 and Gly52. A hydrogen bond forms between G52N-T49O, 2.87 Å; see Figure 56-2. The

side chain of Glu51 and the carbonyl oxygen atom of Pro50 adopt a trans conformation, which is the most favorable energy conformation for a ß-turn. From Asp53 to Lys60 is the ß-strand C. While other hydrophobic residues are pointed to the inside of the BLG molecule, Leu57 is an exception. The side chain of Leu57 approaches the side chains of several hydrophilic residues (Glu44, Glu45, Lys47, Glu55, Gln59, Lys70, Gln68). ß-strand C is sandwiched between ß-strand B and D, the accessible surface areas of the residues are listed in Table 56-1. Gln59, the site of a point mutation between variants A or B and C (Q59H) is largely buried, through hydrogen bonding to the carboxylate group of Glu44. Loop CD has been described in the top portion of BLG.

ß-strand D runs from Ala67 to Lys77. ß-strand D and ß-strand E form an interface between the two ß-sheets, but only one residue is involved in the hydrogen bonding of this interface. In the middle of this ß-strand are three hydrophobic residues (Ile71, Ile72 and Ala73), whose side chains are in the vicinity of the entrance to the calyx. The two isoleucines have relatively large solvent-accessible areas of 51 and 76 Å² at pH 7.1. The closing of loop EF at low pH (and for BLGB at pH 7.1) reduces the solvent accessibility to Ile71. As is apparent in the topology diagram less than half of the ß-sheet potential of strand ß-D is used. As a result, alternation of sidechains between hydrophobic and hydrophilic residues does not occur in strand ß-D.

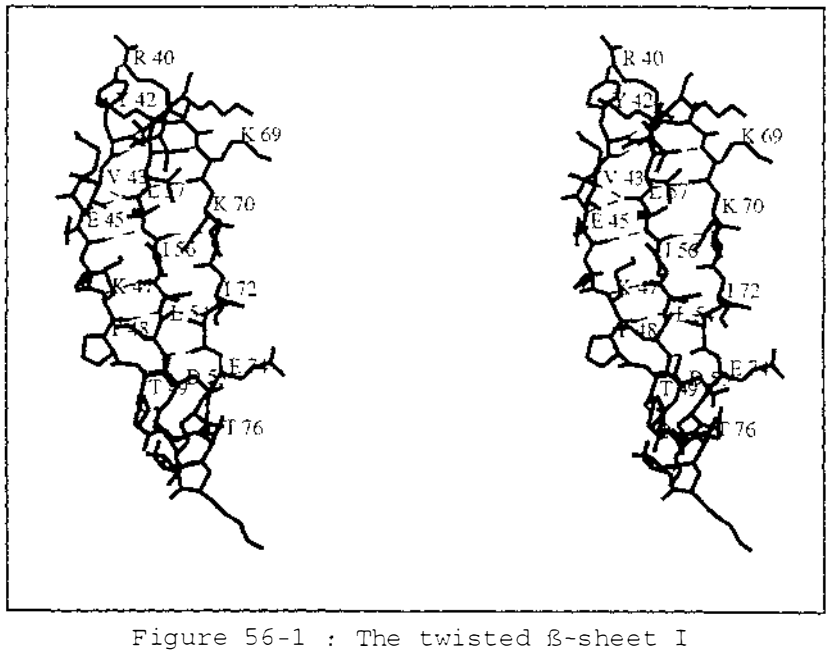


Figure 56-1 : The twisted ß-sheet I

Table 56-1: Accessible surface area of residues of ß-sheet I (Å²)

variant		A pH 8.2	A pH 6.2	B pH 7.1	A pH 7.1	A(in dimer) pH 7.1
ARG	40	21	24	28	23	9
VAL	41	8	12	5	33	33
TYR	42	3	0	1	0	0
VAL	43	4	5	5	5	5
GLU	44	66	57	64	68	68
GLU	45	59	63	51	68	68
LEU	46	14	12	8	13	13
LYS	47	88	73	77	74	74
PRO	48	32	24	27	27	27
THR	49	28	22	22	19	19
PRO	50	136	131	132	126	126
GLU	51	155	140	119	134	134
GLY	52	20	17	18	12	12
ASP	53	33	27	31	34	34
LEU	54	6	9	8	9	9
GLU	55	53	52	48	51	50
ILE	56	15	18	13	17	17
LEU	57	26	36	33	31	31
LEU	58	5	10	3	5	5
GLN	59	2	4	2	3	3
LYS	60	32	19	22	39	39
ALA	67	34	22	39	16	16
GLN	68	128	120	118	121	121
LYS	69	87	75	70	92	92
LYS	70	114	119	130	120	120
ILE	71	39	25	9	51	51
ILE	72	56	78	81	76	76
ALA	73	0	1	0	0	0
GLU	74	89	74	92	74	74
LYS	75	128	123	120	123	123
THR	76	36	44	48	41	41
LYS	77	211	202	199	206	206
Total change between the monomer alone and the monomer in dimer exemplified by BLGA at pH 7.1						14

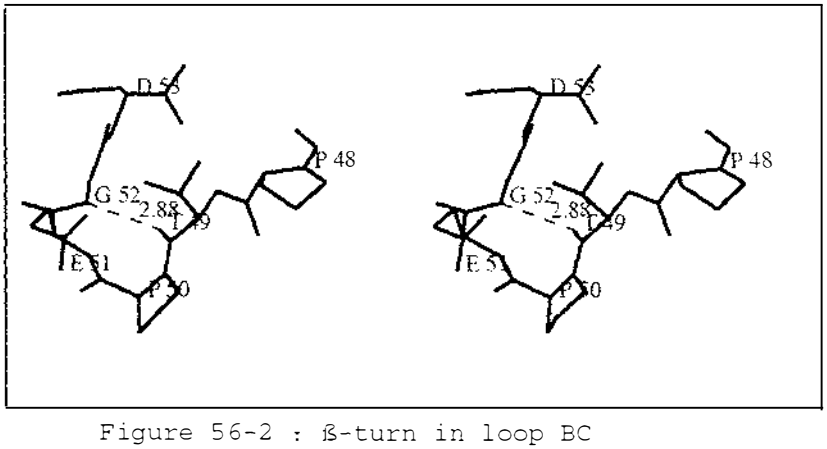


Figure 56-2 : ß-turn in loop BC

5.7 ß-Sheet II of the calyx

ß-Sheet II is made up of strands ß-E, ß-F, ß-G, ß-H and part 2 of ß-A, see Figure 57-1. All of these strands are covered by the calyx handle region. As ß-sheet II is largely shielded from solvent contact, not surprisingly, it is, in general, comprised of hydrophobic residues, no matter which side of the sheet. The loop ends of the strands are an exception, where side chains may be strongly hydrophilic. One example is the loop FG, Asp96 ~ Lys101. The handle region fails to cover this portion of structure, and results in a hydrophilic region. ß-Sheet II contains one disulfide bond Cys106-Cys119, which is well conserved through all BLG. The accessible surface area of residues of ß-sheet II are listed in Table 57-1.

Strand ß-E, which begins at residue 80 and ends at 85, is one of the boundaries of ß-sheet II with ß-sheet I. Only one residue, Lys83, forms a main chain hydrogen bond with the strand ß-D despite the unsued potentiality for hydrogen bonds between the strands ß-E and ß-D. Exposed loop EF with mainly hydrophilic residues has been discussed in detail in section 5.3.4.

Strand ß-F has different length according to pH for variant A. At low pH, ß-F starts from 90, instead of 89 as occurs at higher and neutral pH. The variant B at neutral pH bears similar characteristics to variant A at low pH, whose ß-F also starts from 90. The ß-F strand ends at Asp98. A kink can be identified at Leu95, which is located in the α-helix region in the Ramachandran plot (φ/ψ of Val94, Leu95 and Asp96 are -102/116, -84/-40 and -164/158, respectively). This provides the possibility to form a hydrogen bond between V94N and M7O. In addition, it places the side chain of Asp96 in a position to hydrogen bond with the OH group of Tyr102. Loop FG is close to the bottom region, and stabilized by hydrogen bond links, while Tyr99 adopts an unfavorable Ramachandran conformation (φ/ψ, 68/-38); see Figure 54-2. These hydrogen bonds are detailed in Table 57-2 and are conserved through all pH and BLG phenotypes.

Strand ß-G runs from Lys101 to Ser110. Tyr102 to Met107 is a hydrophobic region. Glu108 is one of few buried hydrophilic residues in BLG. This residue is buried by the N-terminus of BLG. The COO^- moiety forms a salt

bridge with the NH_3^+ group of Leu1. ß-H contains one mutation site, residue 118, which is different according to variant A (Val118) or B (Ala118). The electron density maps are well defined in both variants around residue 118; see Figure 57-2. Strand ß-H ends at Thr125. The last two residues of ß-H, Arg124 and Thr125, have been categorized into the bottom portion of BLG and described in the section 4.4.

Part 2 of strand ß-A runs from Leu22 to Ser27. This part of sheet II provides a hydrogen bond interface to the attached ß-I strand. Introducing a bulky moiety to the small cavity by the free Cys121 may alter the secondary structures around this cavity, including α-H, ß-G, ß-H and part 2 of ß-A. This potential disturbance may be transferred to the ß-I strand by part 2 of ß-A, which may then modify the dimer interface of BLG.

Table 57-1: Accessible surface area of residues of ß-sheet II region (Å²)

variants		A	A	B	A	A(in dimer)
residue		pH 8.2	pH 6.2	pH 7.1	pH 7.1	pH 7.1
LEU	22	0	0	1	0	0
ALA	23	0	0	0	0	0
MET	24	0	0	1	0	0
ALA	25	0	0	0	0	0
ALA	26	0	0	0	0	0
SER	27	4	8	9	7	7
ALA	80	0	0	0	0	0
VAL	81	16	20	15	21	21
PHE	82	0	3	4	0	0
LYS	83	113	117	119	96	96
ILE	84	25	11	4	24	24
LYS	91	48	71	55	63	63
VAL	92	7	11	7	12	12
LEU	93	3	12	8	13	13
VAL	94	8	7	10	12	12
LEU	95	6	3	15	4	4
ASP	96	36	31	45	34	34
THR	97	4	7	4	6	6
ASP	98	40	44	42	35	35
TYR	99	38	37	37	30	30
LYS	100	142	133	158	138	138
LYS	101	73	71	86	70	70
TYR	102	14	17	21	10	10

(continued)

Table 57-1: Accessible surface area of residues of ß-sheet II (Å²)

variant residue		A pH 8.2	A pH 6.2	B pH 7.1	A pH 7.1	A(in dimer) pH 7.1
LEU	103	6	5	5	3	3
LEU	104	3	1	4	2	2
PHE	105	20	19	19	35	35
CYS	106	1	1	0	1	1
MET	107	18	16	15	40	40
GLU	108	5	9	3	4	4
VAL/Ala	118	0	0	8	5	5
CYS	119	1	1	4	1	1
GLN	120	0	0	9	4	4
CYS	121	0	0	0	0	0
LEU	122	0	2	2	0	0
VAL	123	5	3	3	3	3
Total change between the monomer alone and the monomer in dimer exemplified by BLGA at pH 7.1						0

Table 57-2 : The hydrogen bond network stablising the γ-turn at Tyr99

H donor	distance (Å)	H acceptor
Y99O	3.26	R124NE
Y99O	3.44	R124NH2
D98OD1	3.09	Y102N
D98O1	2.87	K101N
D98O	2.77	K100N
Y99OH	2.84	D11N

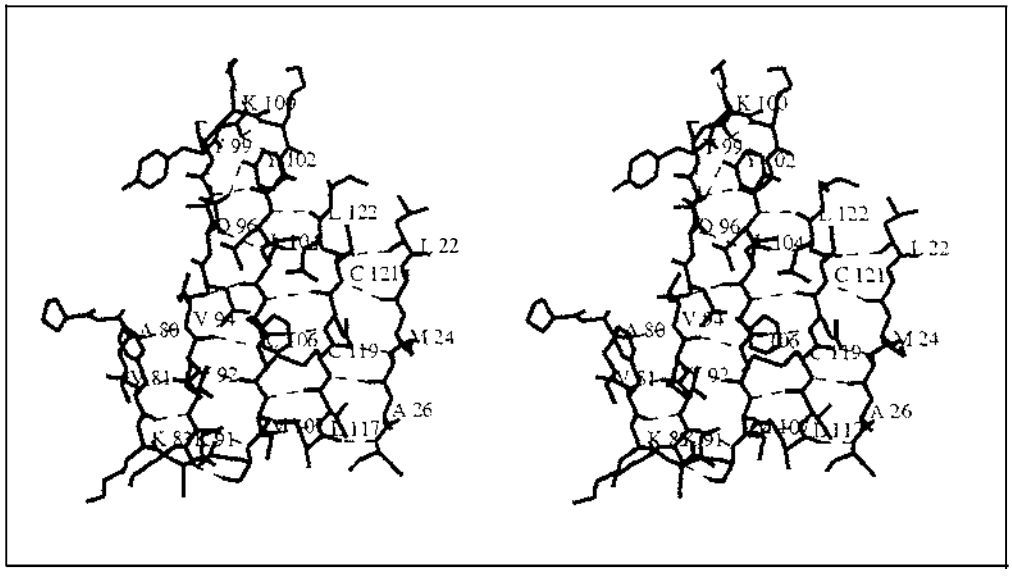


Figure 57-1 : ß-sheet II

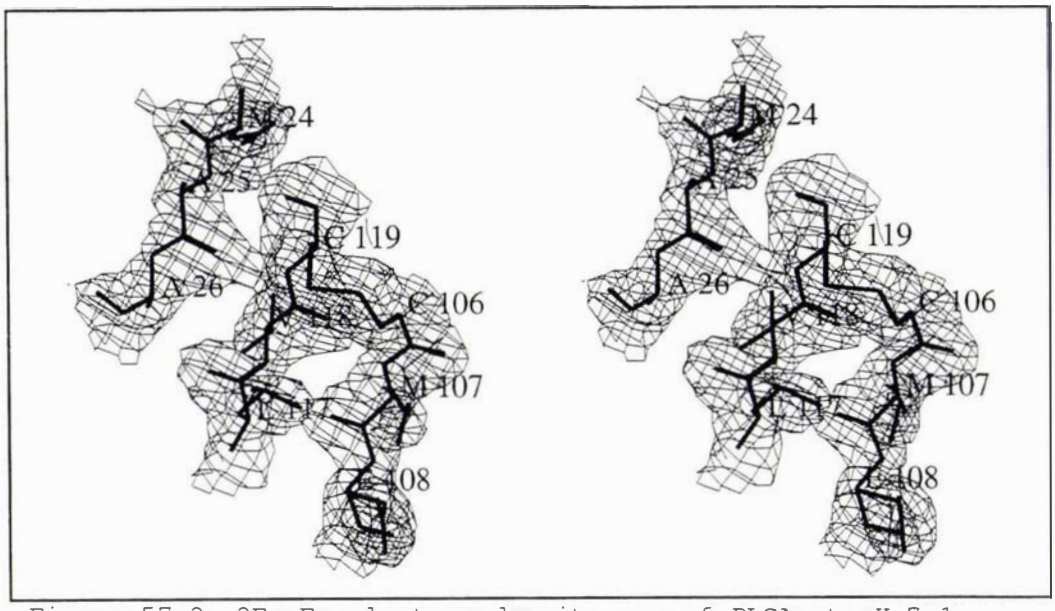


Figure 57-2: 2Fo-Fc electron density map of BLGA at pH 7.1, showing mutation site Val118/Ala118. BLGA (thick lines), BLGB (thin lines). Figure prepared by TURBO [Cambillau *et al.*, 1996].

5.8 Highlights of selected residues

5.8.1 Cysteines and methionines

The disulfide bonds of BLGA in lattice Z are clearly visible in the electron density maps, not only that linking Cys119 to Cys106, but also that linking Cys160 to Cys66. The electron density of disulfide bond Cys160-66 is well displayed in Figure 533-1. At the experimental conditions, the configurations and the accessibilities of these two disulfide bonds are different. The distance between Cα106 and Cα119 is only 3.83 Å, while the distance between Cα160 and Cα66 is 5.91 Å, the former disulfide bond appears more rigid and is unusual in that it links adjacent ß-strands [Thornton, 1981]. The solvent accessibility of the disulfide bond at Cys160-Cys66 is 36.0 $Å^2$ and coincides with a higher than average B factor for this region (for atoms within 6 Å of Cys66 and Cys160, average B = 58.0 $Å^2$), while the other disulfide bond is only 1.0 $Å^2$ exposed and is located in a very well ordered region with lower than average B factors (for atoms within 6 Å of Cys106 and Cys119, average B = 30.6 $Å^2$). Bovine BLGA or BLGB contains four methionines (Met7, Met24, Met107 and Met145) with relatively low B factors.

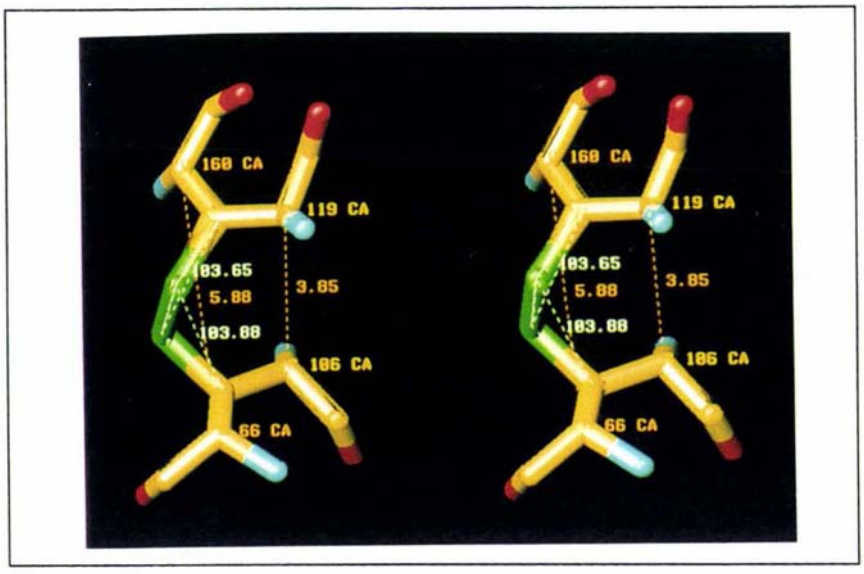


Figure 581-1: Superposition of the two disulfide bonds of BLG, showing the different conformations.

5.8.2 Aromatic residues Tyr, Trp, Phe and His

Solution studies of ligand binding to BLG often utilise the fluorescence from aromatic residues [Gorbunoff, 1967; Townend *et al.*, 1969; Tanford, 1959; Pantaloni, 1965]. Thus, it is important to define the location of these residues and their exposure. BLGA and BLGB contain four Tyr and two Trp residues. In lattice Z, Tyr42 is totally buried, while Tyr20 is the most exposed. Tyr99 and Tyr102 are partially exposed; see Table 54-1, 56-1 and 57-1. Tyr99 is involved in a hydrogen bond complex which allows it to sit in an unfavorable Ramachandran plot area. Four Phe residues exist in the BLGA and BLGB molecule. Two of them, Phe82 and Phe136 are buried residues. The other two, Phe105 and Phe151 are partially exposed, with exposed areas of 35 $Å^2$ and 16 $Å^2$, respectively. The side chain of Trp19 is buried, while that of Trp61 is exposed. The pH in the range 6.2 ~ 8.2 has limited influence on the solvent-accessible areas of these residues. Both of Trp61 and Trp19 are surrounded by hydrophilic residues. An exposed His has a protonated nitrogen at pH below 7. BLG contains only two His; If His146 is assigned as an exposed residue, His161 has to be regarded as a buried one, whose ND1 and NE2 atoms are not accessible. The distribution of aromatic residues in the BLGA molecule is shown in Figure 553-1. However, on dimer formation, His146 becomes substantially buried (For BLGA at pH 7.1, from 126 to 47 $Å^2$; see Table 591-1).

5.8.3 Charged residues: Arg, Lys, Glu and Asp

BLG contains three arginines: Arg40, Arg124 and Arg148. Arg40 is a buried residue which forms several hydrogen bonds with its neighbors: Arg40NH1-OLeu32, Arg40NH1-OAsp33, Arg40N-OAla37, Arg40NE-OHTyr42. Arg124 is a key residue at the bottom of BLG. The lengthy side chain generally brings relatively high B factors to the terminal atoms. However, the special hydrogen bonds of the guanidinium plane of Arg124 lead to low B factors for the terminal atoms of Arg124, see Table 583-1. Bovine BLG contains 15 lysines. As documented in Table 583-2, most Lys have relatively high B values for the terminal NZ atom. However, one exception is Lys8. The reason is that, in lattice Z, Lys8 serves as a "key" and is inserted to the "lock" provided by another molecule. The NZ atom of Lys8 forms several hydrogen bonds, and as a result, the B factors of Lys8 actually decrease along the chain to the end atom, NZ.

In contrast to Arg and Lys, Glu and Asp are negatively charged residues. There are 16 Glu and 11 Asp in BLGA, and one less Asp in BLGB, due to the mutation Asp64Gly, from A to B. Where are these positively charged residues distributed over the BLG molecule? The top, ß-sheet I and α-H are the three regions where Glu and Asp are concentrated; see Figure 583-1. One Glu is buried at pH 7.1, which is Glu108. This residue forms a salt bridge with the α-amino group of residue 1. At low pH and for BLGB at pH 7.1, Glu89 is also buried. In general, the other Glu and Asp point out into the solvent environment.

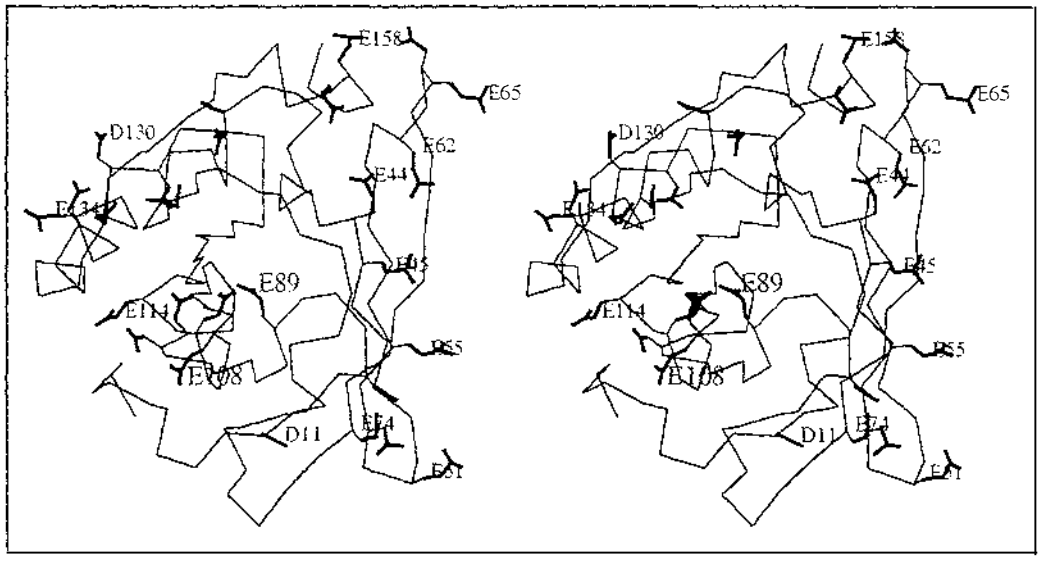


Figure 583-1: Location of Glu and Asp side chains

Table 583-1 : B factors of the atoms of arginine residues

atom of	Arg40		Arg124		Arg148	
	B(Å^2)	access_area (Å^2)	B(Å^2)	access_area (Å^2)	B(Å^2)	access_area (Å^2)
C	37.4	0	38.3	0	34.2	0
N	37.3	0	33.1	1	36.4	7
O	39.3	0	37.2	10	34.4	26
CA	36.2	0	36.5	0	35.4	0
CB	37.5	0	36.2	9	40.0	18
CG	39.9	0	36.0	1	49.6	0
CD	39.8	0	34.5	9	57.3	21
NE	41.6	0	34.8	4	63.7	1
CZ	41.2	0	33.0	5	67.5	2
NH1	42.5	0	33.3	5	69.9	21
NH2	41.7	23	33.9	9	69.5	4

Table 583-2 : B factors of the atoms of lysine residues

Lys	C	N	O	CA	CB	CG	CD	CE	NZ
8	42.4	41.9	42.1	42.3	44.1	45.1	44.2	41.8	40.7
14	35.9	39.6	35.8	38.6	40.9	45.6	48.7	51.9	54.5
47	50.2	48.1	50.0	51.1	55.2	59.5	63.8	66.5	68.2
60	65.6	54.7	68.4	60.5	62.4	66.1	68.6	71.6	73.8
69	61.6	65.7	59.5	62.5	62.3	64.8	66.1	67.5	68.0
70	64.6	63.9	64.7	65.0	66.3	66.9	68.2	67.8	68.0
75	51.0	49.0	51.7	51.6	56.0	60.3	62.8	65.2	67.2
77	48.2	50.1	48.4	50.4	57.2	64.4	69.0	71.1	70.8
83	55.1	46.0	53.8	52.4	56.4	62.1	64.4	66.8	67.4
91	42.7	52.6	41.6	47.1	50.2	56.7	62.2	65.8	67.6
100	46.5	43.1	46.9	45.4	49.4	52.9	58.1	61.3	64.0
101	44.5	45.7	47.1	44.8	49.2	53.8	57.9	60.0	61.9
135	33.9	33.9	34.7	33.0	37.1	38.0	38.5	38.2	44.4
138	41.0	36.5	45.8	40.4	45.8	53.0	60.3	64.3	65.6
141	48.6	43.1	51.5	48.2	50.5	56.3	62.3	65.6	69.3

5.8.4 Aliphatic hydrophobic residues: Leu, Ile, Val and Ala

Not all hydrophobic residues are buried, for example, at the N-terminus, Leu1, and at the C-terminus, Ile162, are well exposed hydrophobic residues. In ß-sheet I, the side chains of hydrophobic residues tend to adopt an orientation towards the interior of the BLG molecule. Leu57 and Ile 72 are two exceptions. Because ß-sheet II is covered by the calyx handle region, except for the loops, ß-sheet II is mainly comprised of hydrophobic residues. The hydrophobic residues of the handle region face the ß-sheet II, except ß-I which is not exposed in solution due to dimer formation, and which contains several hydrophobic residues.

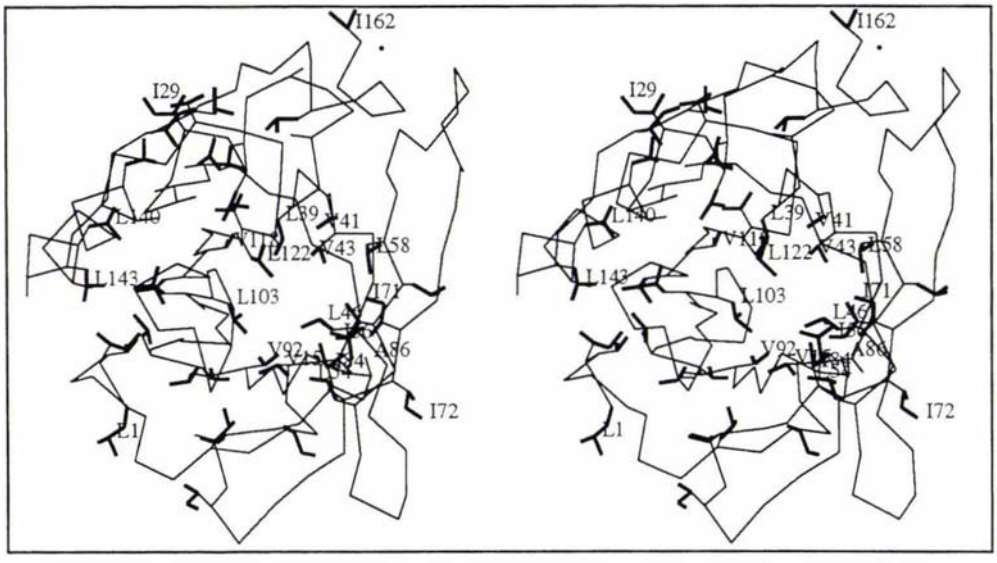


Figure 584-1: Location of aliphatic side-chain residues

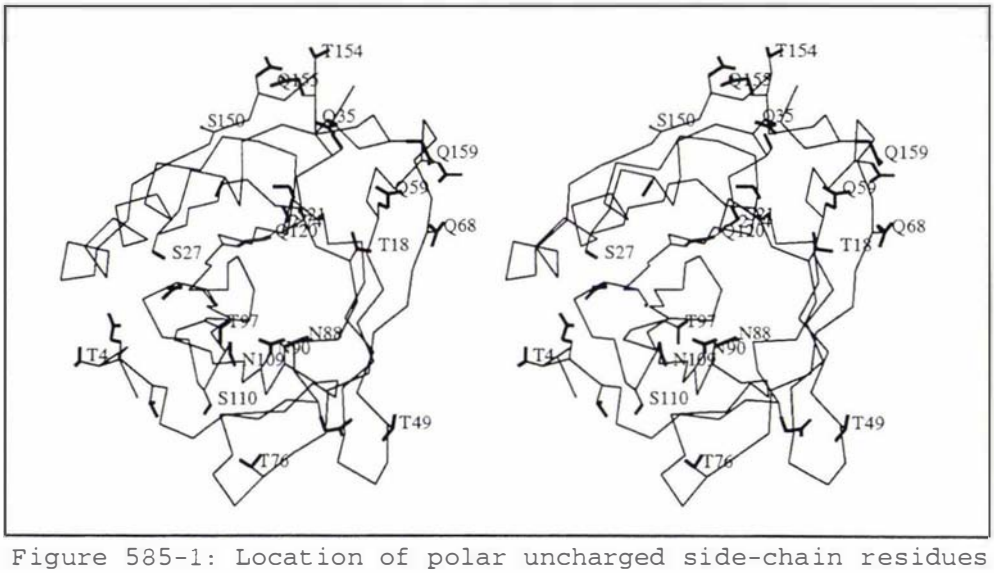


Figure 585-1: Location of polar uncharged side-chain residues

5.8.5 Polar uncharged residues: Thr, Ser, Asn and Gln

There are five Asn, nine Gln, seven Ser and eight Thr in bovine BLGA and BLGB molecules. These residues are concentrated on the N-terminus, helix α-3 and the top region of the molecule. The side chains of these polar uncharged residues generally point to the solvent environment, except for Gln120, Ser21 and Thr97, which have small solvent accessible area of 4 $Å^2$, 0 $Å^2$ and 6 $Å^2$ for BLGA at pH 7.1, respectively. The side chains of Gln120 and Ser21 approach together and form a strong hydrogen bond (Gln120NE2...OG21Ser, 2.74 Å). Note that Thr97 is part of the conserved motif of TDY in all ß-lactoglobulins; see details in section 2.3.4. Gln120, Ser21 and Thr97 are also buried in the pH 6.2 and 8.2 structures

for BLGA and at pH 7.1 for BLGB. Figure 585-1 shows the location of the polar uncharged residues in bovine BLGA at pH 7.1.

5.8.6 Proline and Glycine

The three glycines in BLGA occur at positions 9, 17 and 52. After the point mutation D64G, BLGB has one glycine more than BLGA. There are eight proline residues in BLGA and BLGB. Without exception, all these residues are sitting as expected at turning points, as shown in Figure 586-1. The prolines in BLGB are generally well exposed (solvent accessible area 27 ~ 132 $Å^2$, average 56.6 $Å^2$), as are the glycines (solvent accessible area 18 ~ 71 $Å^2$, average 41.3 $Å^2$).

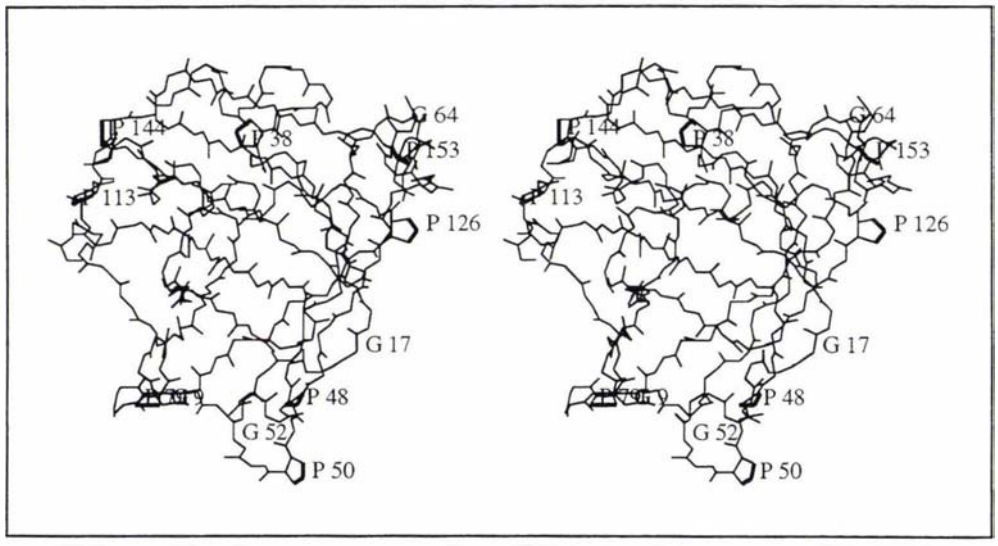


Figure 586-1: Location of prolines and glycines in BLGB

5.9 Molecular packing in lattice Z

The space group of lattice Z is $P3_221$. There are six asymmetric units in the unit cell. Each asymmetric unit contains one molecule. The contents of the unit cell of bovine BLG in lattice Z are shown in Figure 59-1. The BLG molecules in lattice Z are packed layer by layer. Each layer is comprised of parallel linear arrays of BLG monomers with dimer interfaces alternately pointing up and down, as shown in Figure 592-2. The unit cell volume expands with increase in pH in the case of BLGA. BLGB at pH 7.1 has similar crystal parameters to BLGA at pH 8.2 (see Table 467-2), although individual molecules of BLGB have conformation and surface area more similar to BLGA at pH 6.2.

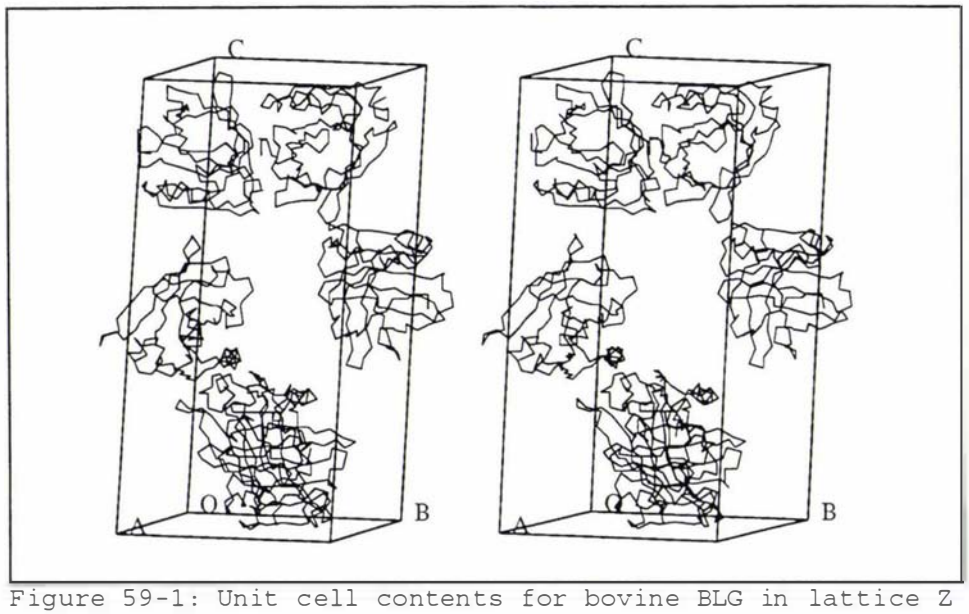


Figure 59-1: Unit cell contents for bovine BLG in lattice Z

5.9.1 Dimer interface

The dimer interface is shown in Figure 591-1, which has been previously described in the structure of BLG in lattice X [Brownlow *et al.*, 1997]. This interface presumably persists in solution and is also found in lattice Y. In lattice Z, the dimer interface is built up from three portions of each monomer which are related by a two-fold crystallographic axis: (i) the ß-I strands of two monomers form the hydrogen-bonded part of the dimer interface, as illustrated in Figure 591-1; (ii) loops AB, especially residues 29-35, of two monomers form part of the dimer interface underneath the ß-I strand, as illustrated in Figure 591-1; (iii) the α-H helices of two monomers form a "V" type channel along the dimer interface above the ß-I strand, as illustrated in Figure 591-1. Ordered water molecules are found in this channel at pH 6.2, 7.1 and 8.2. The dimer interface buries 972 $Å^2$ of the accessible area of two molecules at pH 7.1 (946 $Å^2$ at pH 6.2, 1013 $Å^2$ at pH 8.2). The accessible surface areas of residues of the ß-I strand change substantially on formation of the dimer. Table 591-1 lists these data in detail. Much (334 $Å^2$) of the surface area buried per molecule (total 486 $Å^2$) is associated with residues of strand ß-I; the remainder (132 $Å^2$) is associated with residues of loop AB (see Table 532-1), especially Ile29 and Asp33; and the solvent accessible area of residue Gln155 of helix α-3 decreases upon

the formation of the dimer interface (12 Å^2, Table 553-1).

Table 591-1: Accessible surface area of residues of strand ß-I (Å^2)

variant		A pH 8.2	A pH 6.2	B pH 7.1	A pH 7.1	A(in dimer) pH 7.1
MET	145	36	38	30	34	23
HIS	146	127	130	126	126	47
ILE	147	23	25	23	23	2
ARG	148	94	100	89	100	41
LEU	149	38	33	37	40	0
SER	150	64	59	63	62	7
PHE	151	16	21	20	55	3
ASN	152	70	63	64	67	50
Total change between the monomer alone and the monomer in dimer exemplified by BLGA at pH 7.1						334

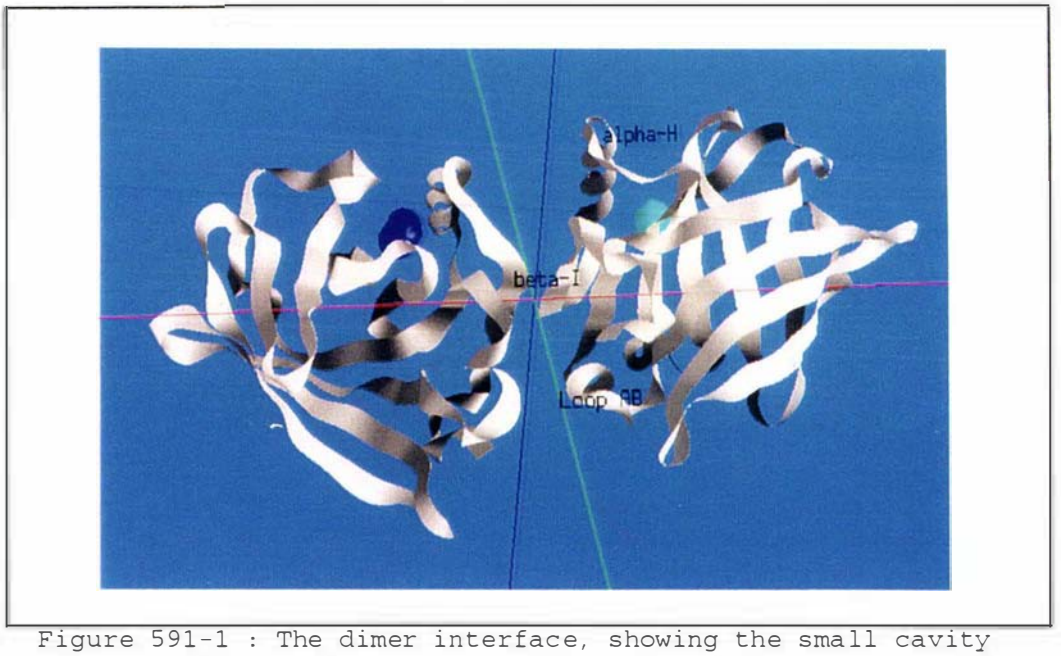


Figure 591-1 : The dimer interface, showing the small cavity around Cys121. Bovine BLG dimer is centred between the two ß-I strands. Figure prepared with GRASP [Nicholls *et al.*, 1991].

5.9.2 "Lock and Key" interface

The interface illustrated in Figure 592-1 involves the side chain of Lys8 from one molecule as a "key", which inserts into a "lock" provided by a neighboring molecule. The "lock" is formed by Tyr20 and Ser21 from the end of first half of the ß-A strand; Val41, Tyr42 and Val43 from the middle of the ß-B strand, and Leu156, Glu157 and Glu158 from the α-3

helix, together with the side chain of His161. The bottom of this hole contains a water molecule which bridges the peptide moieties of residues Val43 and Ser21 via hydrogen bonds. This water molecule is also hydrogen bonded to the inserted lysine's ammonium group, and is conserved in all of the structures of BLGA and BLGB. The insertion of Lys8 leads to formation of several hydrogen bonds, as detailed in Table 592-1. In addition to these hydrogen bonds, van der Waals contacts of this interface are also quite extensive, a total of 1046 Å^2 of accessible area of two molecules is buried at pH 7.1 (1093 Å^2 at pH 6.2, 986 Å^2 at pH 8.2). The "lock and key" interface is quite substantial, and leads to a linear chain array of BLGA molecules in lattice Z, as illustrated in Figure 592-2.

Table 592-1 : Hydrogen bonding in the "lock and key" interface

"key" side residue	atom	distance(Å)	distance(Å)	"lock" side atom	residue
Lys8	NZ	2.86		OE2	Glu44
Lys8	NZ	2.94		O	Val43
Lys8	NZ	2.84		O	Trp19
Lys8	NZ	3.03 water	2.81	N	Ser21
Lys8	NZ	3.03 water	2.81	O	Val43
Thr4	O	2.86		N	Thr18
Thr6	N	3.13		OG1	Thr18
Thr6	O	2.96		OG1	Thr18
Gly9	N	2.91		OE1	Glu157

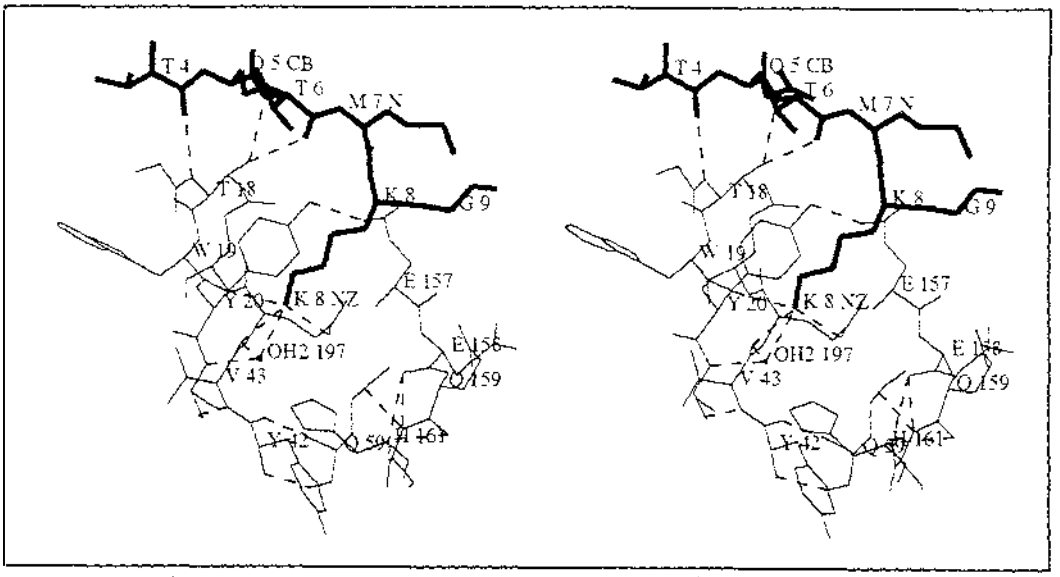


Figure 592-1 : The "lock and key" interface.
Key, thick lines. Lock, thin lines

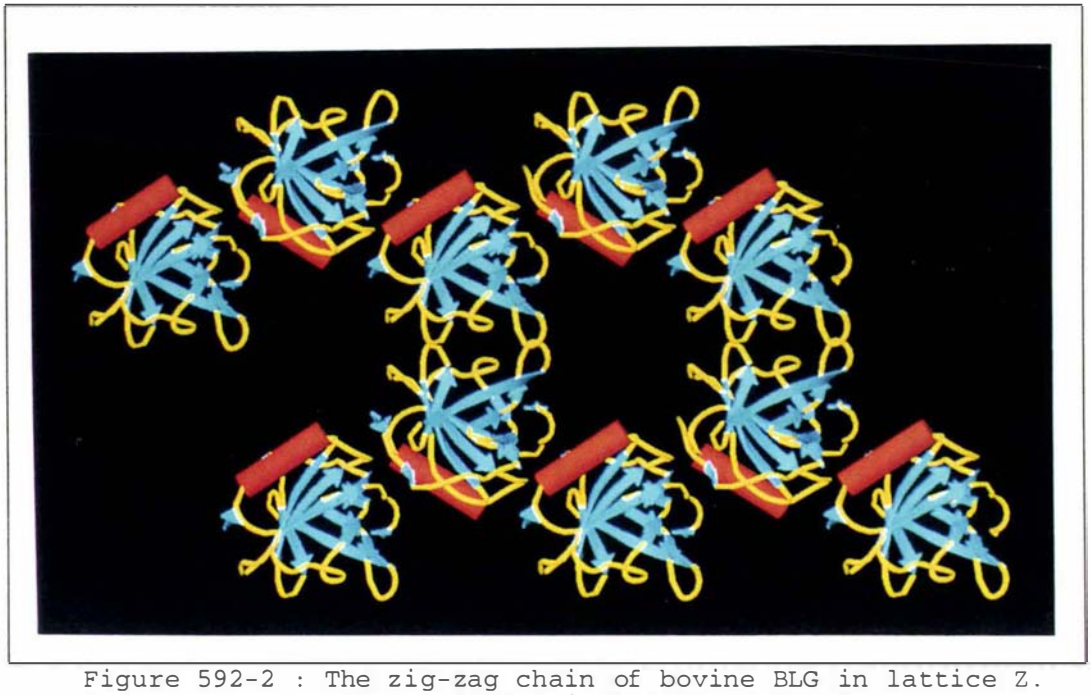


Figure 592-2 : The zig-zag chain of bovine BLG in lattice Z. A crystallographic two-fold axis runs left to right across the middle of the diagram.

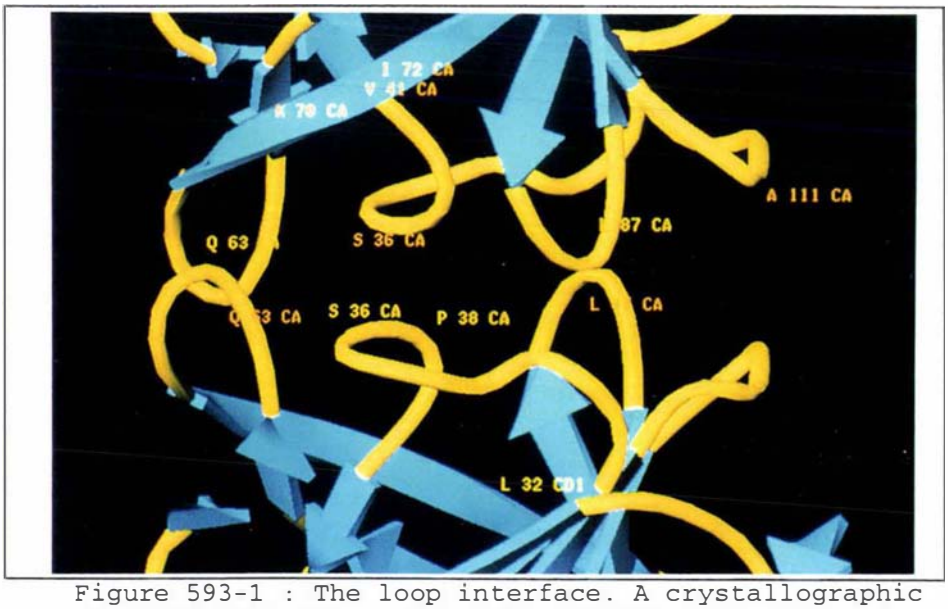


Figure 593-1 : The loop interface. A crystallographic two-fold axis runs left-right through the middle of the diagram.

5.9.3 Loop interface

The loop interface is primarily comprised of the top region of BLG, see Figure 593-1. Space group $P3_221$ has two two-fold crystallographic symmetry axes. One of them lies between the dimer interface described above, while the other two-fold axis places pairs of the loops CD, EF

and GH into mutual contact. There do not appear to be any specific interactions among these rather mobile regions (this is not true in the case of BLGA-BrC12, where Lys69NZ forms a hydrogen bond, 2.68 Å, with Glu62OE2 of its symmetry mate), whose proximity caused some difficulty in the interpretation of electron density maps. These loops, especially loop EF, form the doorway to the interior of BLGA calyx.

5.10 Ligand-binding sites of BLG

5.10.1 Proposed retinol-binding site

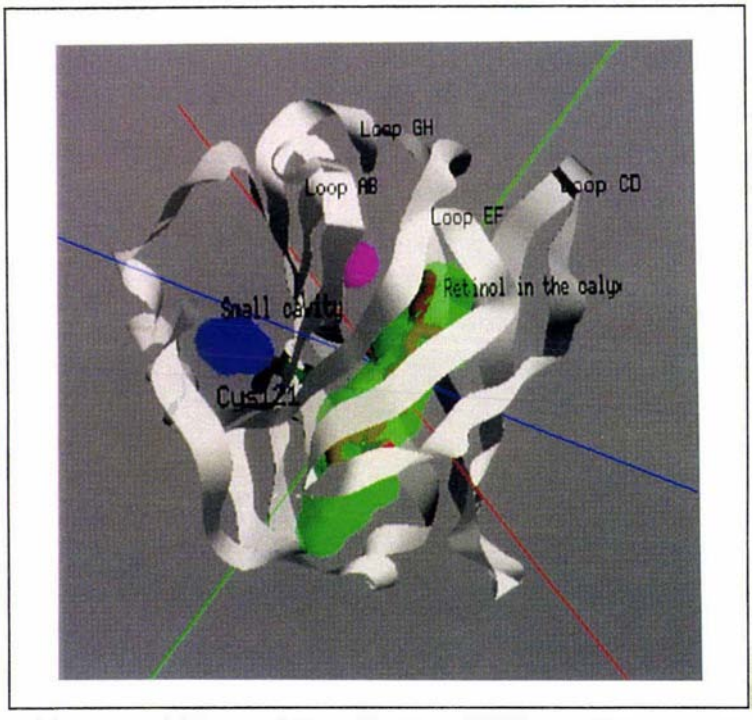


Figure 5101-1: Retinol modelled in the proposed ligand-binding site.

The inside of the ß-lactoglobulin calyx offers a large hydrophobic cavity and would appear to be the major ligand-binding site; see Figure 5101-1. Unfortunately, until very recently, we and others have not obtained crystals where BLGA or BLGB binds ligand in this site. The surface of the interior of the calyx, which involves sidechains of residues Leu10, Ile12, Val15, Val41, Val43, Leu46, Leu54, Ile56, Leu58, Ile71, Ala73, Ala80, Phe82, Ile84, Val92, Val94, Leu103, Phe105 and Met107, appears to complement well to the surface of retinol, which has been fitted mannally into the cavity, as shown in Figure 5101-1. This complementarity is also present at pH 6.2 and pH 8.2, except that at pH 6.2 the retinol may be

totally buried underneath loop EF, and the OE1 atom of Glu89 may form a hydrogen bond with the retinol OH group, while the protonated OE2 atom of Glu89 is hydrogen bonded to the peptide O atom of Ser116. Interaction of the sidechain of Trp19 and the ß-ionone ring of retinol, proposed as important in the binding of retinol, is blocked by the sidechains of residues Leu46 and Leu103 in this model.

The proposed binding site of BLG for retinol is actually filled with something at low pH. A string of water molecules has been identified inside the calyx of bovine BLGA at pH 6.2; see Figure 5101-2. The electron density of this water string is continuous, with the distances between waters around 2.8 ~ 3.0 Å. Except for one water molecule (w1 in Figure 5101-2), which hydrogen bonds to the NZ atom of Lys69 (2.76 Å), the other water molecules hydrogen bond only to each other.

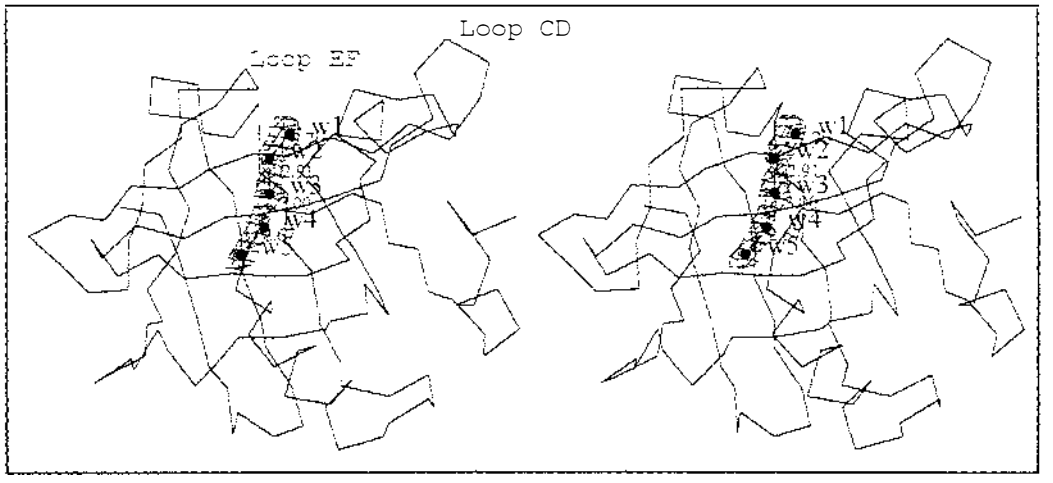


Figure 5101-2: 2Fo-Fc electron density map of string of water molecules (w1 - w5) inside the calyx of BLGA at pH 6.2.

5.10.2 Mercury binding site

Besides the binding site inside the calyx discussed above, another binding site for small molecules appears to exist. In lattice Z, part of ß-sheet II (ß-F, ß-G, ß-H and part 2 of ß-A) and ß-I create a flat ß-sheet, as shown in Figure 5101-1, where the BLGA calyx is on the right side and the loop GH is on its top edge. On the left, this ß-sheet is covered by the N-terminus and alpha-H, thereby creating a small cavity (coloured blue in Figure 5101-1) adjacent to the SG atom of Cys121, in which a methyl mercuric ion binds to the SG atom

of Cys121. This occurs with negligible rearrangement of main-chain and side-chain groups, as shown in Figure 5102-1. However, introducing a bulky moiety to this small cavity may alter the tertiary structure, such as the relative positions of α-H helix and ß-A strand, surrounding this site. The disturbance may be transferred to the ß-I strand through the ß-sheet and indirectly influence the dimer interface of ß-lactoglobulin.

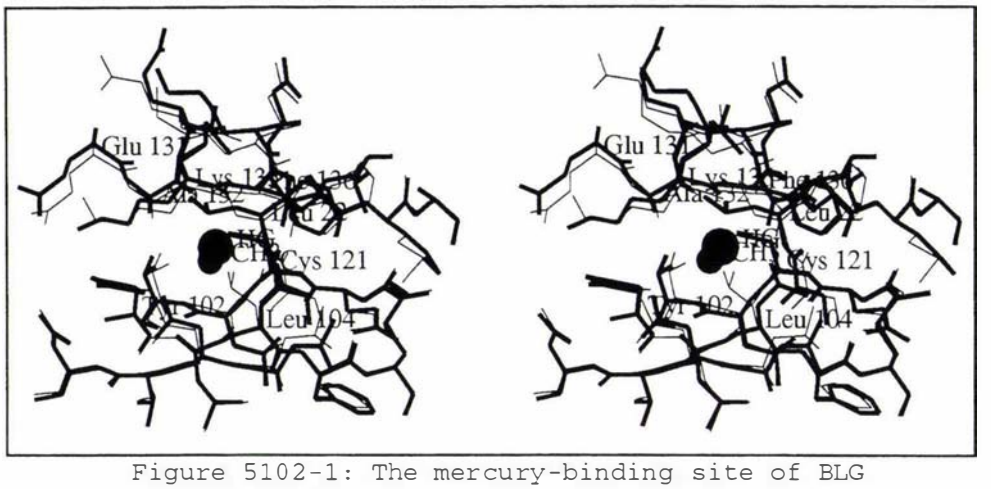


Figure 5102-1: The mercury-binding site of BLG

Side chain of Leu104 rotates 90° to create space for the CH_3 group of the methyl mercury ion. The thin line represents native BLGA in lattice Y; The thick line represents the mercury-bound BLGA in lattice Y. The coordinates are from Jameson (1998) with R=0.19, R_f=0.24.

5.10.3 Binding of fatty acids

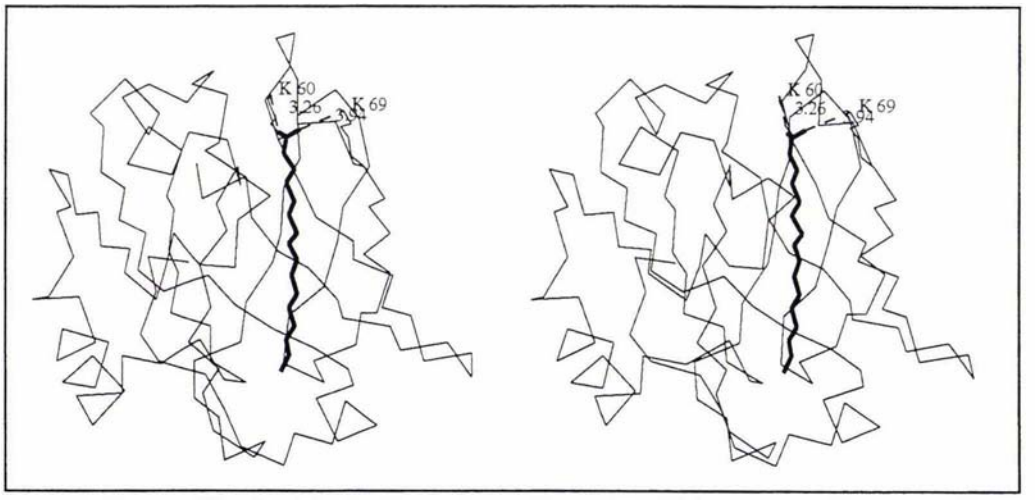


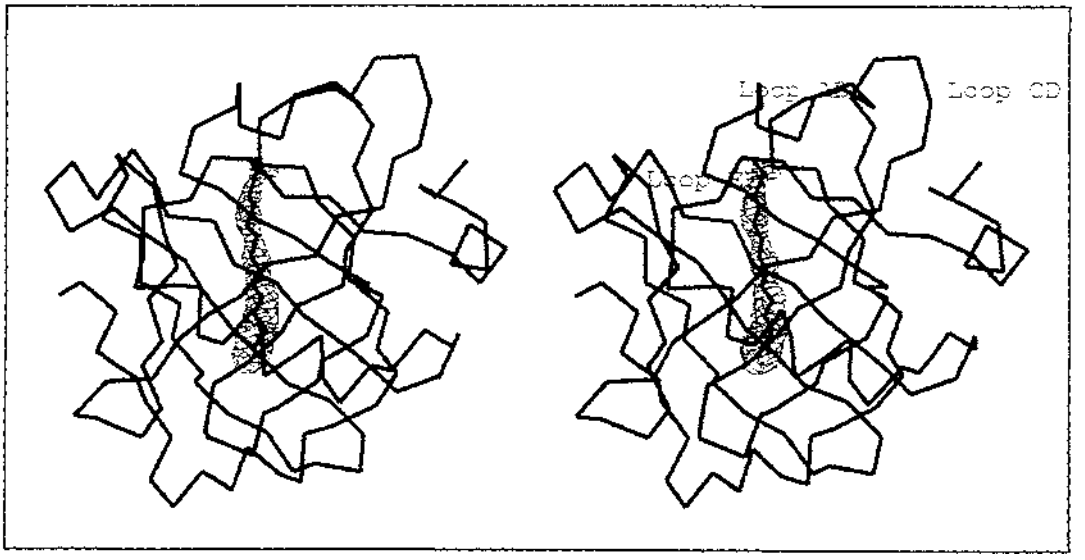
Figure 5103-1: Model of palmitic acid binding inside of the BLG calyx.

As well as binding retinol, bovine BLG binds a range of fatty acids. The aliphatic fatty acid, palmitic acid, is modelled in the major ligand-

binding site, as shown in Figure 5103-1. Lys60 and Lys69 may form a salt bridge with the COO^- group of palmitic acid. As the aliphatic side chain is not rigid, both longer and shorter aliphatic fatty acids than palmitic acid may be accommodated inside the calyx.

5.10.4 Binding of 12-bromododecanoic acid to BLGA at pH 7.3.

Bovine BLGA was successfully cocrystallized with the ligand 12-bromododecanoic acid (BrC12), at pH 7.3. The BLGA-BrC12 model has values for R (R_f) of 23.23% (27.93%) and good geometry, as summarized in Table 512-1. While the existence of the ligand BrC12 in the centre of the BLG calyx was revealed immediately in the difference map calculated following molecular replacement, as detailed in Figure 476-1, the refined structure provides a clearer electron density map. Figure 5104-1 is an omit difference map contoured at +3 σ.



Figur 5104-1: Difference electron density map of BLGA-BrC12, showing BrC12.

Chapter 6 Discussion

6.1 Structural basis of the Tanford transition

6.1.1 Tanford (N⇔R) transition

The Tanford transition was first identified on the basis of a titration study on a mixture of variants A and B in 1959. The results of the titration experiments suggested two possibilities for a titratable residue at ~ pH 7.5 [Tanford *et al.*, 1959a]: (i) the particular side chain group might be an imidazole group; or (ii) a carboxyl group buried at low pH becomes available for titration at higher pH (that is an abnormal COOH group). The purified variants A and B were subsequently studied and found to have similar pH-dependent titration behaviour [Tanford and Nozaki, 1959b], as shown in Figure 611-1. The titration of goat BLG shows similar behaviour to the bovine BLGA and BLGB [Ghose *et al.*, 1968], and an abnormal carboxylate group was again proposed as the titratable group.

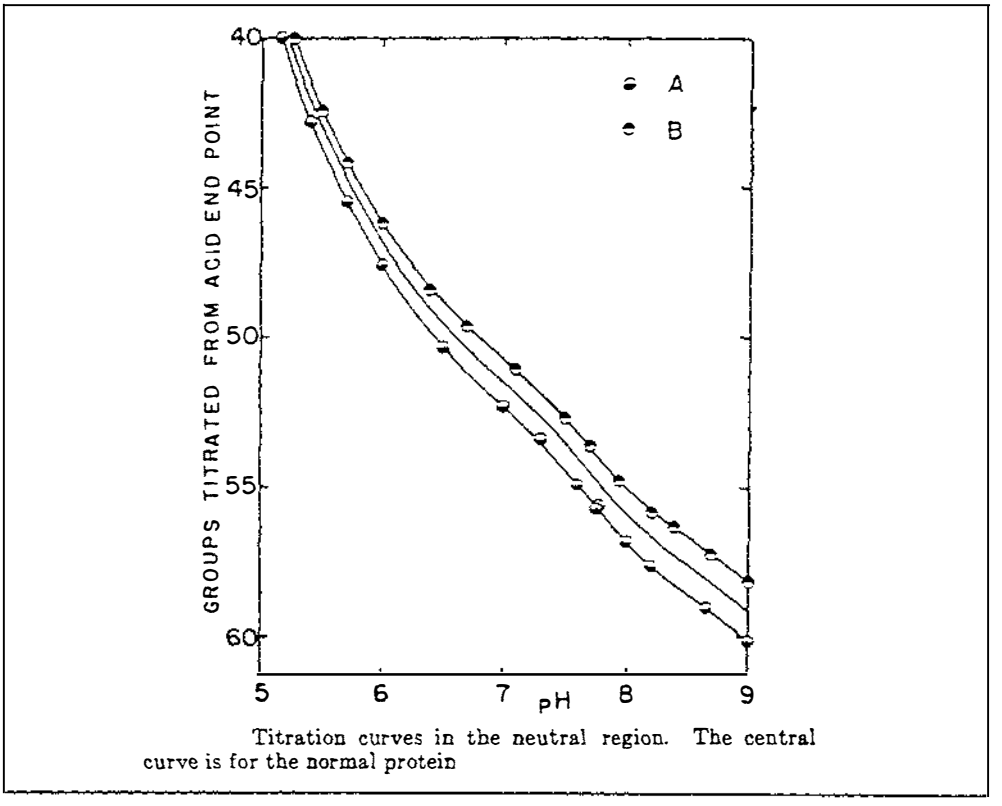


Figure 611-1: Titration curve of bovine BLGA and BLGB [Tanford and Nozaki, 1959b].

Although the titration experiments demonstrate that one carboxylic acid is buried at low pH, they do not identify which carboxylic acid is responsible. To date, no point mutation experiments have been conducted to identify the carboxylic acid and its environment. Until recently, no substantial structural basis had been found for the various pH-dependent phenomena of bovine ß-lactoglobulin, although a tentative suggestion had been made previously [Brownlow *et al.*, 1997] that Glu89 was responsible for the anomalous titration behaviour. The fully determined structures of BLGA at pH 6.2, 7.1 and 8.2 provide the structural basis for answering the question related to the Tanford transition.

6.1.2 pH-dependent pseudo-dynamics of loop EF

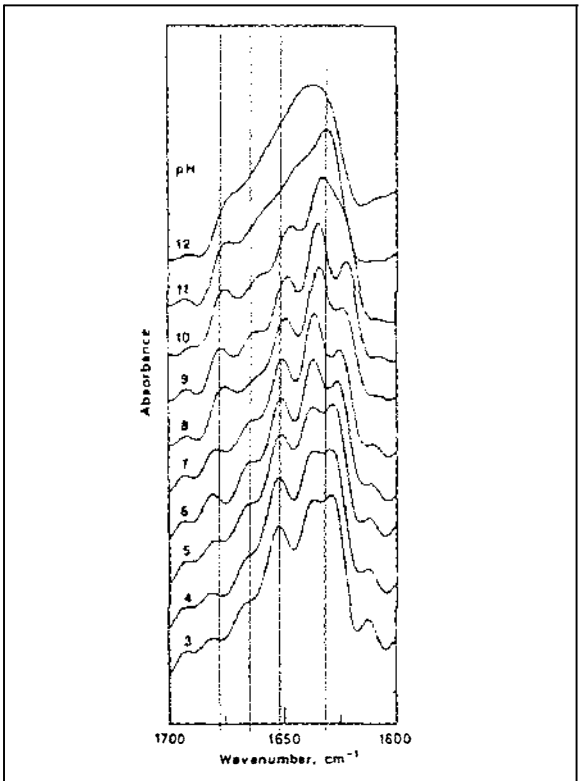
The three structures of bovine BLGA at different pH are superimposed in Figure 521-3. While the BLGA calyx experiences only minor changes, a major conformational change occurs for the loop EF. At low pH, loop EF flips towards the BLG calyx, while at pH above 7.1, the loop EF flips away from the BLG calyx. This conformational change brings a large movement to Glu89. As detailed in Table 534-1, Glu89 is buried at low pH (solvent accessible area 4 $Å^2$ at pH 6.2) and exposed at high pH. Two other protonatable residues are also buried at low pH (Glu108, solvent accessible area 9 $Å^2$; see Table 57-1. His161, solvent accessible area 10 $Å^2$; see Table 553-1), but the latter two residues do not undergo a change from buried situation to exposed situation as a function of pH.

As loop EF in both open and closed conformations is rather mobile, it is important to confirm that one or the other conformation is not an artifact. Electron density maps of loop EF at pH 6.2 and 7.1 are shown in Figure 534-1a and b. Figure 534-1a is the electron density map of BLGA at pH 7.1. It is superimposed with the bovine BLGA structure at pH 7.1 (thick line) and that at pH 6.2 (thin line). At pH 7.1, loop EF clearly adopts the open conformation. Loop EF at pH 8.2 is similar to the case at pH 7.1. This open conformation has not previously been observed in structures of BLG in lattice X. As illustrated in Figure 534-1b, a different conformation for loop EF occurs at pH 6.2, In the electron density map of BLGA at pH 6.2, thick lines represent the final structure of bovine BLGA at pH 6.2, the thin lines represent that at pH 7.1. Some

non-continuous electron density for BLGA at pH 6.2 can be found near loop EF. This density has been modelled as waters. Such density could be interpreted as a trace amount of the open form of loop EF. In any event, the loop EF primarily adopts a closed conformation at low pH.

6.1.3 Conformational change and infra-red spectroscopy

In addition to loop EF, some other regions of BLGA appear to be sensitive to pH over the pH range 6.2 ~ 8.2, for example loop GH. The pH-induced conformational changes of BLGB have been investigated in solution by infra-red spectroscopy. A frequency shift between pH ~ 7 and pH ~ 8 occurs at several typical absorption peaks, as shown in Figure 613-1. One shift, from 1649 cm^{-1} (~ pH 7) to 1647 cm^{-1} (~ pH 8), was attributed to a decrease in the α-helix content and/or an increase in the amount of random coil [Casal *et al.*, 1986] of BLGB. The X-ray structural investigation results for BLGA might provide clues to explain the infra-red spectroscopy behavior of BLGB in solution.

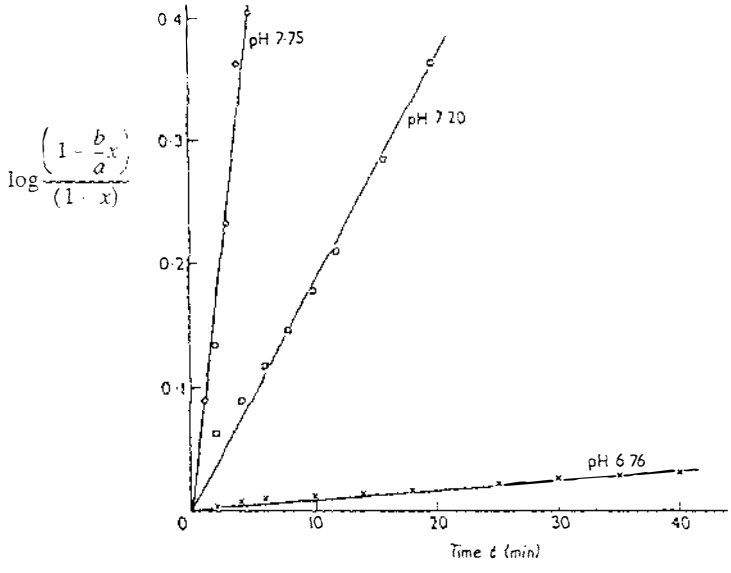


Infrared spectra in the amide I region of β-lactoglobulin (2.5 mM) in phosphate buffer at the indicated pH values after band narrowing.

Figure 613-1: Infrared spectroscopy of BLGB as function of pH. [Casa *et al.*, 1986]

6.1.4 Reactivity of Cys121 at different pH

The reactivity of Cys121 with Hg^{2+} or PCMB (*p*-chloromercuri-benzoate) in solution has been explored by several groups [Lontie and Preaux, 1966; Dunnill and Green, 1965; Boyer, 1954]. Dunnill and Green (1965) revealed that the reactivity of Cys121 with PCMB was strongly pH-dependent. The rate constant, k', for the second-order reaction increased from ~3.5 $l.mole^{-1}sec.^{-1}$ at pH 5.15 - 6.20, to 7.1 $l.mole^{-1}sec.^{-1}$ at pH 6.76, to 91 $l.mole^{-1}sec.^{-1}$ at pH 7.05 - 7.20 and to 430 $l.mole^{-1}sec.^{-1}$ at pH 7.75. Figure 614-1 reveals the time course of the reaction between PCMB and BLG, monitored by optical density. The sharp change of reactivity of Cys121 with PCMB between pH 6.75 and 7.20 was attributed by Dunnill and Green (1965) to the conformational change associated with the anomalous carboxylate (Tanford transition).



Time course of the PCMB-lactoglobulin reaction at several pH. Straight lines are the best fit to the second-order rate equation.
a, Initial concentration of β-lactoglobulin; *b*, initial concentration of PCMB; *x* fraction of PCMB reacted in time *t*.

Figure 614-1: Reactivity of Cys121 at different pH [Dunnill and Green, 1965].

The reactivity increase of Cys121 is not the result of an increase in the solvent-accessible surface area of that residue. Cys121 is a buried residue in the structures of BLGA at the pH 6.2, 7.1 and 8.2; see Table 57-1. Therefore, the reactivity increase of Cys121 is not directly related to the conformational change of BLGA at different pH. The second-

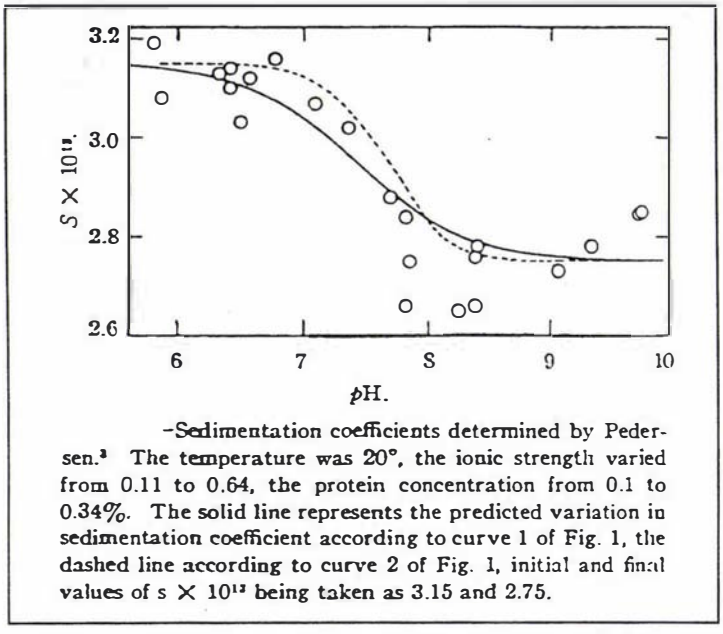
order reaction requires both PCMB and BLGA to approach each other. The electric charge of PCMB and BLGA at different pH may play a crucial role in the increase of reactivity of Cys121 in this reaction.

The SH group of Cys121 points away from the interior of calyx and into the small cavity of BLGA, see Figure 554-2, where a number of charged residues Lys101, Lys135, Arg148, Asp129, Asp137, Glu131 and Glu134 are within 8 Å of the SH group. The surface of BLGA centered at the α-H helix is shown in Figure 554-3. At pH 6.2, these residues provide a less negatively charged environment, a situation unfavorable to the approach of the positively charged PCMB reagent. Increasing the pH to 7.1 and above will change the static charge to a much more negatively charged state, favoring the access of positively charged reagents. In addition, the average B factor of atoms within 8 Å of Cys121 SG atom increases from 16.2 $Å^2$ to 30.0 $Å^2$, from pH 6.2 to pH 7.1. This implies that increasing pH will favor the binding of a positively charged reagent leading to the chemical modification of Cys121.

6.1.5 Sedimentation coefficients

The earliest observed pH-dependent property of BLG was that an increase in pH (from 6 to 8) was accompanied by a decrease in sedimentation coefficients [Pedersen, 1936]. These data have been cited by Tanford et al. in 1959; see Figure 615-1. The change in sedimentation coefficients may be attributed to an overall increase in the molecular volume and/or surface area. The accessible surface area of BLGA increases from 8,235 $Å^2$ at pH 6.2 to 8,458 $Å^2$ at pH 7.1 and to 8,666 $Å^2$ at pH 8.2, as summarised in Table 526-1. In terms of dimers, the probable association state of BLGA in solution over this pH range and at room temperature, the accessible surface area increases from 15524 $Å^2$ at pH 6.2 to 15944 $Å^2$ at pH 7.1, to 16319 $Å^2$ at pH 8.2. Although the accessible surface area is not the same as the molecular volume, the globular characteristic of bovine BLG allows a reliable deduction to be made from the values of the surfaces to a property, sedimentation coefficient, that is mostly but not entirely associated with molecular volume. Note that the pH range covered in the structure analyses spans only the region of change in sedimentation coefficients, and that the movement of the EF loop does

not alter the molecular surface area to any significant extent. In other words changes in sedimentation coefficients may be due at least in part to change in the overall solvent accessible surface area, rather than only the movement of loop EF.



-Sedimentation coefficients determined by Pedersen.[3] The temperature was 20°, the ionic strength varied from 0.11 to 0.64, the protein concentration from 0.1 to 0.34%. The solid line represents the predicted variation in sedimentation coefficient according to curve 1 of Fig. 1, the dashed line according to curve 2 of Fig. 1, initial and final values of s × 10^{13} being taken as 3.15 and 2.75.

Figure 615-1: Sedimentation coefficients of BLG *versus* pH

6.1.6 Thermal behavior of BLGA in the Tanford transition

Qi *et al.* (1996) observed differences in the thermal behavior of BLGA between pH 8.05 and 6.75. One minor peak in the thermalgram at ~ 40 °C is observable at pH 8.05, but not 6.75; another major peak (~ 80 °C) is observable at pH 6.75, but not 8.05; see Figure 616-1. This behavior may be related to infra-red spectroscopic observations by Casal *et al.* (1986) that bovine BLGB undergoes multiple temperature-dependent conformational changes in the range of -100 °C to +90 °C. One change, which occurs abruptly between 58 and 60 °C in the infrared spectrum, leads to the same infrared spectrum as that found as a result of the alkaline denaturation of BLGB [Casal *et al.*, 1986]. Although BLGA was not subjected to temperatures above 25 °C or pH above 8.2 in our research, the changes in the thermalgram [Qi *et al.*, 1996] between pH 8.05 and 6.75 at 40 °C might be correlated to the movement of loop EF of BLGA, assuming that temperature may cause the same conformational change as pH. In other

words, at slightly elevated temperature or at basic pH, loop EF flips open to provide access inside the calyx of bovine BLG.

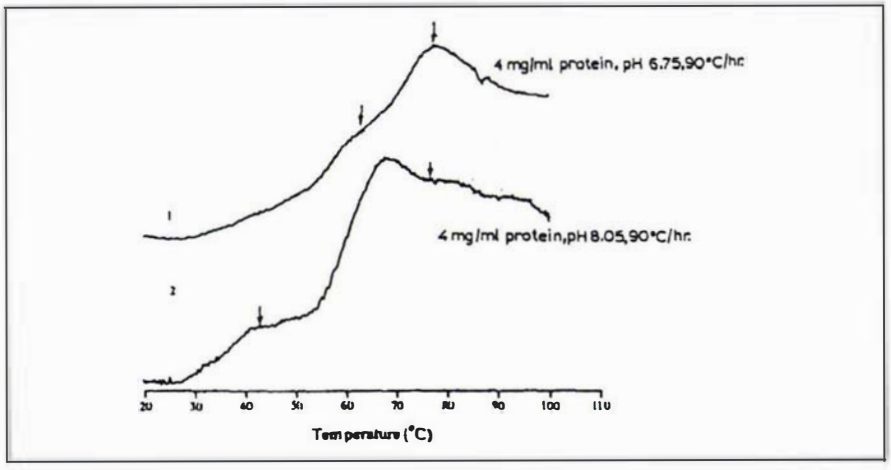


Figure 616-1: Thermalgrams of BLGA at pH 6.75 and pH 8.05. Arrows mark features that are disscussed in the text, [Qi et al., 1996].

6.1.7 Conservation of residue 89 in all BLG

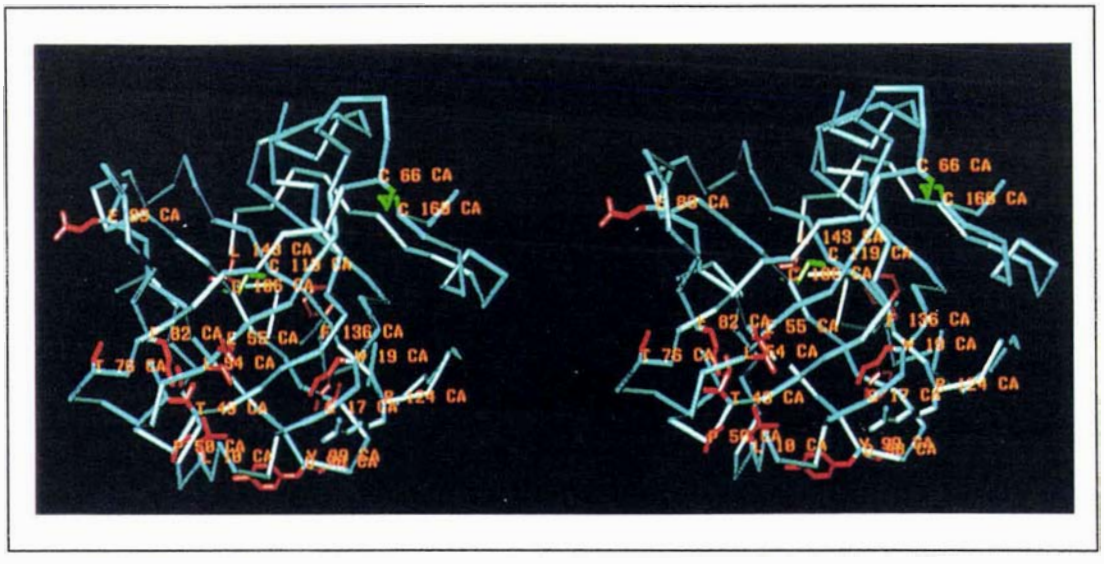


Figure 617-1: Location of conserved residues of bovine BLG

The location of all 21 conserved residues (Table 234-2) in currently sequenced BLG are shown in Figure 617-1. They are Leu10, Gly17, Trp19, Pro38, Leu39, Leu49, Pro50, Leu54, Glu55, Cys66, Thr76, Phe82, Glu89, Asp98, Tyr99, Cys106, Cys119, Arg124, Phe136, Leu143 and Cys160. Five of them (Leu10, Gly17, Trp19, Asp98, Tyr99 and Arg124) are associated with the bottom hydrogen bond network. Four cysteines form two disulfide bonds: Cys66-Cys160 and Cys106-119, which has Phe136 and Leu143 nearby. The other eight residues (Pro38, Leu39, Leu49, Pro50, Leu54, Glu55, Thr76

and Phe82) are in the vicinity of each other. The residue associated with the Tanford transition, Glu89, is isolated from the residues above.

A Tanford-like transition has also been observed in caprine (goat) BLG [Ghose *et al.*, 1968]. Flower (1996) also suggested that a loop scaffold at the entrance of several lipocalins (MUP, BBP) is well adapted to the task of ligand binding, and that the amino acid composition of the pocket and loop scaffold, as well as its overall size and conformation, determine selectivity [Flower, 1995]. Therefore, the conservation of Glu89 in a flipping loop EF might reflect the early proposed model (section 5.10.1): in the putative retinol-BLG complex atom OE1 of Glu 89 in BLGA, as a potential hydrogen-bond acceptor with the OH group of retinol, may stabilize retinol binding.

6.1.8 Functional implication of the Tanford transition

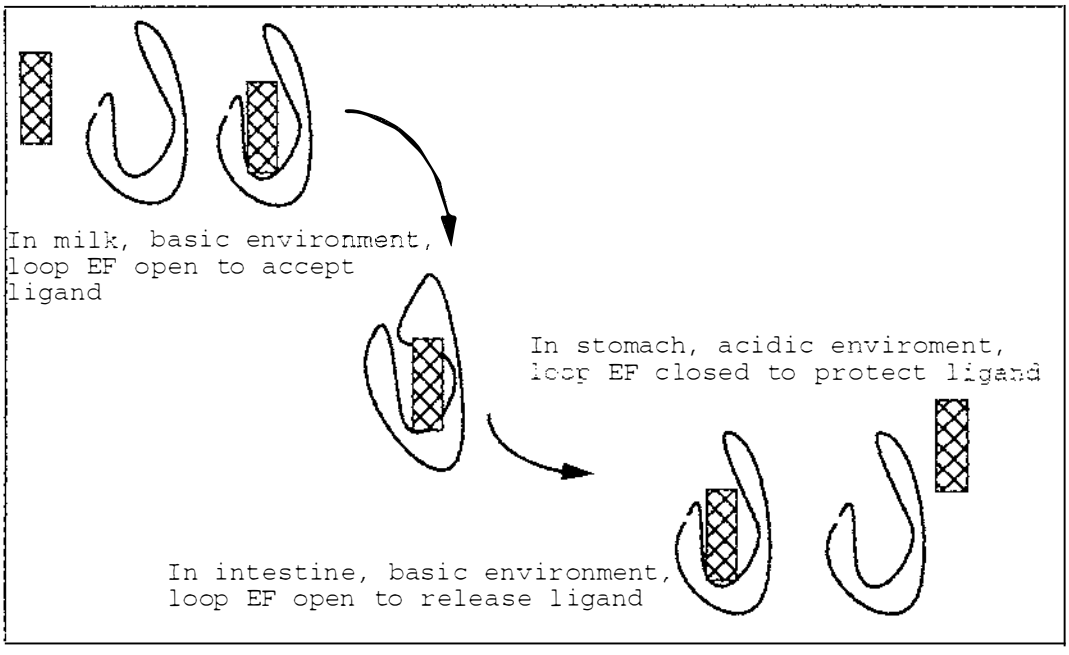


Figure 618-1: Model for molecular transport by BLG

The Tanford transition has been observed in not only bovine ß-lactoglobulins, but also goat ß-lactoglobulin [Ghose *et al.*, 1968]. The key residue, Glu89, of the Tanford transition is conserved through all currently sequenced ß-lactoglobulins, as detailed in Table 234-2. Significantly, there is little conservation of residues adjacent to Glu89. Taking into account the physiological pH conditions of the digestion canal, the Tanford transition may have physiological

implication.

It is well known that pH environment of the digestive canal changes from acidic in the stomach to basic in the small intestine, and that BLG remains intact on passage through the acidic proteolytic environment of the stomach [Yvon *et al.*, 1984]. The pH-dependent conformation of fragment EF may be related to the physiological function of ß-lactoglobulin in vivo: the closed calyx, characteristic of BLG at low pH, protects the ligand in the acidic stomach, while the opened calyx, characteristic of BLG at high pH, permits release of the ligand in the small intestine for absorbing, as detailed in Figure 618-1.

6.2 Structural basis of ligand (fatty acid) binding to bovine BLG

6.2.1 Interaction of bovine ß-lactoglobulin with fatty acids

Spector and Fletcher (1970) reported that bovine ß-lactoglobulin interacts with fatty acids, through measurements of the partition equilibrium of radiolabelled fatty acids. The strength of fatty-acid binding to bovine ß-lactoglobulin decreased in the order palmitic (C_{16}), stearic (C_{18}), oleic (C_{18}, unsaturated), and lauric acids (C_{12}) [Spector and Fletcher, 1970]. A model was proposed that contains one primary site with an association constant of the order of 10^5 to 10^6 M^{-1} and a large number of weak secondary binding sites with affinity constants of ~10^3 M^{-1} [Neuteboom *et al.*, 1992; Frapin *et al.*, 1993; Spector and Fletcher, 1970]. The association constant is higher for fatty acids with long hydrocarbon chains and is reduced when the structure of bovine ß-lactoglobulin is altered by alkylation or esterification [Frapin, 1993]. These findings indicated that interactions with charged groups may play an important role in the binding process [Frapin *et al.*, 1993; Pérez *et al.*, 1993].

In addition, the binding of palmitate analogues, in which the carboxyl group has been modified or removed, is weaker than that of the native acidic form [Spector and Fletcher, 1970]. With a hydrocarbon gel matrix, the affinity of the gel for bovine ß-lactoglobulin is lower than that with fatty-acid like gel matrix [Pérez *et al.*, 1989]. An increase in

ionic strength decreases the amount of protein adsorbed, indicating that polar interactions also are important in the binding process [Pérez *et al.*, 1989]. Thus, Pérez *et al.* (1989) and Spector and Fletcher (1970) postulated that the binding site for fatty acids in bovine ß-lactoglobulin consists of a hydrophobic pocket with a positively charged amino acid near the entrance.

The binding of fatty acids to bovine ß-lactoglobulin increases the resistance of the protein to proteolytic degradation [Puyol *et al.*, 1994a] and to thermal denaturation [Puyol *et al.*, 1994b], indicating that ligand binding may be an important factor in the stabilization of the structure of bovine ß-lactoglobulin. A similar effect is not observed when retinol is bound to bovine ß-lactoglobulin [Puyol *et al.*, 1994a; Puyol *et al.*, 1994b]. The binding of fatty acids to bovine ß-lactoglobulin causes a small change in the ultraviolet fluorescence intensity of bovine ß-lactoglobulin. This change could occur because free fatty acids interact directly with a segment of the molecule that contains tryptophan or as a result of a small conformational change in the region of tryptophan residues that are secondary to the binding of fatty acids [Spector and Fletcher, 1970]. Further studies using palmitate and bovine ß-lactoglobulin have indicated that the binding affinity increases over the pH range 6.5 to 8.5, suggesting a possible correlation with the Tanford transition [Hambling *et al.*, 1994].

ß-Lactoglobulin, when isolated from the milk of ruminants (cows, sheep, and goats) using nondenaturing techniques, also has several bound lipids, mainly triglycerides and fatty acids [Pérez *et al.*, 1989; Díaz de Villegas *et al.*, 1987]. The total amount of fatty acids bound is about 1.0 to 1.4 mol/mol of protein dimer. The predominant fatty acids are palmitic and oleic acids, which together account for approximately 75% of the fatty acids bound to ruminant ß-lactoglobulin. The bound fatty acid ratios resemble those of these fatty acids present in whey and milk [Diaz de Villegas *et al.*, 1987]. These data indicate no important differences in the binding selectivity of ß-lactoglobulin for individual fatty acids, and are inconsistent with the data obtained by the measurements of the partition equilibrium of radiolabelled fatty acids [Spector and Fletcher, 1970].

6.2.2 The BLG calyx as the ligand-binding site

The first binding study was a surface area measurement by Bull and Currie (1946b), who found that eight molecules of sodium lauryl sulfate (SLS) associate with the bovine ß-lactoglobulin dimer. Following this, a wide variety of compounds is now known to bind to bovine ß-lactoglobulin; see details in Table 252-1. However, it was unclear for a long time which regions of the protein were involved in fatty acid binding.

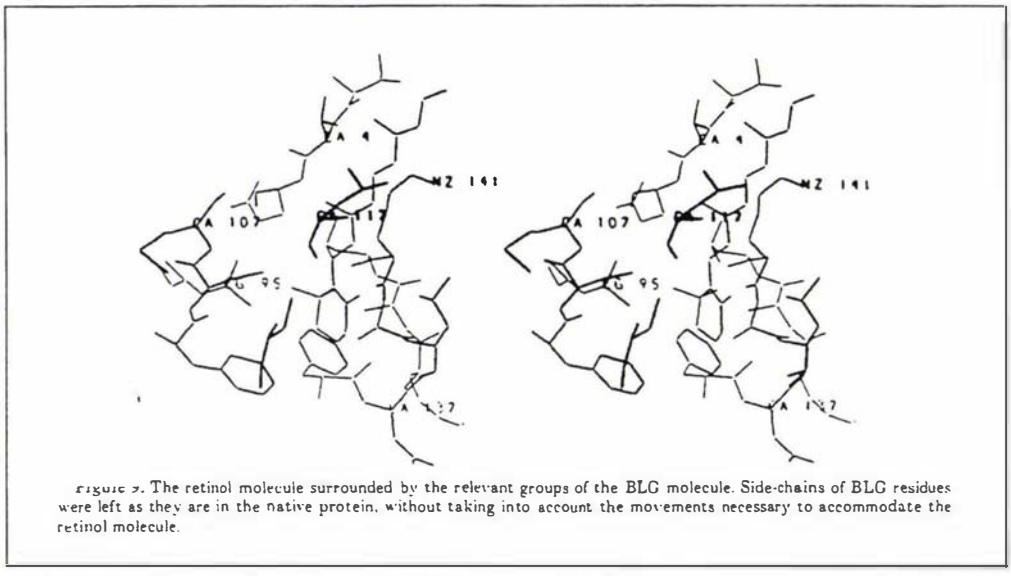


Figure 5. The retinol molecule surrounded by the relevant groups of the BLG molecule. Side-chains of BLG residues were left as they are in the native protein, without taking into account the movements necessary to accommodate the retinol molecule.

Figure 622-1: The proposed retinol-binding region of bovine BLG, [Monaco *et al.*, 1987].

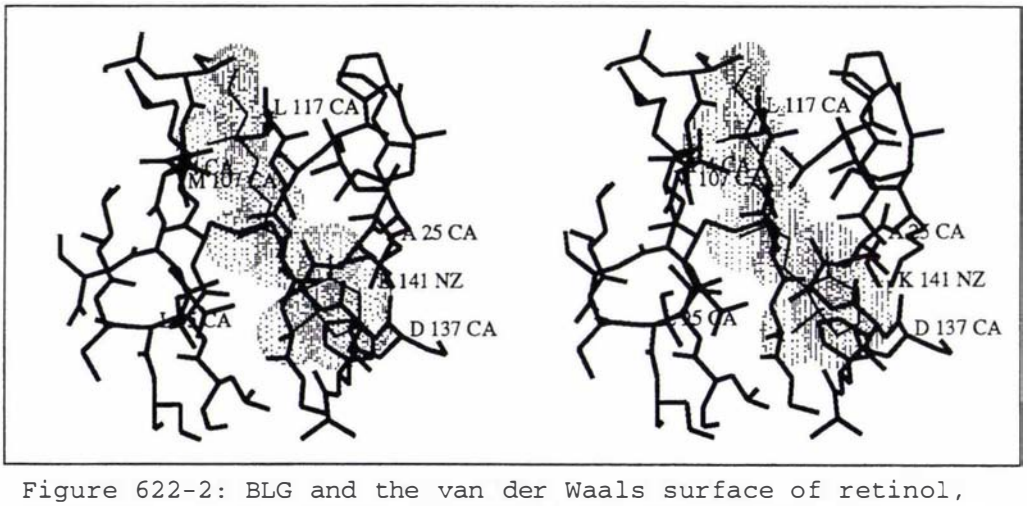


Figure 622-2: BLG and the van der Waals surface of retinol, according to the surface binding model suggested by Monaco *et al.* (1987), showing unacceptable molecular contacts around the retinol.

Monaco *et al.* (1987) proposed that retinol was bound on the surface of bovine BLG. The proposed surface hydrophobic pocket was created by the residues Lys141, Asp137, Leu95, Met107, Leu117, Thr4 and Phe136; see

Figure 622-1. Even though the authors claimed that positive density was found around this region in a difference Fourier map and could be modelled by a retinol molecule, their structure has obvious and severe problems [Brownlow *et al.*, 1997] that made their proposal rather fragile. In the redetermined bovine ß-lactoglobulin models in lattice Z [Qin *et al.*, 1998], this region is close to the minor ligand-binding site (section 5.10.2) near the vicinity of residue Cys 121. However, to enlarge the minor ligand-binding site to host a large and fairly bulky retinol molecule will introduce serious structural disturbance to this region; see Figure 622-2. As is apparent in the unacceptably tight fit of retinol between the α-H helix and N-terminal strand. This structure [Monaco *et al.*, 1987] is now known to be misthreaded [Brownlow *et al.*, 1997].

Based on the well defined bovine ß-lactoglobulin structures in lattice Z, the major ligand-binding site, inside the BLG calyx, has been modelled with retinol and palmitic acid; see section 5.10.1 and 5.10.3. Although the electron density found inside the BLGA calyx at pH 6.2 has been modelled as a string of water molecules, we can not be certain that this density is not due to an aliphatic molecule retained in the purification procedure. All these data led to the refined ligand-binding crystallization trials. The first successful crystallized BLG-ligand complex is BLGA with 12-bromododecanoic acid at pH 7.3. The binding model has been proposed in section 5.10.4.

There are two positively charged residues on the entrance of the BLG calyx: Lys60 and Lys69, and several other hydrophobic residues: Ile71, Ile84, Met107 and Leu39. The NZ atom of Lys60 is near the two O atoms of BrC12 (4.07 and 4.54 Å), while the NZ atom of Lys69 has distances to the two carboxylate O atoms of BrC12 of 5.08 and 6.52 Å. This validates the earlier postulations on the fatty acid binding site of bovine BLG [Pérez *et al.*, 1989; Spector and Fletcher, 1970]. The inside of the BLG calyx is essentially hydrophobic. In ß-sheet I, Val41, Val43, Leu46, Leu54, Ile56, Leu58 and Ile71 are involved in hydrophobic interactions with the hydrocarbon chain of BrC12; see Figure 622-3. In ß-sheet II, Ile84, Val92, Val94, Leu103, Phe105 and Met107 are involved in hydrophobic interactions with the hydrocarbon chain of the BrC12; see Figure 622-4. Within the vicinity of the bromide atom there are four CH3 groups, CG2

of Val94 (3.95 A), CD2 of Leu54 (4.07 A), CD1 of Leu46 (4.01 A) and CD2 of Leu103 (4.02 A). A cavity remains between the bromide atom and the bottom of the calyx, which should allow fatty acid molecules with longer chains than that for dodecanoic acid to bind and be well encapsulated inside the calyx.

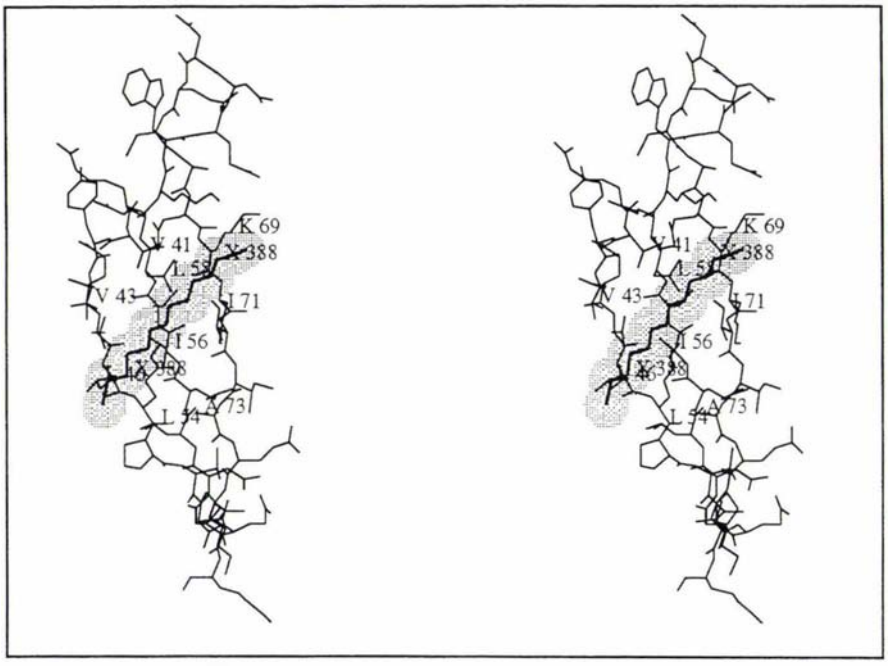


Figure 622-3: ß-sheet I and the van der Waals surface of BrC12

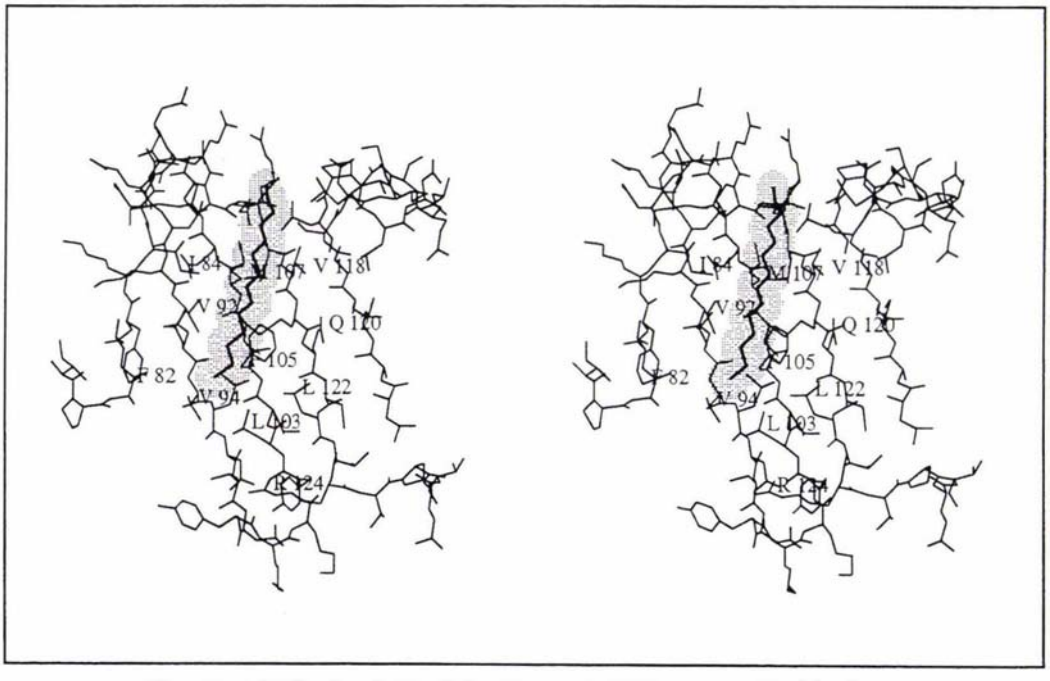


Figure 622-4: ß-sheet II and the van der Waals surface of BrC12

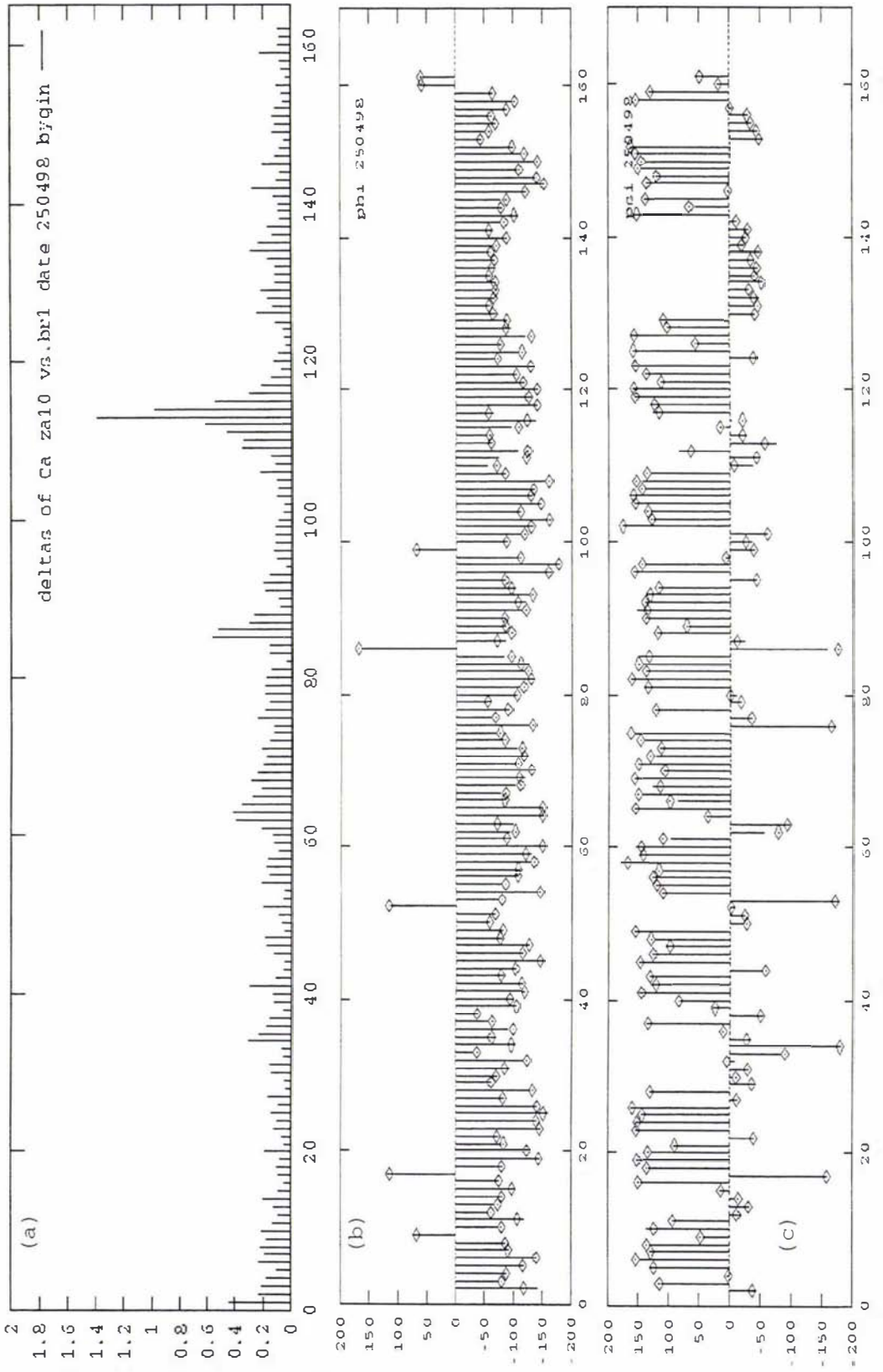


Figure 623-1: Comparison of Cα positions and φ/ψ values for BLGA and BLGA-BrC12

(a) Differences in Cα positions for the superposition of BLGA with BLGA-BrC12.
(b) φ values for BLGA (lines) and BLGA-BrC12 (boxed dots).
(C) ψ values for BLGA (lines) and BLGA-BrC12 (boxed dots).

6.2.3 Influence of BrC12 on the native structure of BLGA

Superposition of the native BLGA structure (BLGA in lattice Z at pH 7.1) onto the ligand-bound BLGA structure reveals minimal structural disturbance on ligand binding to the inside of the BLG calyx; see Figures 623-2 and 623-3: 156 Cα atoms (96.3% out of the total of 162 Cα atoms) have a displacement smaller than 0.5 Å and among these Cα atoms, the rms displacement is only 0.169 Å. Some changes in loops EF and, especially, loop GH, are identifiable. The maximum displacement in Cα positions is 1.38 Å, smaller than the changes from point mutation V118A and D64G; see details in Figure 623-1a. The structural changes of loop GH are displayed on Figure 623-3. However, as both loops EF and GH are relatively mobile the differences may not be functionally significant and may simply be a result of crystal packing and crystal growth conditions.

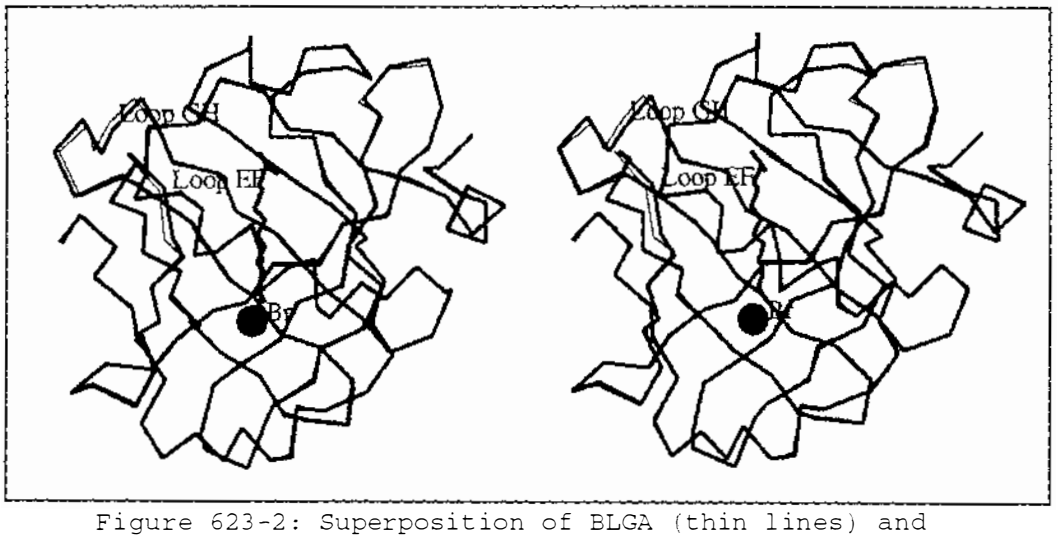


Figure 623-2: Superposition of BLGA (thin lines) and BLGA-BrC12 (thick lines)

Similar to the Cα displacements, the φ/ψ angles of peptide planes also have minimal differences between the native BLGA and ligand-bound BLGA. Figures 623-1b and 623-1c show the differences of the φ/ψ angles. The value of ligand-bound BLGA is represented by the boxed pin points, while the value of native BLGA is represented by the tip of the lines. The negligible influence of ligand (BrC12) has on the overall bovine BLG structure may not be the common issue, as suggested by the binding of 5-doxylstearic acid (5-DSA), 12-doxylstearic acid (12-DSA) and 16-doxylstearic acid (16-DSA) to bovine BLG. These fatty acid derivatives are proposed to bind on the outside of the BLG calyx [Narayan and Berliner, 1997] and this binding is sensitive to modification of Cys121

[Narayan and Berliner, 1998]. If the binding of the DSA series is located at the minor ligand binding site of BLG (see section 5.10.2), profound structural disturbance of bovine BLG is expected on binding.

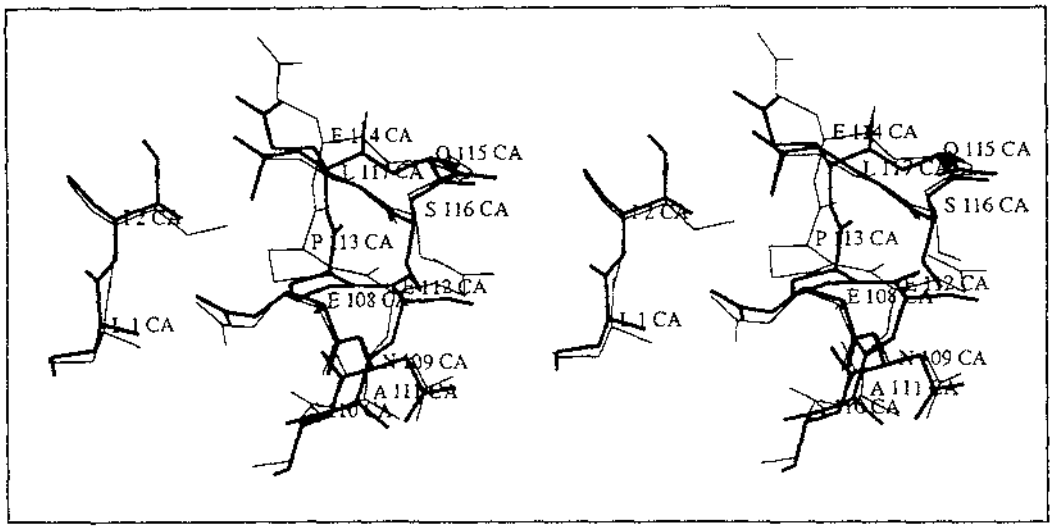


Figure 623-3: Structural change of loop GH after ligand binding. BLGA (thin lines), BLGA-BrC12 (thick lines)

6.2.4 The potential role of fatty acid binding to bovine ß-lactoglobulin

The retinol-binding ability of bovine ß-lactoglobulin has remained a focal point of research for a long time, as described in section 2.5.2. However, despite indirect experimental support for *in vivo* interaction, the presence of retinol bound to bovine ß-lactoglobulin in milk has not so far been detected [Neuteboom *et al.*, 1992]. Most of the retinol in milk is found esterified with fatty acids [Smith and Abraham, 1978]. Retinol bound to bovine ß-lactoglobulin *in vitro* can be displaced by BrC12 [Creamer, 1998]. These data suggest that the binding of fatty acids to bovine ß-lactoglobulin may play a special biological function.

The ruminant pharyngeal lipase, also called pregastric lipase (EC 3.1.1.3), is secreted from glands in the proximal-dorsal side of the tongue, the soft palate and the anterior portion of the esophagus [Moreau *et al.*, 1988]. Its activity is higher in newborn animals and is followed by a marked decrease in activity as the animal becomes older. This lipase participates in lipid digestion in the stomach, and the optimal pH ranges from 4.0 to 6.0. Pregastric lipase is very stable to low pH and pepsin degradation and is very important in newborn and young animals because

their pancreatic lipase and bile salts are low [Merchant *et al.*, 1987]. However, pregastric lipase is subjected to strong inhibition by free fatty acids, the products of the lipase. This inhibition can be prevented only when free fatty acids are removed [Bernback *et al.*, 1990].

Remarkably, the activity of ruminant pregastric lipase more than doubles in the presence of bovine ß-lactoglobulin concentrations ranging from 10 to 20 mg/ml, which correspond to protein levels found during the colostral period [Pérez *et al.*, 1992; Pervaiz and Brew, 1985]. Even though bovine serum albumin binds free fatty acids and thereby increases the activity of pregastric lipase [Bernback *et al.*, 1990; Kaminsky *et al.*, 1988], the concentration of albumin required to double the activity of lipase is between 5 ~ 10 mg/ml, which is much greater than that found either in colostrum or in the milk of ruminants [Pérez *et al.*, 1989]. Thus, the potential biological role of bovine ß-lactoglobulin could be to aid milk-fat digestion in newborn animals by promoting pregastric lipase activity [Pérez *et al.*, 1992].

6.3 The structural consequences of point mutations of BLG

6.3.1 Crystal structure of BLGA and BLGB

The unit cell parameters of crystals of BLGB in lattice Z at pH 7.1 are similar to those of BLGA in lattice Z at pH 8.2, as detailed in Table 467-2. Because the surface of BLGB at pH 7.1 is different from that of BLGA at pH 8.2, due to the loop EF adopting a different conformation in these two kinds of molecules, this similarity is irrelevant to the molecular surface; see accessible surface area of BLGs in Table 525-1. BLGB's accessible surface area bears a similarity to that for BLGA at pH 6.2.

The BLGB molecules are packed in the same manner in lattice Z as the BLGA molecules. The three types of interface of BLG in lattice Z are conserved. From the diagram showing the displacements of Cα atoms of BLGB and BLGA at pH 7.1 (Figure 631-1), conformational changes for residues involved in these three interfaces are revealed. The "dimer interface" and the "lock and key" interfaces are largely unaffected by the point mutations, while the "loop interface" differs.

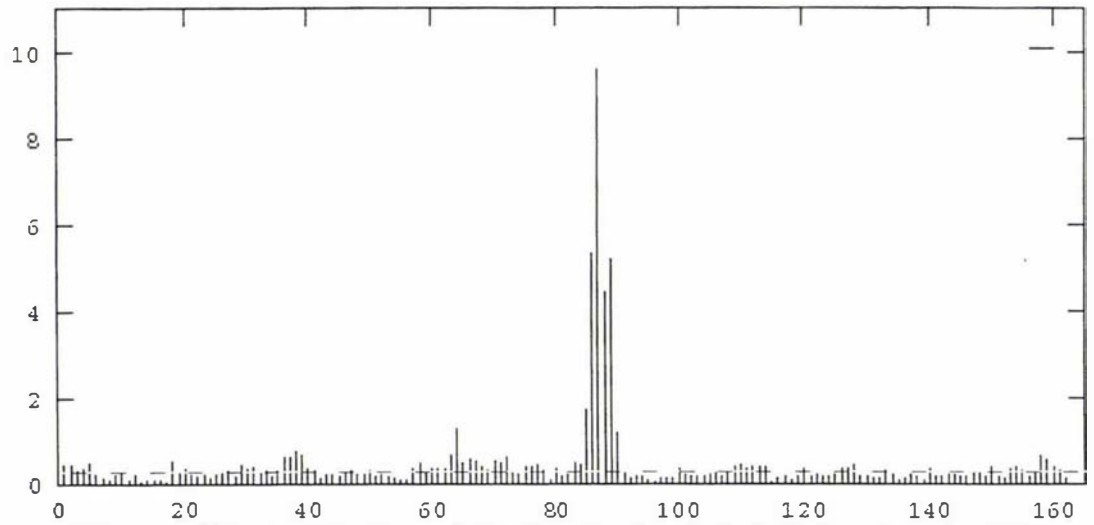


Figure 631-1: Displacements of Cα atoms for the superposition of BLGA at pH 7.1 and BLGB at pH 7.1. Twice the rms displacement is drawn in, as a reference line for significance of change.

6.3.2 Displacements and the torsional changes due to point mutations

In BLGA, Asp64 has been replaced by Gly64 in BLGB. One less negative charge at neutral pH and a less bulky side chain moiety have brought subtle changes to these two variants around position 64. The difference in Cα positions for residue 64 is 1.32 Å, the largest difference except for loop EF. Other Cα displacements have the maximum smaller than 0.7 Å, in general. Several regions of minor displacement, for example, around residues 18, 38, 64, 87, 110, 128 and 158, can be isolated from the Cα pairwise displacement plot, Figure 631-1. Around residue 64, the displacement may be the result of the point mutation D64G, while other regions of minor displacement might be random, except for residue 89 and around which are the results of conformational change of loop EF. Like bovine BLGA, the loop CD, C-terminus and disulfide bond Cys66-160 are well defined regions of bovine BLGB; see Figure 632-1. The orientation of the side chain of Glu65 in BLGB is parallel to a plane formed by the Cα atoms of residues 62 ~ 66, while that of BLGA is perpendicular to this plane. The torsion angles of the side chains of Glu62 and Asn63 are also different between these variants, but not as obviously as Glu65. The influence of the point mutation D64G on the atom coordinates is clear for the loop CD. As the loop CD is on the surface of bovine BLG molecule, the electrical charge distribution on the surface of bovine BLGA and

BLGB, especially around residue 64, should not be the same.

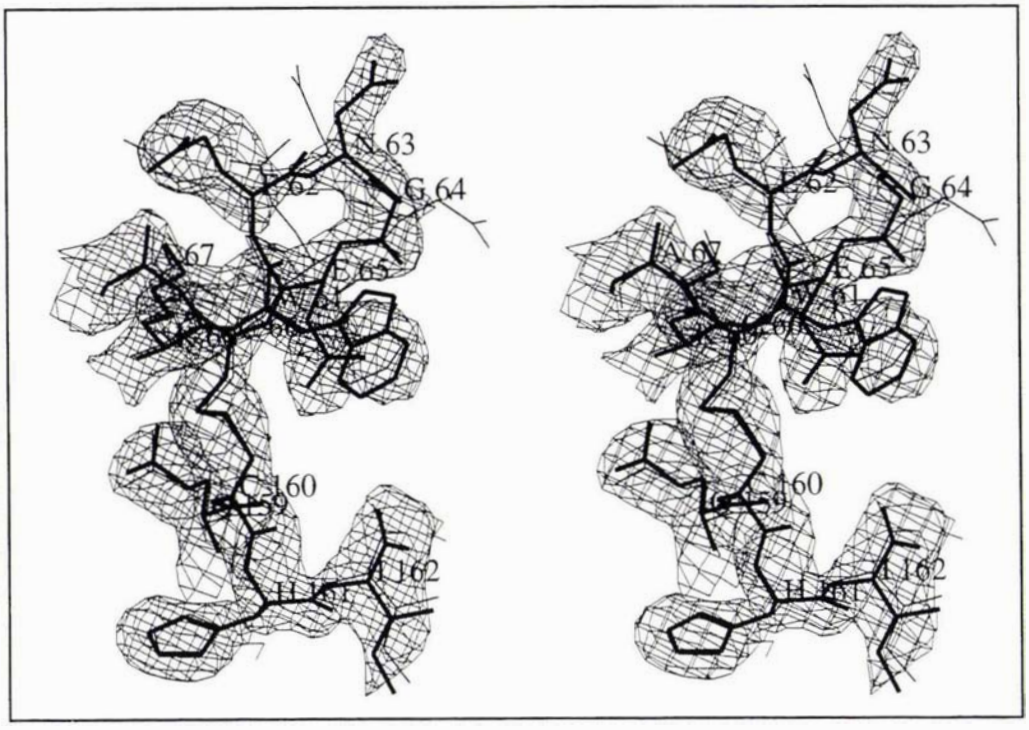


Figure 632-1 : 2Fo-Fc electron density map of loop CD of BLGB (thick lines), for reference, BLGA (pH 7.1) is drawn in thin lines.

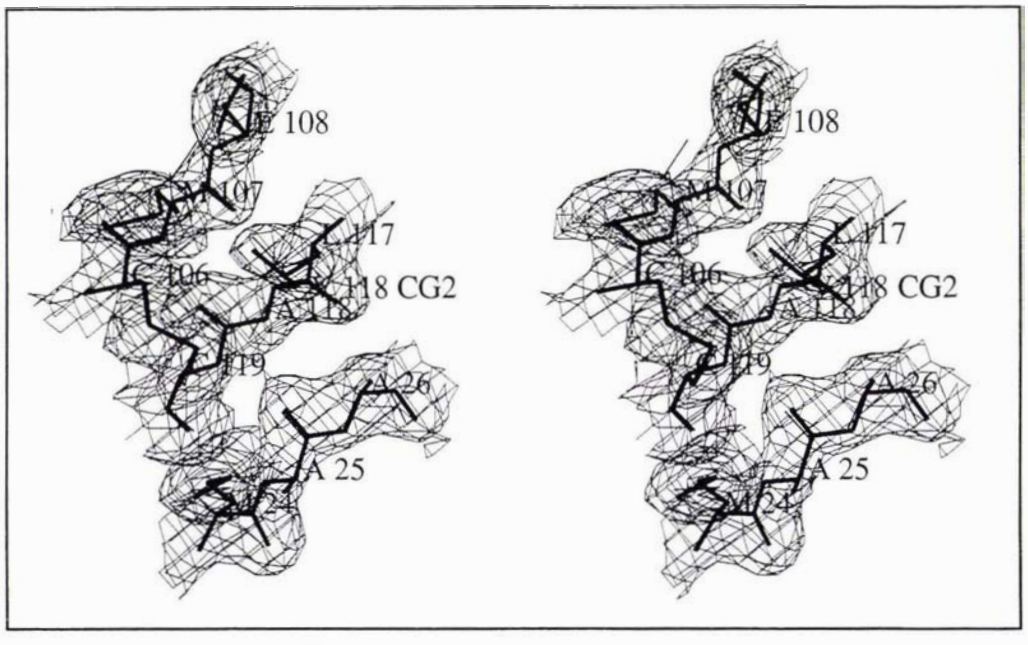


Figure 632-2: 2Fo-Fc electron density map around A118 of BLGB (thick lines), for reference, BLGA (pH 7.1) is drawn in thin lines.

Another point mutation site V118A is located on the ß-sheet II of bovine BLG molecule. This mutation site is near the disulfide bond Cys106-Cys119, and is loosely sandwiched between two Met residues (Met107 and Met24), as shown in Figure 632-2. In contrast to loop CD (average B factor

82.41 Å^2), ß-sheet II is a quite rigid region in the bovine BLG molecule with an average B factor only 42.23 Å^2. Except for the torsion angle of the side chain of Met107, no noticeable changes occur around the point mutation site V118A.

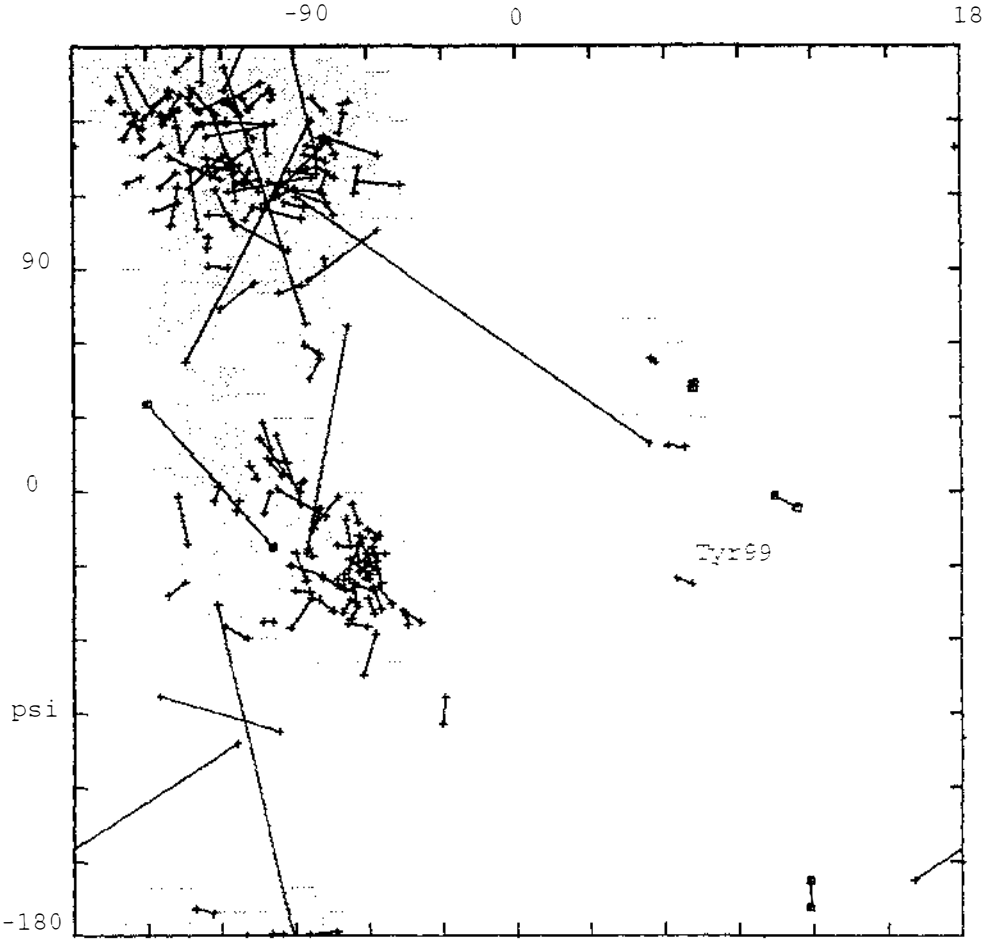


Figure 632-3: Conformational properties of BLGB compared with BLGA in lattice Z at pH 7.1.
(a). Ramachandran plot of BLGB and BLGA in lattice Z at pH 7.1, Glycines are shown as squares.

BLGB has only one residue, Tyr99, falling into a disallowed Ramachandran plot area; see Figure 632-3a. As with BLGA, Tyr99 of BLGB is at the center of a γ-turn which is stabilized by a special hydrogen bond network in the bottom region of bovine BLG. The pairwise φ/ψ angles BLGA and BLGB are shown in Figure 632-3b and 632-3c. The tips of the lines represent the values of φ/ψ angles for BLGA, while the boxed dots represent the values of φ/ψ angles for BLGB. Of all the residues of BLGA and BLGB, differences are noticeable only at a few residues and most of these have no significance whatsoever. For example, residues 87, 89 and 90 are different in their ψ angles, while residues 87 and 89 have significantly different φ angles. These differences are the result of the different

conformations of loop EF in BLGB and BLGA at pH 7.1.

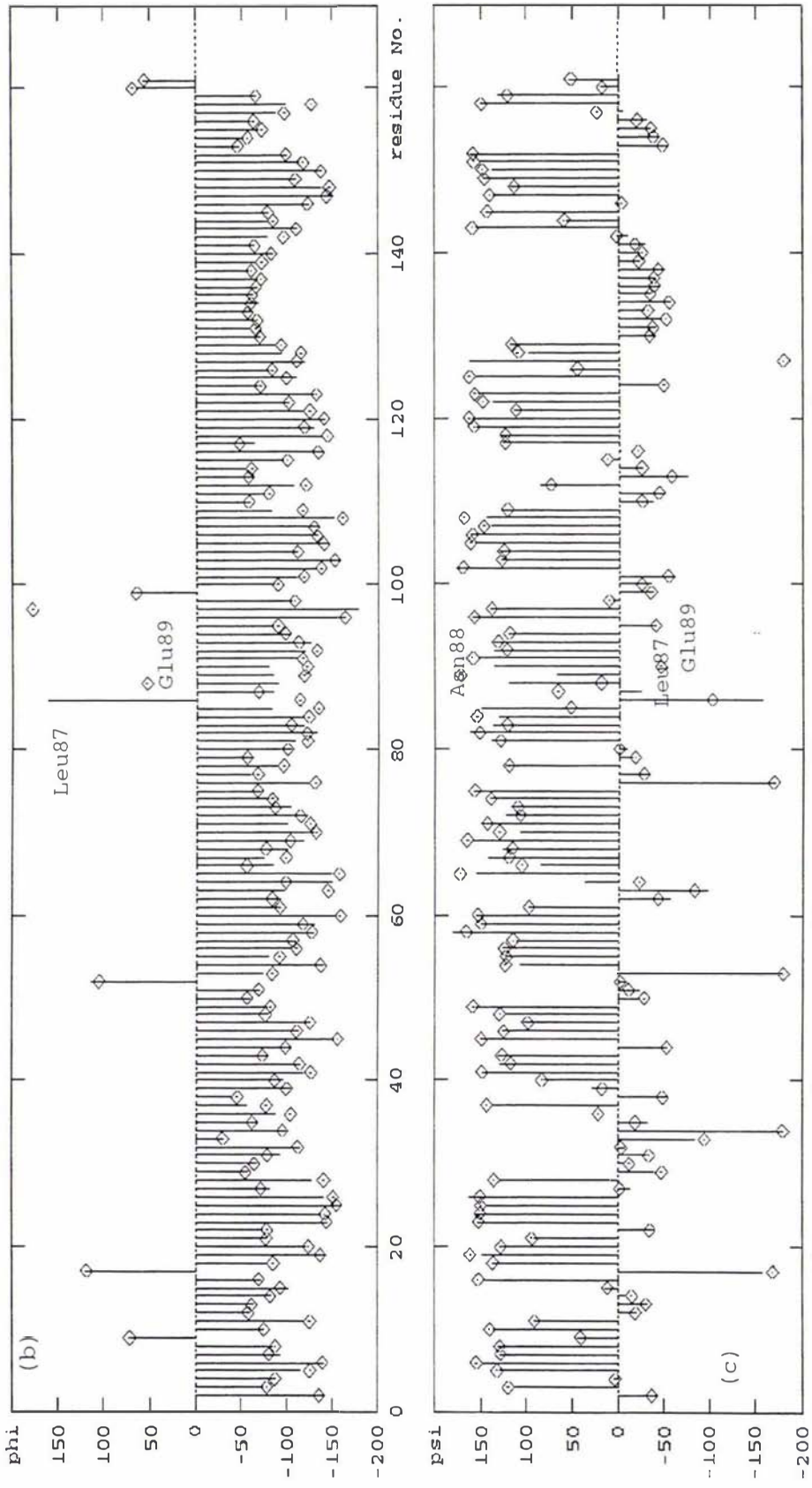


Figure 632-3: Conformational properties of BLGB compared with BLGA at pH 7.1 in lattice Z.
(b) φ values for BLGA (lines) and BLGB (boxed dots) at pH 7.1. Significant changes are marked by numbered residues.
(c) ψ values for BLGA (lines) and BLGB (boxed dots) at pH 7.1. Significant changes are marked by numbered residues.

6.3.3 BLG calyx and handle

The calyx of BLG is comprised of ß-sheet I, ß-sheet II and bottom of BLG. The ß-sheet II contains one of the point mutations, V118A. Negligible disturbance is introduced to the BLG structure by this point mutation. The average displacements of the Cα atoms of these (BLGB vs. BLGA at pH 7.1) regions are relatively insignificant. From Figure 631-1, the displacements of Cα atoms after the loop EF, are relatively less than those of Cα before loop EF. The overall B factors of ß-sheet I of both BLGA and BLGB at pH 7.1 are larger than those of ß-sheet II, as evidenced in Table 524-1.

6.3.4 Top region of BLG molecule

The top region of BLG is comprised of several loops. Loop CD contains a point mutation, D64G, which has been discussed in 6.3.2. Another loop in the top region, loop EF, differs greatly between variants A and B at the same pH 7.1; see Figure 634-1.

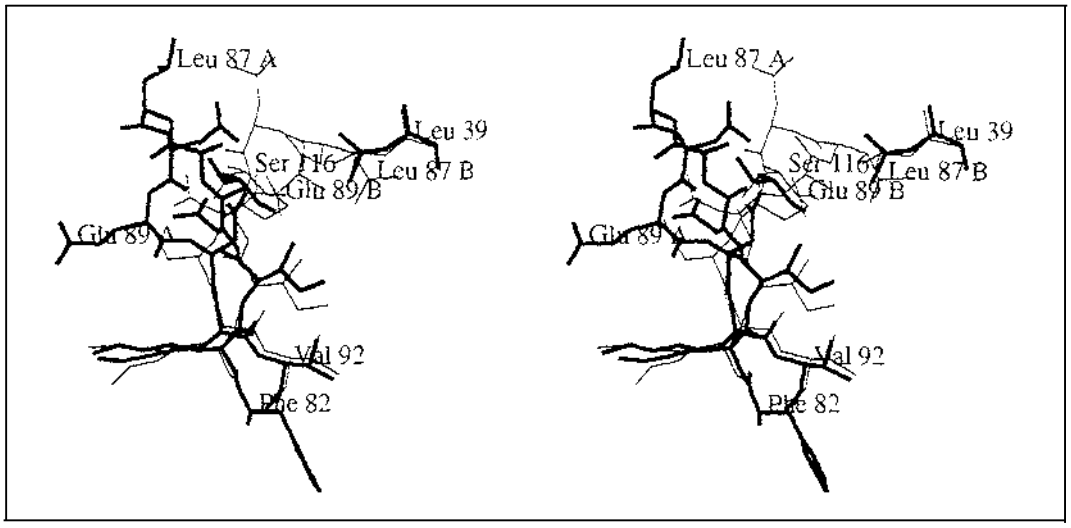
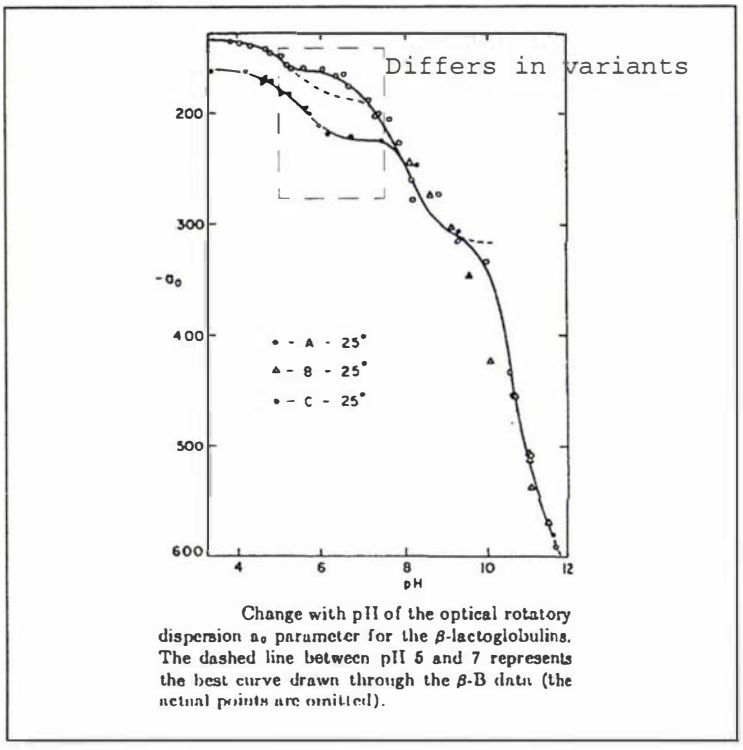


Figure 634-1 : Loop EF in BLGA (thin lines) and BLGB (thick lines) at pH 7.1, showing changes in conformation.

Loop EF in variant A at pH 7.1 is open, allowing access to the interior of the BLGA calyx. However, in BLGB at pH 7.1 loop EF is closed. This provides further evidence that loop EF is a pH-sensitive part of the BLG molecule. Even though the loop EF in BLGB at pH 7.1 and that in BLGA at pH 6.2 are both closed, the conformations are not the same; see Figure 534-3. Such data suggest that Glu89 in BLGB is released at a higher pH condition than that in BLGA. The crystallographic result for loop EF is

inconsistent with early titration researches [Tanford and Nozaki, 1959; Basch and Timasheff, 1967]. The titration experiments suggest that both variants A and B have similar pH turning points. However, the optical rotation spectra for variants A and B over pH range 6 to 7 are different; only above pH 7 do the spectra becomes indistinguishable [Timasheff *et al.*, 1966], as shown in Figure 634-2.



Change with pH of the optical rotatory dispersion a_0 parameter for the β-lactoglobulins. The dashed line between pH 5 and 7 represents the best curve drawn through the β-B data (the actual points are omitted).

Figure 634-2: Optical rotation spectra of BLGA and BLGB [Timasheff *et al.*, 1966]

The distance from the Cα atom of Gly64 to that of Leu87 in BLGB is 18.82 Å; from the Cα atom of Ala118 to that of Leu87 is 10.86 Å; the corresponding distances in BLGA at pH 7.1 are 25.63 Å, and 17.70 Å. These numbers suggest that loop EF is far from both point mutation sites. An explanation of why loop EF adopts different conformations in variants A and B involves a detailed understanding of the Tanford transiton. Even though the structural basis of the Tanford transition has been demonstrated previously [Qin *et al.*, 1998], the mechanism of this transition is uncertain at this stage. The difference of conformations of loop EF in variants A and B may be the result of the point mutations, or, more likely, it is the result of the crystallization process. The pH is close to that for the mid point of the Tanford transition for both variants A and B in solution. Because the Tanford transition is reversible, around this pH, both conformations for loop EF are likely to

exist in solutions of BLGB and BLGA. The crystallization process acts as a selector that accepts only one conformation for loop EF in the crystal lattice, and thereby gives a false impression that in solution and solid state there are intrinsically different conformations of loop EF for BLGA and BLGB.

6.4 Structural comparisons of bovine BLG in lattices X, Y and Z

6.4.1 Crystal parameters and molecular packing

The PDB file for BLG in lattice X was retrieved (code number, 1beb, deposited by Yewdall *et al.* with resolution range 15 ~ 1.8 Å and R = 0.181 (R_f = 0.243)). The protein sample in the lattice X crystals is a mixture of variants A and B. The primary structure of the PDB file reflects this fact: Gly64 and Val118. Gly64 is characteristic of BLGB and Val118 is characteristic of BLGA. The PDB file for BLGA in lattice Y has been provided by G.B.Jameson with resolution range 10 ~ 1.9 Å and R = 0.19 (R_f = 0.24). The similarities and differences in crystal parameters of these three lattices have been summarized in Table 641-1. The molecular packing of BLG is quite different in these three lattice types. Based on the molecular packing of BLGA in lattice Z, as detailed in section 5.9, the molecular packing of bovine BLG in lattices Y and X is discussed below.

Table 641-1 : Crystal parameters of BLG in lattice X, Y and Z

Lattice type code	X	Y	Z
BLG variant	AB	A	A
pH	4.6	7.5	7.1
Resolution range Å	15 ~ 1.8	10 ~ 1.9	15 ~ 2.04
crystal space	P1	$C222_1$	$P3_221$
unit cell a b c Å	37.5 42.5 51.9	85.44 82.03 44.79	83.860 83.860 110.408
unit cell α β γ°	103.4 97.5 111.7	90.00 90.00 90.00	90.0 90.0 120.00
Volume	71697	97114	75442
molecules asymmetric unit	1	1	1
asymmetric unit/unit cell	2	8	6
V_M (Matthews, 1968)	2.09	2.11	2.01 Å³ Dalton

The space group of lattice X of BLG is P1; the unit cell contains only one asymmetric unit with two molecules inside. The packing is illustrated in Figure 641-1. These two BLG molecules are related by a local non-crystallographic two-fold symmetry axis, and form a dimer interface similar but not identical to that for BLG in lattice Z. The V-shaped gorge of the dimer interface of BLG in lattice X is more open than that in lattice Z; see Figure 641-2. The "lock and key" interface and the loop interface do not exist in lattice X. The "key", Lys8 is moved away from the surface hole, as shown in Figure 525-5 for BLGA in lattice Z, and is close to loop CD from one molecule and loop GH from another molecule.

The BLG dimers in lattice Z are chained together in a linear array through the "lock and key" interfaces; see section 5.9. The BLG dimers in lattice X are also packed in a linear array, but in different way. The bottom region of chain A in the dimer faces the bottom region of chain B of another dimer, and *vice versa*. Through such "bottom:bottom" interfaces, the BLG dimers in lattice X form a linear array which is parallel to the c axis of the unit cell of lattice X; see Figure 641-1.

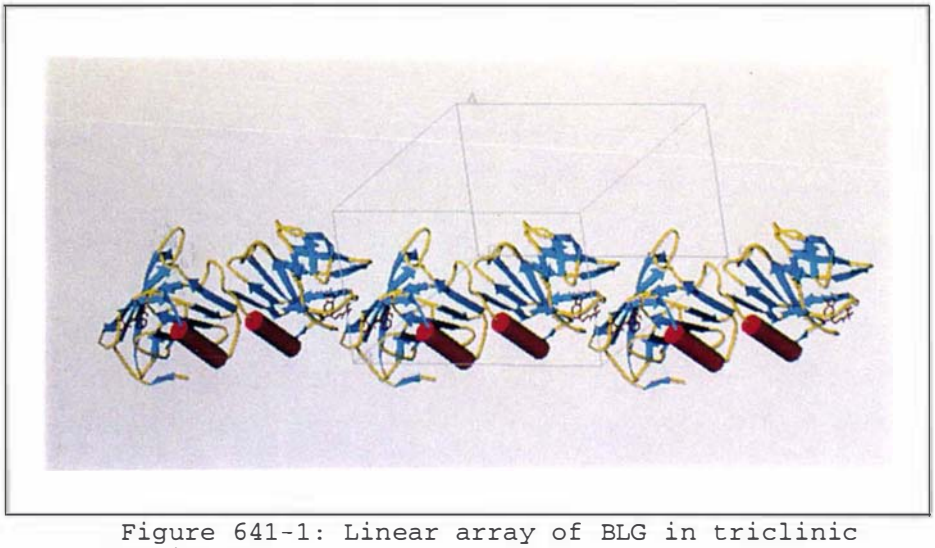


Figure 641-1: Linear array of BLG in triclinic lattice X

The space group of lattice Y of bovine BLG is $C222_1$; the unit cell contains eight asymmetric units each with one molecule inside. The dimer interface of BLGA in lattice Y is also conserved. Relative to BLGA in lattice Z, there is an even larger rotation of the second monomer of the dimer for lattice Y compared to lattice X, as illustrated in Figure 641-2. In contrast to the dimer interfaces in lattices X and Z, where the V-

shaped gorges are free of other molecules (except for water molecules), the wide gorge of the dimer interface of lattice Y contains by two Lys77 side chains from neighboring molecules, which are related by two-fold symmetry. A similar loop interface is observable in lattice Y, with loop EF approaching the calyx of its symmetry-related BLG molecule.

The dimer in lattice Y forms chains in a manner similar to that in lattice X. Therefore, the "bottom:bottom" interface is also observed in lattice Y. There are two "bottom:bottom" dimer arrays related by a two-fold rotation axis found in lattice Y. These two symmetrically related dimer arrays further generate their two-fold symmetry mates in lattice Y.

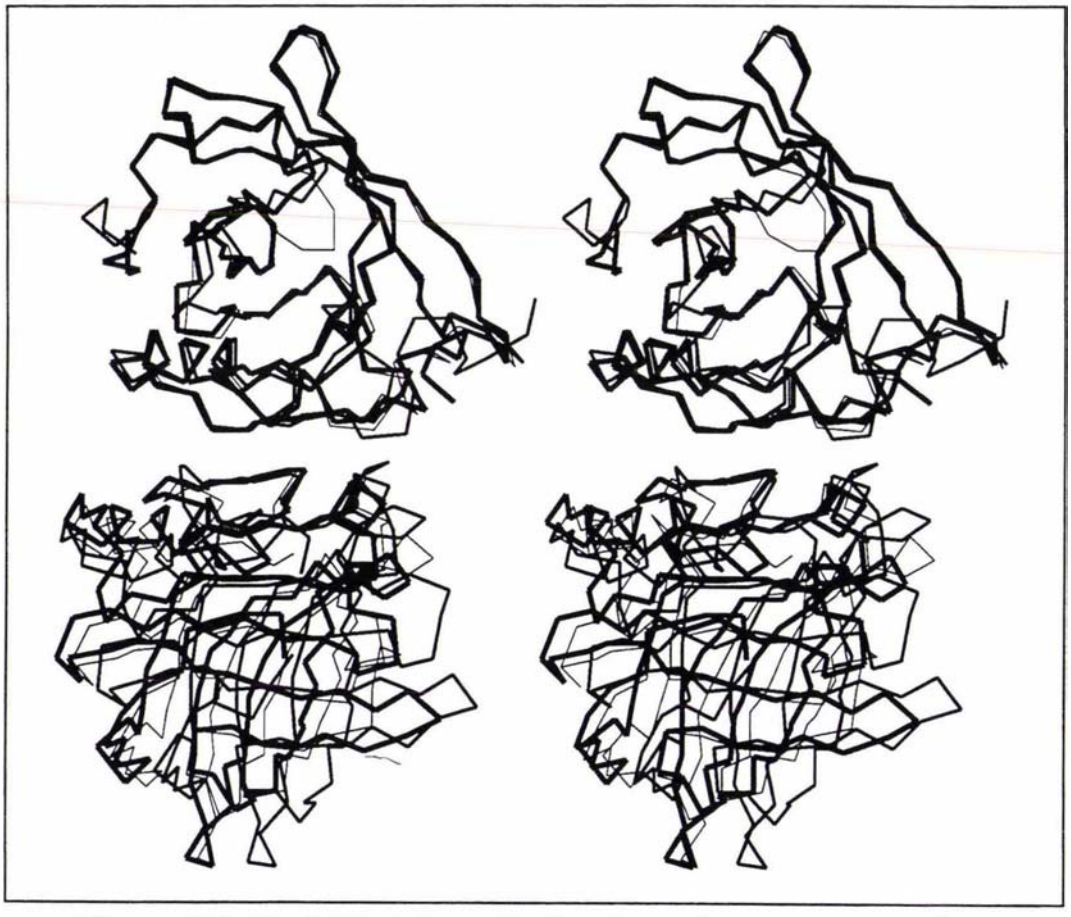


Figure 641-2: Differences in the dimer interface for BLG in lattices X, Y and Z

The BLG dimers in lattices X, Y and Z are superimposed to minimise the Cα displacements of one of the monomers. As a result, the Cα displacements of the other monomer in the BLG dimer will be amplified. Lattices X (thin lines), Y (medium lines) and Z (thick lines).

6.4.2 The BLG calyx

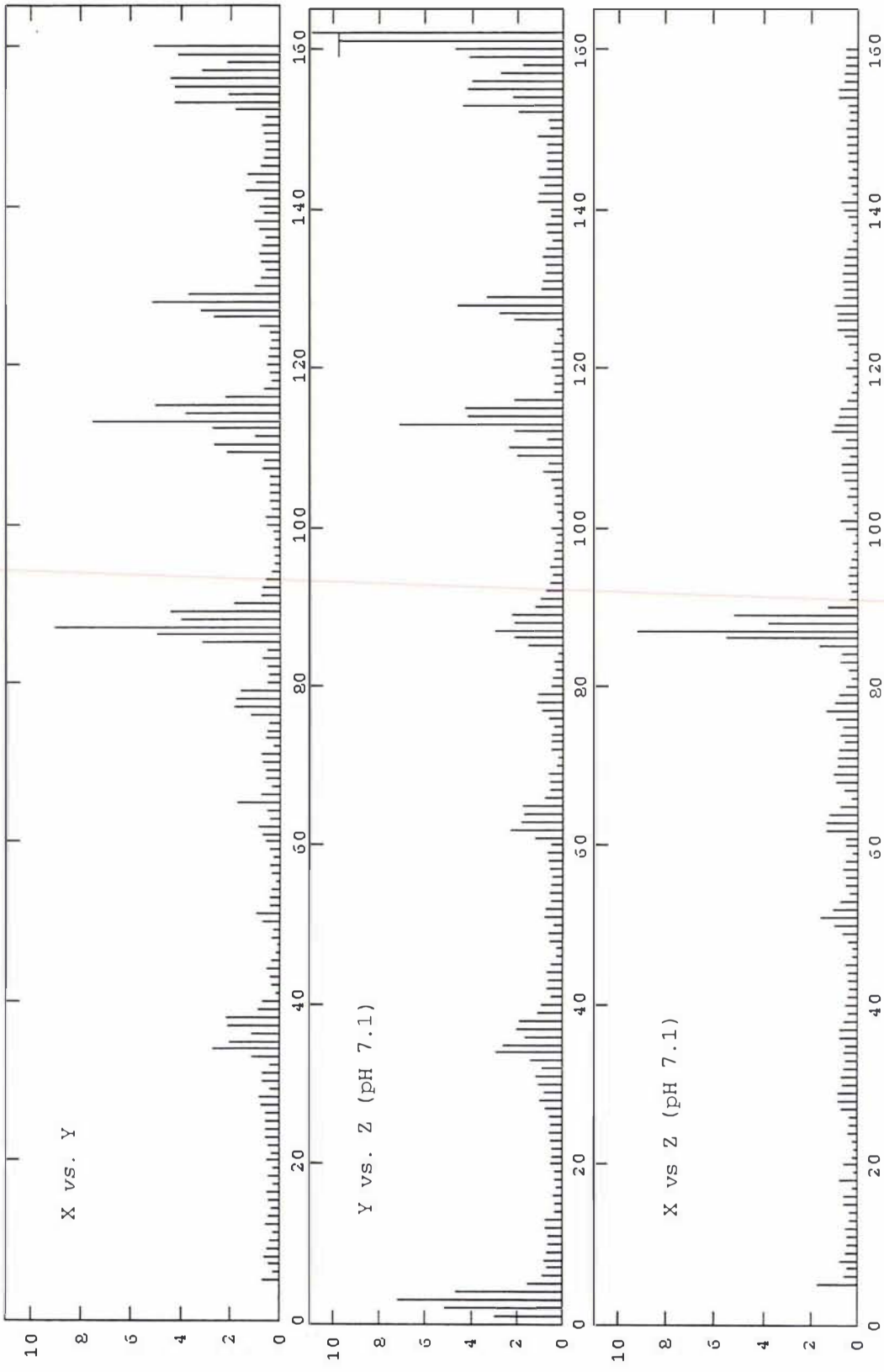


Figure 642-1: Displacement plots of Cα atoms of bovine BLG molecules in different lattices.

The BLG calyx is comprised of ß-sheet I, ß-sheet II, and bottom region.

Trivial differences can be identified in these three regions, as detailed in Figure 642-1, the plot of Cα displacements of bovine BLG in different lattices. This suggests that the BLG calyx is conserved through different lattice types. However, the calices in different lattices may vary in terms of the entrance accessibilities. In lattice Z, where the entrances to the calyx face to each other, it is not as easy for the ligand to diffuse into the calyx as in lattice Y. The entrance to the BLG calyx in lattice X is blocked, due to the closed conformation of loop EF. The difficulty in the ligand binding experiment for crystalline BLG may be caused by the impossibility for ligands to diffuse into the binding site.

6.4.3 Handle region

While the handle regions of BLG in lattice X and Z are conserved, the handle region of the BLG molecules in lattice Y is twisted compared to those of lattice X and Z. The α-H helix of BLG in lattice Y is shifted down towards the bottom region by 0.75 Å ~ 1.05 Å, as shown in Figure 643-1. This causes the region from residue 125 to 130 to adopt a quite different conformation compared to lattices X and Z; as is also apparent in the plot of Cα displacements (Figure 642-1).

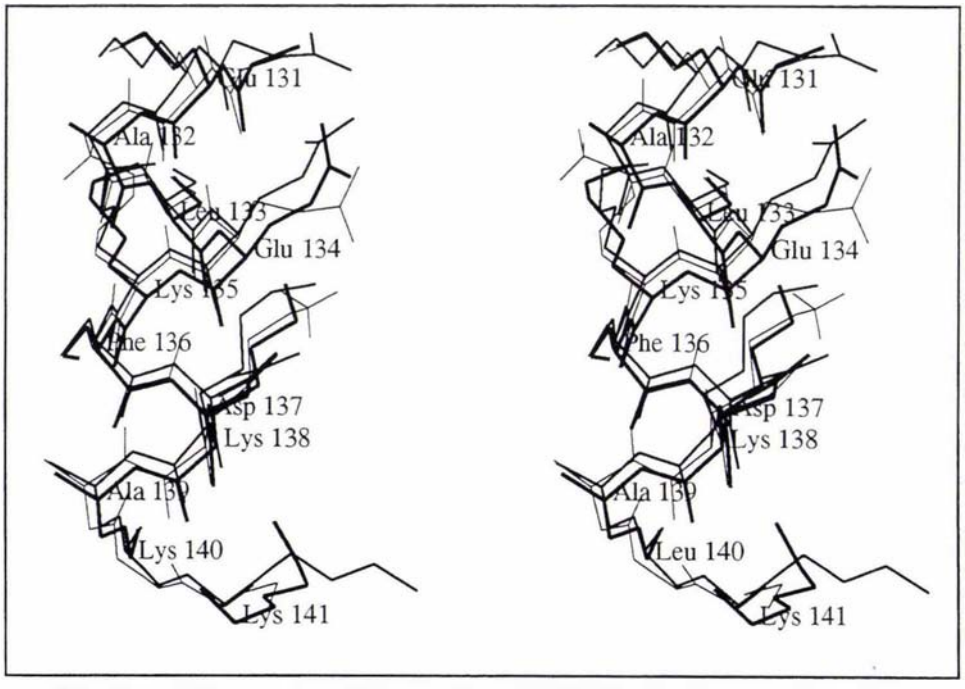


Figure 643-1 : Position of α-H helix in three lattices X (thin lines), Y (medium lines) and Z (thick lines), relative to superpositions of the whole molecule.

Strand ß-I and C-terminus of BLG in lattice Y are also different from those of bovine BLG in lattices X and Z. The α-3 helix is poorly defined in lattice Y, but exists in both lattices X and Z. Another part of handle region, the N-terminus, which is omitted in lattice X, is quite different in lattices Y and Z. This can be evidenced by the plot of Cα displacements of Figure 642-1.

6.4.4 Top region

The top region of BLG contains four loops: loop AB (residues 28 ~ 39), CD (residues 60 ~ 66), EF (residues 86 ~ 90) and GH (residues 109 ~ 116). The displacements of Cα atoms in this region are shown in Figure 642-1. While the change of conformations of loops AB, CD and GH is observable, the displacements of the Cα atoms of loop EF are distinctive, especially comparing lattice Y with lattice X. The loop EF of BLG in lattice Y adopts an open conformation, while that of BLG in lattice X is closed, as shown in Figure 644-1.

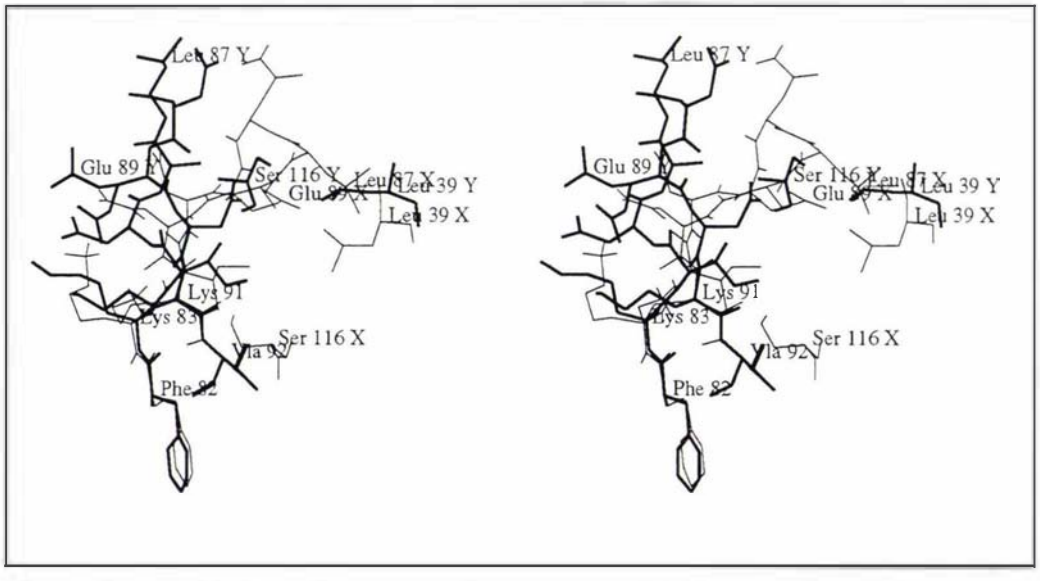


Figure 644-1: Loop EF of BLG in lattice X (thin) and Y (thick), superposition based on the whole molecule. (Note the side chain of Glu89 in lattice Y is disordered)

Taking into account the pH conditions of crystallization of lattice X (acid pH side of the Tanford transition) and lattice Y (slightly above neutral pH), the conformation of loop EF is consistent with the results of the structures of BLGA in lattice Z at pH 7.1 and pH 6.2. This gives further support to the hypothesis that the conformation changes in loop EF are the pH-dependent rather than lattice type-dependent.

Based on the facts that the Tanford transition of BLG requires a buried COOH group at low pH and that the Glu89 is the only buried residue with a COOH moiety observable in the lattice X structure, Brownlow *et al.* (1997) made a tentative suggestion about the role of Glu89 in the Tanford transition. However, with the N-terminus undefined in lattice X, leading to the apparent exposure of Glu108, which in lattice Z is buried by the N-terminus, the evidence supporting the assigment of Glu89 to the Tanford transition was equivocal, until the structures of BLGA in lattice Z [Qin *et al.*, 1998].

6.4.5 Disulfide bonds

The disulfide bond Cys106-Cys119 is observable in the three lattice types. This disulfide bond is located in ß-sheet II, in a region where the polypeptide chain of bovine BLG is relatively very rigid. Another disulfide bond, Cys66-Cys160, links flexible loop CD to the C-terminus. While this disufide bridge and neighboring residues are clearly defined in lattice Z, they are less defined in lattices X [Brownlow *et al.*, 1997] and Y. Figure 642-1 shows clearly that BLGA in lattice Y follows a somewhat different path than that in lattice Z (or X), beginning at residue 152.

6.5 Structural relationship of BLG with other lipocalins

6.5.1 Source of tertiary structures and the crystal parameters

The PDB file for retinol-binding protein, RBP, retrieved from the PDB was deposited by Zanotti *et al.* Refinement used data in the range 9 ~ 1.9 Å, leading to R = 0.184 (R_f not monitored). The PDB file for bilin-binding protein, BBP, retrieved from the PDB was deposited by Huber *et al.* Refinement used data in the range 25.0 ~ 2.0 Å, leading to R = 0.2 (R_f not monitored). The pdb file for odorant-binding protein, OBP, retrieved from PDB bank [Bianchet *et al.*, 1996]. Refinement used data in the range 6.0 ~ 2.2 Å, leading to R = 0.19 (R_f NULL). Bovine ß-lactoglobulin, which is able to produce several crystal forms, is represented by BLGA in lattice Z at pH 7.1; see details in Table 651-1.

Table 651-1 : Crystal parameters of selected lipocalins

Protein	BLG	RBP	OBP	BBP
PDB Code	N.A.	1fem	1pbo	1bbp
R(R_f)	0.233(0.279)	0.184(N.A.)	0.19(NULL)	0.20(N.A.)
Resolution (Å)	15~2.24	9~1.9	6.0~2.2	25.0~2.0
Crystal system	Trigonal	Orthorhombic	Monoclinic	Orthorhombic
Space group	$P3_221$	$P2_12_12_1$	$P12_11$	$P2_12_12$
a,b,c (Å)	54.0 54.0 112.4	46.2 48.8 75.8	55.9 65.5 42.7	132.1 121.9 63.8
α,β,γ (°)	90.0 90.0 120.0	90.0 90.0 90.0	90.0 98.8 90.0	90.0 90.0 90.0
molec./asym.	1	1	2	4
asym./cell	8	4	2	4

6.5.2 Primary structure analysis

Table 652-1: Structural alignment of four lipocalins with respect to secondary structure

```
                 N-terminus   13   [    ß-A         ]            39
-------------------------------------------------------------------------
    BLG    .......LIV TQTMKGLDIQ K. [ VAGTWYSLAM AAS ] DISLLDAQSA PL
    RBP    ..ERDCRVSS FRVKENFDKA RF [ .AGTWYAMAK KD. ] ......PEG. ..
    BBP    NVYHDGACPE VKPVDNFDWS N. [ YHGKWWEVAK YP. ] ..NSVEKYGK ..
    OBP    .......AQE EEAEQNLSE. .. [ LSGPWRTVYI GST ] NPEKIQ.ENG PF
Consensus  .........e .....NfD.. .. [ .aGtWy.vAk ... ] .......eg. ..
          [      ß-B          ] 51[      ß-C         ]   65[    ß-D            ]
-------------------------------------------------------------------------
    BLG   [ ....RVYVEE LKPTP ] E [ .GDLEILLQK WE... ] NGE [ ....CAQKKI IAEKTK ]
    RBP   [ LFLQDNIVAE FSVDE ] N [ .GHMSATAKG RVRLL ] N.. [ NWDVCADMVG TFTDTE ]
    BBP   [ .....CGWAE YTPE. ] G [ KSVKVSNYHV IH... ] .G. [ .....KEYFI EGTAY. ]
    OBP   [ ....RTYFRE LVFDD ] E [ KGTVDFYFSV KRD.. ] .G. [ KWK...NVHV KATKQD ]
Consensus [ ....r.yvaE l.pd. ]   [ .G.......v ..... ] .G. [ .....a...i .aTkt. ]
          78     [    ß-E        ]   88[      ß-F         ]100[  ß-G          ]
-------------------------------------------------------------------------
    BLG   IP.... [ AVFKID.... .. ] ALN [ ...ENKVLV. LDTDY ] K [ KYLL..FCME NS ]
    RBP   DP.... [ AKFKMKYWGV .. ] ASF [ LQKGNDDHWI IDTDY ] E [ TFAVQYSCR. .. ]
    BBP   .PVGDS [ .KIGKIYHKL TY ] G   [ VTKENVFNV. LSTDN ] K [ NYIIGYYCKY DE ]
    OBP   D..... [ GTYVADY... .. ] EG  [ ..GQNVFKI. VSLSR ] T [ .HLVAHNINV DK ]
Consensus .P.... [ akfk.dY... .. ]     [ ..geNvf.v. ldTDy ]   [ .ylv.y.C.. d. ]
          111     [          ß-H        ]   128 [       α-H          ]
-------------------------------------------------------------------------
    BLG   AEPEQSL [  VC.....QCL V.RT... ] PEV   [ DDEALEKFDK ALK.... ]
    RBP   LLNLDGT [  .CADSYSFVF A.R.... ] DPSGF [ SPEV.QKIVR QRQEELC ]
    BBP   DK      [  KGHQDFVWVL S.R.... ] SKVLT [ .GEA..KTAV ENYLIG. ]
    OBP   H       [  GQTTELTG.L FVKLNVE ]       [ .DEDLEKFWK LTEDKGI ]
Consensus         [  ........vL ..R.... ]       [ ..EaleKf.k ....... ]
                  [    ß-I     ] 153 C-terminus        Total residues
-------------------------------------------------------------------------
   BLG    ....ALP [ MHIRLSFN.. ] PTQLEEQCHI            162
   RBP    ...LAR. [ .QYRLIPHN. ] GY.....CDG KSERNIL    183
   BBP    SPVVDSQ [ ...KLVYSDF ] ...SEAAC.. KVN        173
   OBP    ....DKK [ ..NVVNFL.. ] ....ENEDHP HPE        159
Consensus ....... [ ...rL.f... ] ....E..C.. K.....    20/72
-------------------------------------------------------------------------
```

Comparison of the primary structures for bovine lipocalins BLG, RBP and OBP was shown in section 2.3.3. Based on knowledge of the secondary structure, these lipocalins plus BBP are aligned further in Table 652-1. If a residue is conserved in three out of the four sequences, the consensus residue is labelled in a capital letter. Bovine BLGA involves 16 such residues: Asp11, Gly17, Trp19, Glu45, Gly52, Gly64, Pro79, Asn90, Thr97, Asp98, Cys106, Arg124, Glu131, Lys135, Leu149 and Glu158. The importance of five residues, Gly17, Trp19, Thr97, Asp98 and Arg124, is interpretable. All of these residues are part of the bottom region complex.

6.5.3 The lipocalin fold

RBP, BBP and BLG have the typical lipocalin fold, as shown in Figure 231-1. A central ß-barrel is comprised of 8 anti-parallel ß-strands. Outside the central ß-barrel is a three-turn α-helix, which is right on the top of the ß-sheet II in terms of the bovine BLG structure. The tertiary structure of OBP has some exceptions to the typical lipocalin fold. While the basic characteristics of the central ß-barrel are still conserved in the OBP molecule, the three-turn α-helix (almost four-turn) extends away from the OBP molecule, to assist in dimer formation as shown in Figure 653-1.

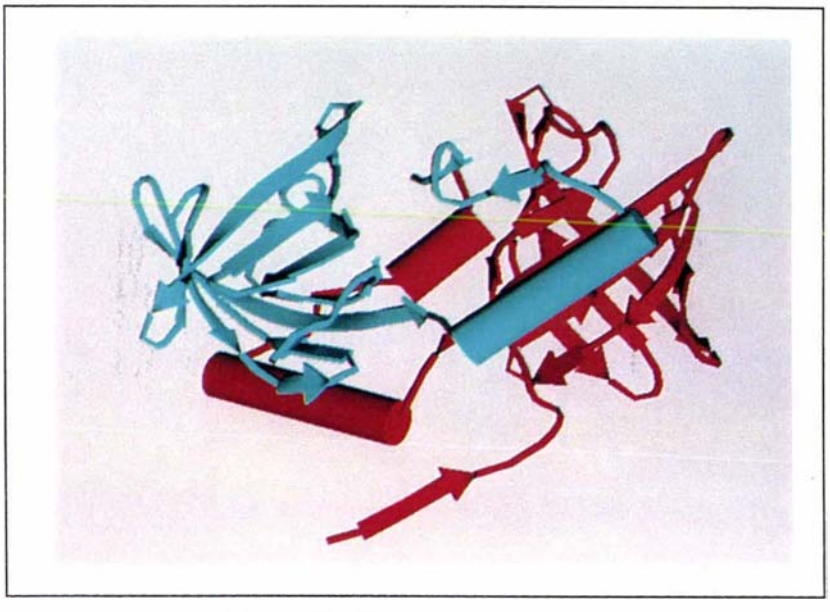


Figure 653-1: OBP dimer, showing non-conventional lipocalin fold

6.5.4 Ligand-binding sites

The ligands of BBP (bilin), Figure 654-1, RBP (retinol), Figure 654-2, OBP (pyrazine), Figure 654-3, and BLG (BrC12), Figure 654-4, are all bound inside the central ß-barrel. All the calyces are aligned in similar orientation. In BBP, the side chain of Trp133 is parallel and close to

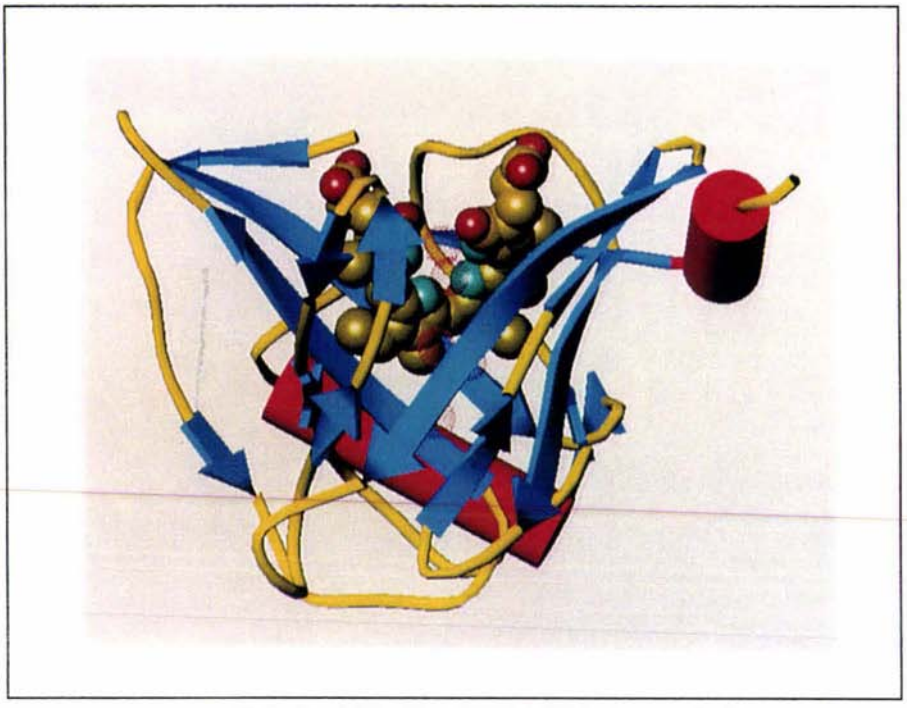


Figure 654-1: BBP-bilin complex

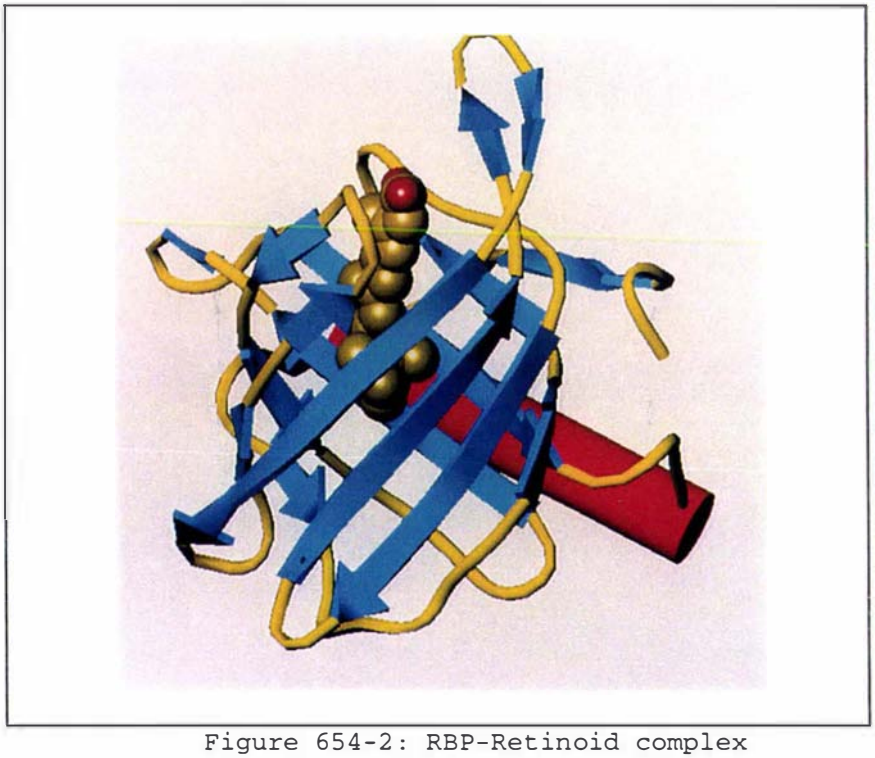


Figure 654-2: RBP-Retinoid complex

the porin plane of the bilin ligand; π-electron cloud stacking may be involved. The binding site of BBP is much more spacious that those of other lipocalins accommodating the bulky bilin molecule, as well as several water molecules. Whereas the ligand-binding site for BBP is somewhat polar inside the calyx, OBP, RBP and BLG provide the ligands with an absolutely hydrophobic environment.

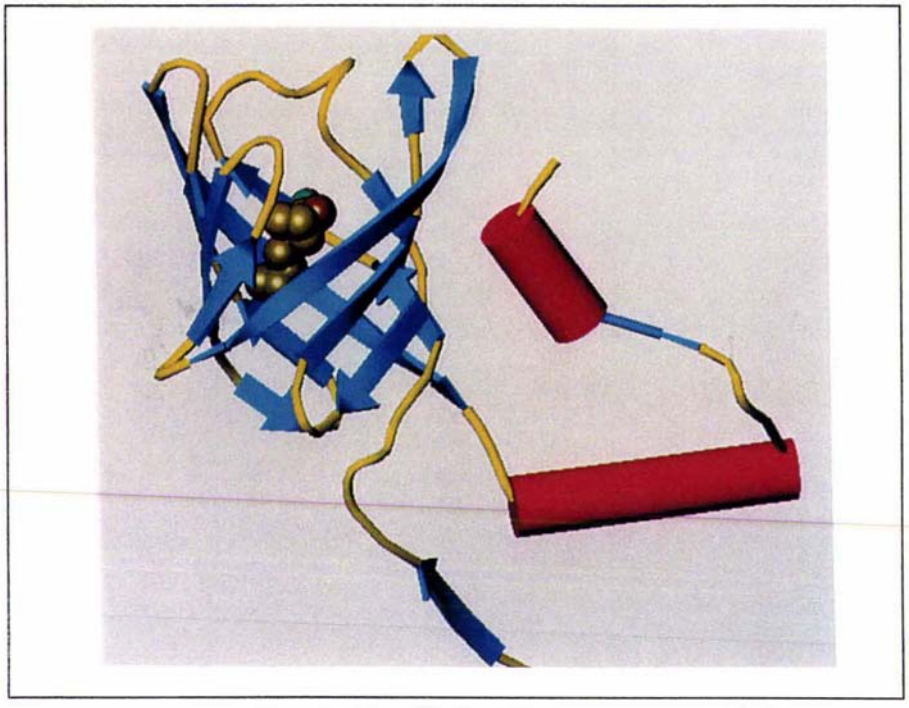


Figure 654-3: OBP-pyrazine complex

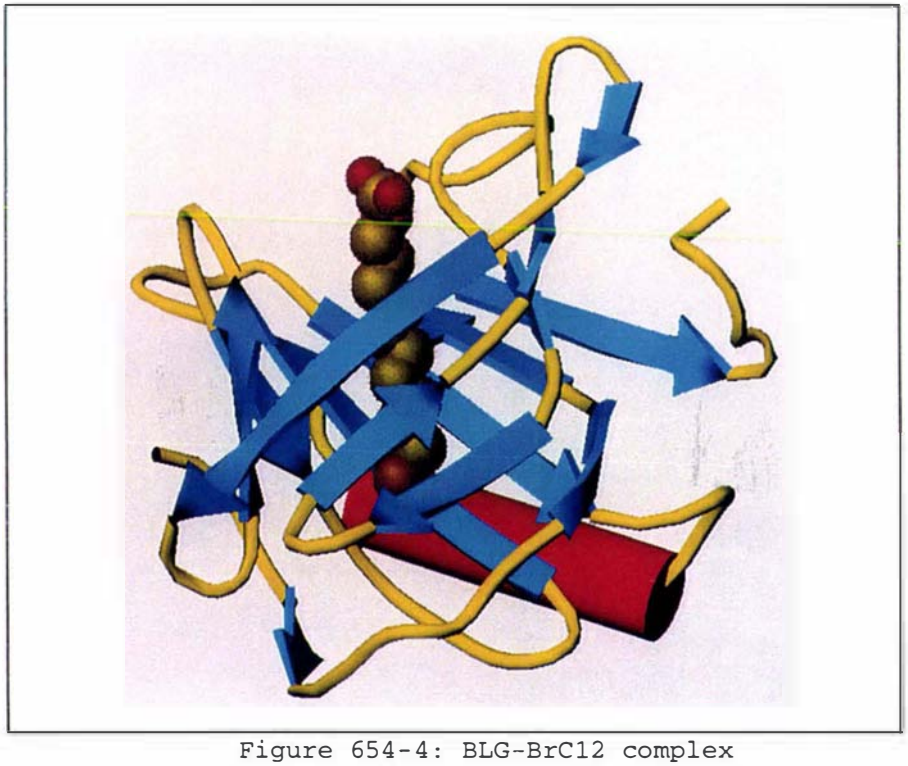


Figure 654-4: BLG-BrC12 complex

6.5.5 Bottom region and handle region

As analysed in section 6.5.2, many of the residues involved in the bottom region of BLG do not exist in the OBP sequence. This causes the OBP molecule to have a more open style of ß-barrel; see Figure 655-1. While the bottom complex is almost intact in RBP, the corresponding structure in BBP is twisted. In lattice Z of BLGA, the ß-strand I is followed by the short helix α-3 near the C-terminus. No corresponding helix has been found in retinol-binding protein.

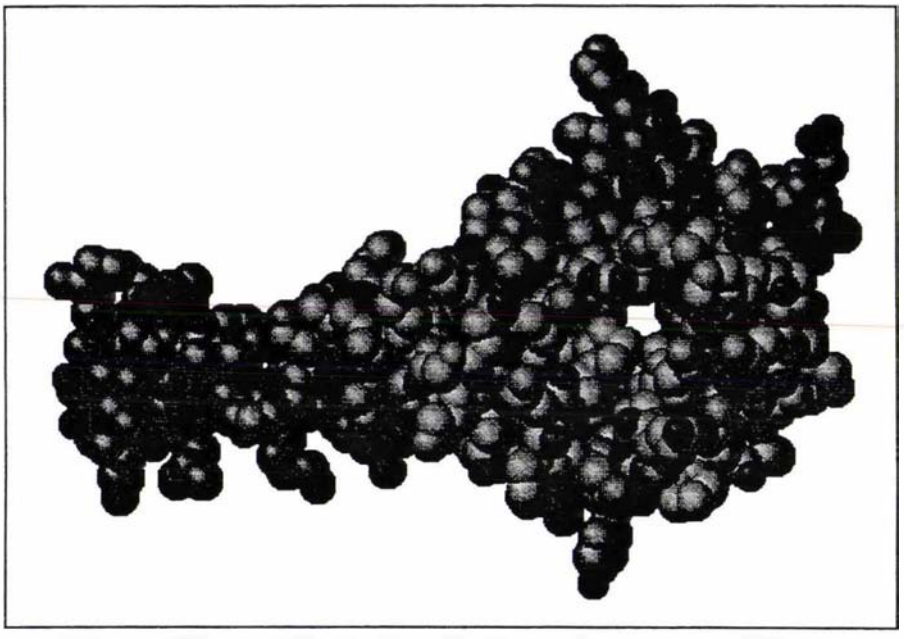


Figure 655-1: CPK model of the ß-barrel of OBP

6.5.6 Disulfide bonds

RBP contains three disulfide bonds: Cys4-Cys160, Cys70-Cys174 and Cys120-Cys129. The last two correspond to the disulfide bonds Cys160-Cys66 and Cys119-Cys106 of BLG, the third disulfide bond of RBP anchors the N-terminus of this molecule. BBP contains two disulfide bonds: Cys43-Cys175 and Cys9-Cys119. The first disulfide bond in BBP corresponds to Cys66-Cys160 of BLG, and the second disulfide bond is similar to that Cys4-Cys120 of RBP. OBP has no disulfide bond, or even any cysteine residues.

Chapter 7 Conclusions and Future Prospects

7.1 Summary of current research on bovine BLG

Bovine ß-lactoglobulin was isolated first in 1934 by Palmer from skim milk. In biological properties, this protein belongs to the lipocalin superfamily. Bovine ß-lactoglobulin shares a similar tertiary structure with other lipocalins, despite a low sequence homology that would suggest a low structural similarity between each other. On the other hand, the lipocalin fold of bovine ß-lactoglobulin has distinctive characteristics compared to that of other lipocalins. The ß-barrel of bovine ß-lactoglobulin is assembled from two cross-oriented ß-sheets with only a single pair of hydrogen bonds between strand D and strand E. In addition BLG has longer and more flexible loops than other lipocalins. These result in more flexible binding properties for bovine ß-lactoglobulin than other lipocalins.

The consequences of point mutation depend on their environment (rigid or mobile) plus the mutation characteristics themselves (e.g. change of the side chain's charge and bulk). Comparing bovine BLGA with BLGB, the mutation D64G, which occurs in a mobile region, leads to substantial conformational change around this mutation site; the mutation V118A, which occurs in a rigid region, leads to no substantial conformational change around the mutation site. However, loop EF in BLGA and BLGB at pH 7.1 has different conformations which result in different molecular volumes for bovine BLGA and BLGB.

Bovine ß-lactoglobulin can be crystallized in several crystal forms, called lattice types. Investigation of three lattice types (X, Y and Z) of BLG suggests that the BLG dimer is the basic functional unit of the BLG molecule. In lattice Z, the BLG dimers form zig-zag chains through "lock and key" interfaces. In lattices X and Y, the BLG dimers form chains through "bottom:bottom" interfaces. The detailed structural comparison of the BLG molecule in lattices X, Y and Z reveals overall similarity of BLG in lattices X and Z, while the handle region of BLG in lattice Y is twisted compared to that of BLG in lattices X and Z.

Two ligand-binding sites are revealed in the structures: the major and

the minor sites. The major site is inside the BLG calyx, which has been observed to host 12-bromododecanoic acid (BrC12) or waters. The binding of ligand BrC12 has only a minor structural impact on several regions of BLG. The minor site lies between the ß-sheet II and the BLG handle, at Cys121. The binding of CH_3Hg^{3+} has minimal impact on the conformation of the handle region.

Bovine ß-lactoglobulin can be conveniently disassembled into several portions: the top region, which involves loops AB, CD, EF, and GH; the bottom region; the calyx handle region which partially covers the ß-barrel; and the ß-barrel which is formed by ß-sheets I and II. While the major portions of bovine ß-lactoglobulin remain unchanged with change in pH, loop EF, which contains the Glu89, experiences a critical conformational change. This transition causes the side chain of Glu89, which is buried at pH 6.2, to become exposed at pH 7.1 and pH 8.2. This conformational change provides a structural basis for a variety of pH-dependent phenomena which are collectively known as the Tanford transition.

7.2 Future research

To date bovine ß-lactoglobulin provides the only structural data on the ß-lactoglobulin family. Compared with the available BLG sequences (currently seventeen), this is only the beginning of the structure determinations for ß-lactoglobulin. One interesting point about the ß-lactoglobulin structure is the signal peptide, which is removed from bovine ß-lactoglobulin in the mature form of the protein, but remains in porcine ß-lactoglobulin. Does the signal peptide play a role in the potential physiological function of ß-lactoglobulin? Does it relate to the fact that fatty acids can not be bound to porcine ß-lactoglobulin, but to bovine ß-lactoglobulin?

The existence of BLG in kangaroo suggests that this secreted protein has a long evolutionary history. The role of BLG in animals which diverged from ruminants ~50 million years ago may provide some insight into the physiological function of BLG. The simple argument is that the physiological function of BLG may have been lost during the evolutionary history of mammals (since human and rodent lack BLG), but should have

some impact on its early appearance. Future research about the BLG structures of long-diverged species may facilitate the elucidation of the physiological function of BLG.

Even though the role of Glu89 in the Tanford transition has been proposed based on the structure analyses of this thesis, gene manipulation about Glu89 is anticipated to provide more insight into the role of Glu89 in the Tanford transition. Understanding the Tanford transition is important not only for the physiological function of BLG, but also for potential pharmaceutical applications. Bovine ß-lactoglobulin may act as a shuttle to carry safely acid-sensitive medicine through stomach. As the residues inside the BLG calyx are not strongly conserved, suitable gene manipulation to these residues may customize the binding ability of BLG required in future applications.

References

Ali, S. and Clark, A.J. (1988). Characterization of the gene encoding ovine ß-lactoglobulin. *J. Mol. Biol.*, 1988;199:415-426.

Ali, S., Mcclenaghan, M., Simons, J.P. and Clark, A.J. (1990). Characterisation of the alleles encoding ovine beta-lactoglobulins A and B. *Gene*, 1990;91:201-207.

Alexander, L.J., Levine, W.B., Teng, C.T. and Beattie, C. W. (1992). Cloning and sequencing of the porcine lactoferin cDNA. *Anim. Genet*, 1992;23:251-256.

Alexander, L.J. and Beattie, C.W. (1992). Sequence of porcine ß-lactoglobulin cDNA. *Anim. Genet.*, 1992;23:263-265.

Anema, S.G. and Creamer, L.K. (1993). Effect of the A and B variants of both alpha s1- and kappa-casein on bovine casein micelle solvation and kappa-casein content. *J. Dairy Res.*, 1993;60:505-516.

Askonas, B. A. (1954). Crystallization of goat ß-lactoglobulin. *Biochem. J.*, 1954;58:332-337.

Aschaffenburg, R. (1968). Genetic variants of milk protein: their breed distribution. *J.Dairy Res.*, 1968;35:447-460.

Aschaffenburg, R. and Drewry, J. (1955). Occurrence of different ß-lactoglobulins in cow's milk. *Nature*, 1955;176:218-219.

Aschaffenburg, R. and Drewry, J. (1957). Improved method for the preparation of crystalline ß-lactoglobulin and α-lactoalbumin from cow's milk. *Biochem. J.*, 1957;65:273-277.

Aschaffenburg, R., Green, D.W. and Simmons, R.M. (1965). Crystal forms of ß-lactoglobulin. *J. Mol. Biol.*, 1965;13:194-201.

Axelsson, I., Jakobsson, I., Lindberg, T. and Beneditsson, B. (1986). Bovine ß-lactoglobulin in the human milk. *Acta Paed. Scand.*, 1986;75:702-

707.

Baker, E.N. and Hubbard, R. E. (1984). Hydrogen bonding in globular proteins. *Prog. Biophys. Molec. Biol.*, 1984;44:97-179.

Barlow, W. (1897). A mechanical cause of homogeneity of structure and symmetry. *Sci. Pro. Roy. Dublin Soc.*, 1897;8:527-690.

Basch, J.J. and Timasheff, S.N. (1967). Hydrogen ion equilibria of the genetic variants of bovine ß-lactoglobulin. *Arch. Biochem. Biophys.*, 1967;118:37-47.

Bell, K. (1962). One dimensional starch gel electrophoresis of bovine skim milk. *Nature*, 1962;195:705-706.

Bell, K. and McKenzie, H.A. (1964). ß-Lactoglobulins. *Nature*, 1964;204:1275-1278.

Bell, K. and McKenzie, H.A. (1967). The whey proteins of ovine milk: ß-lactoglobulins A and B. *Biochim. Biophys. Acta*, 1967;147:109-122.

Bell, K., McKenzie, H.A. and Murphy, W.H. (1970). ß-Lactoglobulin (Drought master). A unique protein variant. *Biochim. Biophys. Acta*, 1970;214:427-436.

Bell, K., McKenzie, H.A. and Shaw, D.C. (1981). Bovine ß-lactoglobulin E, F and G of Bali (Banteng) Cattle, *Bos (Bibos) javanicus*. *Aust. J. Biol. Sci.*, 1981;34:133-147.

Bernback, S., Blackberg, L. and Hernell, O. (1990). The complete digestion of human milk triacylglycerol *in vitro* requires gastric lipase, pancreatic colipase-dependent lipase, and bile salt-stimulated lipase. *J. Clin. Invest.*, 1990;85:1221-1226.

Berni, R., Stoppini, M., Zapponi, M.C., Meloni, M.L., Monaco, H.L. and Zanotti, G. (1990). The bovine plasma retinol-binding protein, amino acid sequence, interaction with transthyretin, crystallization and preliminary X-ray data. *Eur. J. Biochem.*, 1990;192:507-513.

Bewley, M.C., Qin, B.Y., Jameson, G.B., Sawyer, L. and Baker, E.N. (1998). Bovine ß-lactoglobulin and its variants: a three dimensional structure perspective. In *intl. Dairy Fed. Bull.*, 1998; Special Issue. "Milk Protein Polymorphism," IDF, 101-109.

Bianchet, M.A., Bains, G., Pelosi, P., Pevsner, J., Snyder, S.H., Monaco, H.L. and Amzel, L.M. (1996). The three-dimensional structure of bovine odorant binding protein and its mechanism of odor recognition. *Nat. Struct. Biol.*, 1996;3:934-939.

Bignetti, E., Tirindelli, R., Rossi, G.L., Bolognesi, M., Coda, A. and Gatti, G. (1985). Crystallization of an odorant-binding protein from cow nasal mucosa. *J. Mol. Biol.*, 1985;186:211-212.

Blomhoff, R., Green, M.H., Berg, T. and Norum, K.R. (1990). Transport and storage of vitamin A. *Science*, 1990;250:399-403.

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

Boyer, P.D. (1954). Spectrophotometric study of the reaction of protein sulfhydryl groups with organic mercurials. *J. Amer. Chem. Soc.*, 1954;76:4331-4337.

Brand, E., Saidel, L.J., Goldwater, W.H., Kassell, B. and Ryan, F.J. (1945). Empirical formula for ß-lactoglobulin. *J. Amer. Chem. Soc.*, 1945;67:1524-1532.

Braunitzer, G., Chen, R., Schrank, B. and Stangl, A. (1973). The sequence of ß-lactoglobulin. *Hoppe-Seylers. Z. Physiol. Chem.*, 1973;354:867-878.

Brew, K. and Campbell, P.N. (1967). The characterization of the whey proteins of guinea-pig milk. *Biochem. J.*, 1967;102:258-264.

Brignon, G. and Ribadeau-Dumas, B. (1969). Chemical and physico-chemical

characterization of genetic variant D of bovine ß-lactoglobulin D. *Arch. Biochem. Biophys.*, 1969;129:720-727.

Brignon, G. and Ribadeau-Dumas, B. (1973). Localisation dans la chaine peptidique de la ß-lactoglobulin bovine de la substitution Glu/Gln differenciant les varients genetiques B et D. *FEBS Lett.*, 1973;33:73-76.

Brown, E.M. and Farrell, H.M. Jr. (1978). Interaction of ß-lactoglobulin and cytochrome c: complex formation and iron reduction. *Arch. Biochem. Biophys.*, 1978;185:156-164.

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

Brünger, A.T. (1988-1992). V3.1f. Yale University.

Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, 1992;355:472-475.

Bull, H.B. and Currie, B.T. (1946a). Monolayers of ß-lactoglobulin I film molecular weight. *J. Amer. Chem. Soc.*, 1946;68:745-747.

Bull, H.B. and Currie, B.T. (1946b). Monolayers of ß-lactoglobulin II Mixed monolayers of ß-lactoglobulin and sodium lauryl sulphate. *J. Amer. Chem. Soc.*, 1946;68:747-748.

Cambillau, C., Roussel, A., Inisan, A.G. and Knoops-Mouthuy, E. (1996). TURBO-FRODO version 5.5 Bio-Graphics.

Cannan, R.K., Palmer, A.H. and Kibrick, A.C. (1941). The hydrogen ion dissociation curve of ß-lactoglobulin. *J. Biol. Chem.*, 1941;142:803-833.

Casal, H.L., Kohler, U. and Mantsch, H.H. (1988). Structural and conformational change of ß-lactoglobulin B: an infrared spectroscopic study of the effect of pH and temperature. *Biochim. Biophys. Acta*, 1988;957:11-20.

Cecil, R. and Ogston, A.G. (1949). The sedimentation constant, diffusion constant and molecular weight of ß-lactoglobulin. *Biochem. J.*, 1949;44:33-35.

Chibnall, A.C. (1946). The contribution of the analytical chemist to the problem of protein structure (second proctor memorial lecture). *J. Intern. Soc. Leather Trades Chem.*, 1946;30:1-20.

Cho, Y.J., Batt, C.A. and Sawyer, L. (1994). Probing the retinol-binding site of bovine ß-lactoglobulin. *J. Biol. Chem.*, 1994;269:11102-11107.

Cho, Y., Gu, W., Watkins, S., Lee, S.P., Kim, T.R., Brady, J.W. and Batt, C.A. (1994). Thermostable variants of bovine ß-lactoglobulin. *Protein Eng.*, 1994;7:263-270.

Chobert, J.M., Dalgalarrondo, M., Dufour, E., Bertrand-Harb, C. and Haertle, T. (1991). Influence of pH on the structural changes of ß-lactoglobulin studied by tryptic hydrolysis. *Biochim. Biophys. Acta*, 1991;1077:31-34.

Clark, A.J., Simons, P., Wilmut, I. and Lathe, R. (1987). Pharmaceuticals from transgenic livestock. *TIBTech.*, 1987;5:20-24.

Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst.*, 1994;D50:760-763.

Collet, C., Joseph, R. and Nicholas, K. (1991). A marsupial ß-lactoglobulin gene: characterization and prolactin-dependent expression. *J. Mol. Endocrinol.*, 1991;6:9-16.

Collet, T.C. and Joseph, R. (1995). Exon organization and sequence of the genes encoding α-lactalbumin and ß-lactoglobulin from the tammar wallaby (*Macropodidae, Marsupialia*). *Biochem. Genet.*, 1995;33:61-72.

Conti, A., Napolitano, L., Cantisani, A.M., Davoli, R. and Dall'Olio, S. (1988). Bovine ß-lactoglobulin H: isolation by preparative isoelectric focusing in immobilized pH gradients and preliminary characterization.

J. Biochem. Biophys. Meth., 1988;16:205-214.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucl. Acids. Res.*, 1988;16:10881-10890.

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

Creamer, L.K. and MacGibbon, A.K.H. (1996). Some recent advances in the basic chemistry of milk proteins and lipids. *Int. Dairy Journal*, 1996;6:539-568.

Creamer, L.K. (1998). personal communication.

Crowfoot, D. and Riley, D. (1938). Crystal structures of the proteins. *Nature*, 1938;141:521-524.

Díaz de Villegas, M.C., Oria, R., Sala, F.J. and Calvo, M. (1987). Lipid binding by ß-lactoglobulin of cow milk. *Milchwissenschaft*, 1987;42:357-361.

Drayna, D. (1986). Cloning and expression of human apolipoprotein D cDNA. *J. Biol. Chem.*, 1986;261:16535-16539.

Drenth, J. (1995). Principles of protein X-ray crystallography. 2nd ed. Springer, New York. p74-181.

Dufour, E. and Haertle, T. (1990). Alcohol-induced change of ß-lactoglobulin —— retinol-binding stoichiometry. *Protein Engineering*, 1990;4:185-190.

Dunnill, P. and Green, D.W. (1965). Sulphydryl groups and the N ⇔ R conformational change in ß-lactoglobulin. *J. Mol. Biol.*, 1965;15:147-151.

Erhardt, G., Godovac-Zimmermann, J. and Conti, A. (1989). Isolation and complete primary sequence of a new ovine wild-type ß-lactoglobulin C.

Biol. Chem. Hoppe-Seyler, 1989;370:757-762.

Ewald, P.P. (1921). Das 'reziproke gitter' in der strukturtheorie. Teil I: Das Reziproke eines einfachen gitters. [The 'reciprocal lattice' in the theory of structure. Part I. The reciprocal of a primitive lattice.] *Z. Krist.*, 1921;56:129-156.

Farrell, H.M., Bede, M.J. and Enyeart, J.A. (1987). Binding of *p*-nitrophenyl phosphate and other aromatics by ß-lactoglobulin. *J. Dairy Sci.*, 1987;70:252-258.

Fernandez, F.M. and Oliver, G. (1988). Proteins present in llama milk. I. quantitative aspects and general characteristics. *Milchwissenschaft*, 1988;43:299-302.

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

Flower, D.R. (1995). Multiple molecular recognition properties of the lipocalin protein family. *J. Mol. Recognit.*, 1995;8:185-195.

Flower, D.R., North, A.C.T. and Attwood, T.K. (1991). Mouse oncogene protein 24p3 is a member of the lipocalin protein family. *Biochem. Biophys. Res. Commun.*, 1991;180:69-74.

Flower, D.R., North, A.C.T. and Attwood, T.K. (1993). Structure and sequence relationships in the lipocalins and related proteins. *Protein Science*, 1993;2:753-761.

Frapin, D., Dufour, E. and Haertlé, T. (1993). Probing the fatty acid binding site of ß-lactoglobulins. *J. Protein Chem.*, 1993;12:443-449.

Fugate, R.D. and Song, P.S., (1980). Spectroscopic characterization of ß-lactoglobulin-retinol complex. *Biochim. Biophys. Acta*, 1980;625:28-42.

Futterman, S. and Heller, J. (1972). The enhancement of fluorescence and the decreased susceptibility to enzymatic oxidation of retinol complexed with bovine serum albumin, ß-lactoglobulin, and the retinol-binding

protein of human plasma. *J. Biol. Chem.*, 1972;247:5168-5172.

Ghose, A.C., Chaudhuri, S. and Sen, A. (1968). Hydrogen ion equilibria and sedimentation behaviour of goat ß-lactoglobulins. *Arch. Biochem. Biophys.*, 1968;126:232-243.

Godovac-Zimmermann, J., Conti, A., James, L. and Napolitano, L. (1988). Micro-analysis of the amino-acid sequence of monomeric ß-lactoglobulin I from donkey (*Equus asinus*) milk. *Biol. Chem. Hoppe-Seyler*, 1988; 369:171-179.

Godovac-Zimmermann, J., Conti, A., Liberatori, J. and Braunitzer, G. (1985). The amino-acid sequence of ß-lactoglobulin II from horse colostrum (*Equus caballus, Perissodactyla*): ß-lactoglobulins are retinol-binding proteins. *Biol. Chem. Hoppe-Seyler*, 1985;366:601-608.

Godovac-Zimmermann, J. and Shaw, D. (1987). ß-Lactoglobulin identified in marsupial milk. the primary structure, binding site and possible function of ß-lactoglobulin from eastern grey kangaroo (*macropus giganteus*). *Biol. Chem. Hoppe-Seyler*, 1987;368:879-886.

Godovac-Zimmermann, J., Conti, A. and Napolitano, L. (1987). The complete amino-acid sequence of dimeric ß-lactoglobulin from mouflon (*Ovis ammon musimon*) milk. *Biol. Chem. Hoppe-Seyler*, 1987;368:1313-1319.

Godovac-Zimmermann, J, Conti, A., Sheil, M. and Napolitano, L. (1990). Covalent structure of the minor monomeric ß-lactoglobulin II component from donkey milk. *Biol. Chem. Hoppe-Seyler*, 1990;371:871-879.

Godovac-Zimmermann, J., Krause, I., Buchberger, J., Weiss, G. and Klostermeyer, H. (1990). Genetic variants of bovine beta-lactoglobulin. A novel wild-type ß-lactoglobulin W and its primary sequence. *Biol. Chem. Hoppe-Seyler*, 1990;371:255-260.

Goodman, D.S., Sporn, M.B., Roberts, A.R. and Goodman, D.S. (1984). In the retinoids. Vol 2 Eds., Academic Press, Orlando, FL, pp. 41-88

Gorbunoff, M.J. (1967). Exposure of tyrosine residues in protein:

Reaction of cyanuric fluoride with ribonuclease, α-lactalbumin, and ß-lactoglobulin. *Biochemistry*, 1967;6:1606-1615.

Green, D.W. and Aschaffenburg, R. (1959). Two fold symmetry of the ß-lactoglobulin molecule in crystals. *J. Mol. Biol.*, 1959;1:54-64.

Green, D.W., Aschaffenburg, R., Camerman, A., Coppola, J.C., Dunnill, P., Simmons, R.M., Komorowski, E.S., Sawyer, L., Turne, E.M.C. and Woods, K.F. (1979). Structure of bovine ß-lactoglobulin at 6 Å resolution. *J. Mol. Biol.*, 1979;131:375-397.

Grosclaude, F. (1966). Mise en evidence de deux variants supplementaires des proteines du lait de vache: αs1-CnD et LgD. *Ann. Biol. Anim. Biochem. Biophys. Acta*, 1966;295:352-363.

Groves, M.L., Hipp, N.J. and McMeekin, T.L. (1951). Effect of pH on the denaturation of ß-lactoglobulin and its dodecyl sulfate derivative. *J. Amer. Chem. Soc.*, 1951;73:2790-2793.

Hahn, T. (1995). 4th ed. International tables for crystallography, Vol A. Kluwer Academic Publishers, Boston.

Harker, D. (1956). The determination of the phases of the structure factors of non-centrosymmetric crystals by the method of double isomorphous replacement. *Acta Cryst.*, 1956;9:1-9.

Halliday, J.A., Bell, K., McKenzie, H.A. and Shaw, D.C. (1990). Feline whey proteins: identification, isolation and initial characterization of α-lactalbumin, ß-lactoglobulin and lysozyme. *Comp. Biochem. Physiol.*, 1990;B95:773-779.

Halliday, J.A., Bell, K., McAndrew, K. and Shaw, D.C. (1993). Feline ß-lactoglobulin I II and III, and canine ß-lactoglobulins I and II: Amino acid sequences provide evidence for the existence of more than one gene for ß-lactoglobulin in the cat and dog. *Protein Seq. Data Anal.*, 1993;5:201-205.

Halliday, J.A., Bell, K. and Shaw, D.C. (1991). The complete amino acid

sequence of feline ß-lactoglobulin II and a partial revision of the equine ß-lactoglobulin II sequence. *Biochim. Biophys. Acta*, 1991;1077:25-30.

Hambling, S.G., McAlpine, A.S. and Sawyer, L. (1992). In "Advanced Dairy Chemistry: 1.Proteins" (Fox,P.F.,ed.) pp 141-190, Elsevier Appl. Sci.

Haque, Z., Kristjannson, M.M. and Kinsella, J.E. (1987). Interaction between κ-casein and ß-lactoglobulin: possible mechanism. *J. Agric. Food Chem.*, 1987;35:644-649.

Hattori, M., Ametani, A., Katakura, Y., Shimizu. M. and Kaminogawa, S. (1993). Unfolding/refolding studies on bovine ß-lactoglobulin with monoclonal antibodies as probes. *J. Biol. Chem.*, 1993;268:22414-22419.

Hemley, R., Kohler, B.E. and Siviski, P. (1979). Absorption spectra for the complexes formed from vitamin-A and ß-lactoglobulin. *Biophys. J.*, 1979;28:447-455.

Hennighausen, L.G. and Sippel, A.E. (1982). Mouse whey acidic protein is a novel member of the family of 'four-disulphide core' proteins. *Nucl. Acid Res.*, 1982;10:2677-2684.

Hill, J.P., Motion, R.L., Buckley, P.D. and Blackwell, L.F. (1994). The effect of *p*-(chloromercuri) benzoate modification of cytosolic aldehyde dehydrogenase from sheep liver. Evidence for a second aldehyde binding site. *Arch. Biochem. Biophys.*, 1994;310:256-263.

Hill, J.P., Boland, M.J., Creamer, L.K., Anema, S.G., Otter, D.E., Paterson, G.R., Lowe, R., Motion, R.L. and Thresher, W.C. (1996). Effect of the bovine ß-lactoglobulin phenotype on the properties of ß-lactoglobulin, Milk Composition, and Dairy Products. Chapter 22, Macromolecular Interactions in Food Technology, published by the American Chemical Society.

Huang, X.L., Catignani, G.L., Foegeding, E.A. and Swaisgood, H.E. (1994). Relative structural stabilities of ß-lactoglobulins A and B as determined by proteolytic susceptibility and differential scanning calorimetry. *J.*

Agric. Food Chem., 1994;42:1064-1067.

Huber, R., Schneider, M., Mayr, I., Muller, R., Deutzmann, R., Suter, F., Zuber, H., Falk, H. and Kayser, H. (1987). Molecular structure of the bilin binding protein (BBP) from *Pieris brassicae* after refinement at 2.0 Å resolution. *J. Mol. Biol.*, 1987;198:499-513.

Huber, R., Schneider, M., Epp, O., Mayr, I., Messerschmidt, A., Pflugrath, J. and Kayser, H. (1987). Crystallization, crystal structure analysis and preliminary molecular model of the bilin binding protein from the insect *Pieris brassicae*. *J. Mol. Biol.*, 1987;195:423-434.

Hunter, A.K. (1995). Submitted to EMBL/GENBANK/DDBJ DATA BANKS.

Hunziker, H.G. and Tarassuk, N.P. (1965). Chromatographic evidence for heat induced interaction of α-lactalbumin and ß-lactoglobulin. *J. Dairy Sci.*, 1965;48:733-744.

Jakob, E. and Puhan, Z. (1992). Technological properties of milks as influenced by genetic polymorphism of milk proteins —— a review. *Int. Dairy J.*, 1992;2:157-178.

Jakob, E., Sievert, C., Sommer, S. and Puhan, Z. (1995). Automated determination of total nitrogen in milk by the Dumas method. *Z. Lebensm Unters Forsch*, 1995;200:239-243.

Jakob, E. (1994). Genetic polymorphism of milk proteins. *Bull. Int. Dairy Fed.*, 1994;298:17-27.

Jakobsson, I., Lindberg, T., Benediktsson, B. and Hansson, B.G. (1985). Dietary bovine ß-lactoglobulin is transferred to human milk. *Acta Paediatr Scand.*, 1985;74:342-345.

Jakobsson, I. (1985). Unusual presentation of adverse reactions to cow's milk proteins. *Klin. Padiatr*, 1985;197:360-362.

Jang, H.D. and Swaisgood, H.E. (1990). Characteristics of the interaction of calcium with casein submicelles as determined by analytical affinity

chromatography. *Arch. Biochem. Biophys.*, 1990;283:318-325.

Jameson, G.B. (1998). unpublished results.

Jeffrey, P.A. (1997). PREDICT. shareware.

Jenness, R. (1985). Biochemical and nutritional aspects of milk and colostrum. In Lactation (Larson, B.L. ed.). Iowa State University Press, pp. 164-197.

Jenness, R., Phillip, N.I. and Kalan, E.B., (1967). Immunochemical comparison of ß-lactoglobulins. *Fed. Proc.*, 1967;26:340-343.

Jones, S. and Thorton, J.M. (1995). Protein-protein interactions: a review of protein dimer structures. *Prog. Biophys. Molec. Biol.*, 1995;63:31-65.

Jones, T. A. and Kjeldgaard, M. (1993). v5.9, Department of Molecular Biology, BMC, Uppsala University, Sweden.

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

Kaminsky, S., Smith, L.J. and D'Souza, S.W. (1988). Human gastric lipase. Effects of fatty acid and bovine serum albumin on *in vitro* activity. *Scand. J. Clin. Lab. Invest.*, 1988;48:583-587.

Katakura, Y., Totsuka, M., Ametani, A. and Kaminogawa, S. (1994). Tryptophan-19 of ß-lactoglobulin, the only residue completely conserved in the lipocalin superfamily, is not essential for binding retinol, but relevant to stabilizing bound retinol and maintaining its structure. *Biochim. et Biophys. Acta*, 1994;1027:58-67.

Kella, N.K. and Kinsella, J.E. (1988). Enhanced thermodynamic stability of ß-lactoglobulin at low pH. *Biochem. J.*, 1988;255:113-118.

Kim, S.C., Olson, N.F. and Richardson, T. (1990). Polymerization and

gelation of thiolated ß-lactoglobulin at ambient temperature induced by oxidation with potassium iodate. *Milchwissenschaft.*, 1990;45:627-631.

Kolde, H.-J., Liberatori, J. and Braunitzer, G. (1981). The amino acid sequence of the water buffalo ß-lactoglobulin. *Milchwissenschaft*, 1981;36:83-86.

Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.*, 1991;24:946-950.

Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structure. *J. Appl. Cryst.*, 1993;26:283-289

Lee, S-P., Cho, S. and Batt, C.A. (1993). Enhancing the gelation of ß-lactoglobulin. *J. Agric. Food. Chem.*, 1993;41:1343-1348.

Lee, K.H. (1987). Isolation of an olfactory cDNA: similarity to retinol-binding protein suggests a role in olfaction. *Science*, 1987;235:1053-1056.

Li, C.H. (1946). Electrophoretic inhomogeneity of crystalline ß-lactoglobulin. *Biochem. J.*, 1946;68:2746-2747.

Liberatori, J., Guetti, L.M. and Conti, A. (1979). Immunological evidence of ß-lactoglobulins in human colostrum and milk. *Boll. Soc. It. Biol. Sper.*, 1979;55:822-825.

Liberatori, J., Morisio, L., Guetti, L.M., Conti, A. and Napolitano, L. (1979). ß-Lactoglobulins in the mammary secretions of camel (*Camelus dromedarius*) and she-ass. Immunological detection and preliminary physico-chemical characterization. *Boll. Soc. It. Biol. Sper.*, 1979;55:1369-1373.

Liberatori, J., Guidetti, L.M. and Conti, A. (1979). Immuno-chemical studies on ß-lactoglobulins precipitin reactions of sow's and mare's mammary secretions against anti-bovine ß-lactoglobulin antiserum. *Boll. Soc. It. Biol. Sper.*, 1979;55:815-821.

Lontie, R. and Preaux, G. (1966). Polarimetric investigation of ß-lactoglobulin A and B and of the reactivity of their thiol groups. *Protides of the Biological Fluids*, 1966;14:476-481.

Lovrien, R., and Anderson, W.F. (1969). Resolution of binding sites in ß-lactoglobulin. *Arch. Biochem. Biophys.*, 1969;131:139-144.

Lyster, R.L.J., Jenness, R. and Phillips, N.I. (1966). Comparative biochemical studies of milks ---- III. Immuno-electrophoretic comparisons of milk proteins of the *Artiodactyla*. *Comp. Biochem. Physiol.*, 1966;17:967-971.

Mailliart, P. and Ribadeau-Dumas, B. (1988). Preparation of ß-lactoglobulin and ß-lactoglobulin-free proteins from whey retentate by NaCl salting out at low pH. *J. Food Sci.*, 1988;53:743-752.

Manderson, G., Hardman, M.J. and Creamer, L.K. (1995). Thermal denaturation of bovine ß-lactoglobulin A, B, and C. *J. Dairy Sci.*, Suppl. issue, 1995;78:132-132.

Matthews, B.W. (1968). Solvent content of protein crystals. *J. Mol. Biol.*, 1968;33:491-497.

Maurer, H.R. (1966). Simple decolorizing apparatus for disc electrophoresis. *Z. Klin. Chem.*, 1966;4:85-86.

Maurer, H.R. and Dati, F.A. (1972). Polyacrylamide gel electrophoresis on micro slabs. *Anal. Biochem.*, 1972;46:19-32.

McAlpine, A.S. and Sawyer, L. (1990). ß-Lactoglobulin: a protein drug carrier? *Biochem. Soc. Trans.*, 1990;197:407-417.

McDonald, I., Naylor, D., Jones, D. and Thornton, J. (1993). HBPLUS: hydrogen bond calculator v2.25.

McDougall, E.I. and Stewart, J.C. (1976). The whey proteins of the milk of red deer (*cervus elaphus L.*). *Biochem. J.*, 1976;153:647-655.

McKenzie, H.A. Ed. (1971). In Milk Proteins Chemistry and Molecular Biology; Academic Press: New York. p274-331.

McKenzie, H.A., Müller, V.J. and Shaw, D.C. (1983). 'Whey' proteins of milk of the red (*Macropus rufus*) and Eastern grey (*Macropus giganteus*) kangaroo. *Comp. Biochem. Physiol.*, 1983;B74:259-271.

McKenzie, H.A. and Sawyer, W.H. (1967). Effect of pH on ß-lactoglobulins. *Nature*, 1967;214:1101-1104.

McKenzie, H.A., Sawyer, W.H. and Smith, M.B. (1967). Optical rotatory dispersion and sedimentation in the study of association-dissociation: bovine ß-lactoglobulins near pH 5. *Biochim. Biophys. Acta*, 1967;147:73-92.

McKenzie, H.A. and Sawyer, W.H. (1972). On the dissociation of bovine ß-lactoglobulins A, B, and C near pH 7. *Aust. J. Biol. Sci.*, 1972;25:949-961.

Mclean, D.M., Graham, E.R.B., Ponzoni, R.W. and McKenzie, H.A. (1987). Effects of milk protein genetic variarts and composition on heat stability of milk. *J. Dairy Res.*, 1987;54:219-235.

McMeekin, T.L., Polis, B.D., DellaMonica, E.S. and Custer, J.H. (1949). A crystalline compound of ß-lactoglobulin with dodecyl sulfate. *J. Amer. Chem. Soc.*, 1949;71:3606-3609.

McSwinery, M., Singh, H., Campanella, O. and Creamer, L.K. (1994). Thermal gelation and denaturation of bovine ß-lactoglobulins A and B. *J. Dairy Res.*, 1994;61:221-232.

Merchant, Z., Jiang, L.X., Lebenthal, E. and Lee, P.C. (1987). Pancreatic exocrine enzymes during the neonatal period in postmature rats. *Int. J. Pancreatol*, 1987;2:325-335.

Mills, O.E. (1976). Effect of temperature on tryptophan fluorescence of ß-lactoglobulin B. *Biochim. Biophys. Acta*, 1976;434:324-332.

Miranda, G. and Pelissier, J.P. (1983). Kinetic studies of *in vivo* digestion of bovine unheated skim-milk proteins in rat stomach. *J. Dairy Res.*, 1983;50:27-36.

Monaco, H.L. and Zanotti, G. (1992). Three-dimensional structure and active site of three hydrophobic molecule-binding proteins with significant amino acid sequence similarity. *Biopolymers*, 1992;32:457-465.

Monaco, H.L., Zanotti, G., Spadon, P., Bolognesi, M., Sawyer, L. and Eliopoulos, E.E. (1987). Crystal structure of the trigonal form of bovine ß-lactoglobulin and of its complex with retinol at 2.5 Å resolution. *J. Mol. Biol.*, 1987;197:695-706

Monti, J.C. and Jolles, P. (1982). Characterization of a thermosensitive protein from human milk whey. *Experientia*, 1982;38:1211-1213.

Monti, J.C., Mermoud, A.F. and Jolles, P. (1989). Anti-bovine ß-lactoglobulin antibodies milk. *Experientia*, 1989;45:178-180.

Moreau, H., Gargouri, Y., Lecat, D., Junien, J.L. and Verger, R. (1988). Screening of preduodenal lipases in several mammals. *Biochim. Biophys. Acta*, 1988;959:247-252.

Mulvihill, D.M. and Kinsella, J.E., (1987). Characteristics of whey proteins and ß-lactoglobulin. *Food Technol.*, 1987;419:102-111.

Narayan, M. and Berliner, L.J. (1997). Fatty acids and retinoids bind independently and simultaneously to ßlactoglobulin. *Biochemistry*, 1997;36:1906-1911.

Narayan, M. and Berliner, L.J. (1998). Mapping fatty acid binding to ß-lactoglobulin: Ligand binding is restricted by modification of Cys121. *Protein Science*, 1998;7:150-157.

Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Cryst.*, 1994;A50:157-163.

Neuteboom, B., Giuffrida, M.G. and Conti, A. (1992). Isolation of a new ligand-carrying casein fragment from bovine mammary gland microsomes. *Fed. Eur. Biol. Soc. Lett.*, 1992;3:189-193.

Newcomer, M.E. (1993). Structure of the epididymal retinoic acid binding protein at 2.1 Å resolution. *Structure*, 1993;1:7-18.

Nicholls, A., Sharp, K. and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins*, 1991;11:281-296.

North, A.C.T. (1991). Protein structure, prediction and design. (Kay, J., Lunt, G., and Osguthorpe, D. Ed.) Portland Press, 1990, London. *Biochem. Soc. Symp.*, 1991;57:35-48.

Nozaki, Y., Bunville, L.G. and Tanford, C. (1959). Hydrogen ion titration curves of ß-lactoglobulin. *J. Amer. Chem. Soc.*, 1959;81:5523-5529.

Oatley, S.J. (1994). TOM v2.4 UCSD

O'Neill, T.E. and Kinsella, J.E. (1987). Binding of alkanone flavors to ß-lactoglobulin: effects of conformational and chemical modification. *J. Agric. Food Chem.*, 1987;35:770-774.

Otwinowski, Z. (1996-97). Data processing system Version 1.9.1. University of Virginia Patent Foundation.

Palmer, A.H. (1938). The preparation of a crystalline globulin from the albumin fraction of cow's milk. *J. Biol. Chem.*, 1938;104:359-372.

Pantaloni, D. (1965). Structure et changements de conformations de la ß-lactoglobulin en solution. Doctural Thesis, University of Paris.

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

Parry-Smith, D.J. and Attwood, T.K., (1992). ADSP —— a new package for computational sequence analysis. *Comput. Appl. Biosci.*, 1992;8:451-459.

Patterson, A.L. (1934). A Fourier series method for the determination of the components of interatomic distances in crystals. *Phys. Rev.*, 1936;46: 372-376.

Patterson, A.L. (1935). A direct method for the determination of the components of interatomic distances in crystals. *Z. Krist.*, 1935;A90:517-542.

Pedersen, K.O. (1936). Introduction and preliminary results with fractionation of skim milk. *Biochem. J.*, 1936;30:948-960.

Pedersen, K.O. (1936). The lactoglobulin of Palmer: Detected changes in sedimentation coefficient at pH 5.0 and 7.5. *Biochem. J.*, 1936;30:961-970.

Pelosi, P. (1982). Identification of s specific olfactory receptor for 2-isobutyl-3-methoxypyrazine. *Biochem. J.*, 1982;201:245-248.

Pelosi, P. (1989). In Chemical Senses, Vol 1. (Brand, J.G., Teeter, J.H., Cagan, R.H. and Kare, M.R., Eds.) Marcel Dekker, New York.

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

Pérez, M.D. and Calvo, M. (1995). Interaction of ß-lactoglobulin with retinol and fatty acids and its role as possible biological function for this protein: a review. *J. Dairy Sci.*, 1995;78:979-988.

Pérez, M.D., Sanchez, L., Aranda, P.P., Ena, J.M. and Calvo, M. (1990). Synthesis and evolution of concentration of ß-lactoglobulin and α-lactoalbumin from cow and sheep colostrum and milk throughout early lactation. *Cell Mol. Biol.*, 1990;36:205-212.

Pérez, M.D., Sanchez, L., Aranda, P., Ena, J.M., Oria, R. and Calvo, M.

(1992). Effect of ß-lactoglobulin on the activity of pregastric lipase —— a possible role for this protein in ruminant milk. *Biochim. Biophys. Acta*, 1992;1123:151-155.

Pérez, M.D., Diaz de Villegas, C., Sanchez, L., Aranda, P., Ena, J.M. and Calvo, M. (1989). Interaction of fatty acids with ß-lactoglobulin and albumin from ruminant milk. *J. Biochem. (Tokyo)*, 1989;106:1094-1097.

Pervaiz, S. and Brew, K. (1986). Purification and characterization of the major whey proteins from the milks of the bottlenose dolphin (*Tursiops truncatus*), the Florida manatee (*Tricheclus manatus latirostris*) and the beagle (*Canis familiaris*). *Arch. Biochem. Biophys.*, 1986;246:846-854.

Pervaiz, S. and Brew, K. (1985). Homology of ß-lactoglobulin, serum retinol-binding protein, and protein HC. *Science*, 1985;228:335-337.

Pervaiz, S. and Brew, K. (1987). Homology and structure-function correlations between α1-acid glycoprotein and serum retinol-binding protein and its relatives. *FASEB. J.*, 1987;1:209-214.

Pevsner, J. (1990). Odorant-binding protein (characterization of ligand binding). *J. Biol. Chem.*, 1990;265:6118-6125.

Pevsner, J. (1990). Odorant-binding protein (characterization of ligand binding). *J. Biol. Chem.*, 1990;265:6118-25

Pevsner, J. and Reed, R.R. (1988). Molecular cloning of odorant-binding protein member of a ligand carrier family. *Science*, 1988;241:336-339.

Pevsner, J., Trifiletti, R.R., Strittmatter, S.M. and Snyder, S.H. (1985). Isolation and characterization of an olfactory receptor protein for odorant pyrazines. *Proc. Natl. Acad. Sci. USA*, 1985;82:3050-3054.

Polis, B.D., Schmukler, H.W., Custer, J.H. and McMeekin, T.L. (1950). Isolation of an electrophoretically homogeneous crystalline component of ß-lactoglobulin. *J. Amer. Chem. Soc.*, 1950;72:4965-4968.

Preaux, G., Braunitzer, G., Schrank, B. and Stangl, A. (1979). The amino acid sequence of goat ß-lactoglobulin. *Hoppe-Seylers Z. Physiol. Chem.*, 1979;360:1595-1604.

Puyol, P., Pérez, M.D., Ena, J.M. and Calvo, M. (1991). Interaction of bovine ß-lactoglobulin and other bovine and human whey proteins with retinol and fatty acids. *Agric. Biol. Chem.*, 1991;55:2515-2520.

Puyol, P., Pérez, MD., Mata. L. and Calvo, M. (1994). Study on interaction between ß-lactoglobulin and other bovine whey proteins with ascorbic acid. *Milchwissenschaft*, 1994;49:25-31

Puyol, P., Pérez, MD., Mata. L., Ena, J.M. and Calvo, M. (1993). Effect of retinol and fatty acid binding to bovine ß-lactoglobulin on its resistance to trypsin digestion. *Int. Dairy J.*, 1993;3:589-595.

Puyol, P., Pérez, M.D., Peiro, J.M. and Calvo, M. (1994). Effect of retinol and fatty acid binding to bovine ß-lactoglobulin on its resistance to thermal denaturation. *J. Dairy Sci.*, 1993;77:1494-1503.

Qi, X.L., Brownlow, S., Holt, C. and Sellers, P. (1995). Thermal denaturation of ß-lactoglobulin: effect of protein concentration at pH 6.75 and 8.05. *Biochim. Biophys. Acta*, 1995;1248:43-49.

Qin, B.Y. (1996). The structure determination of ß-lactoglobulin from bovine milk whey. First year report, 1996 June. Masey University, Palmerstom North, New Zealand.

Qin, B.Y., Creamer, L.K., Bewley, M.C., Baker, E.N. and Jameson, G.B. (1998). Structural basis of the Tanford transition of bovine ß-lactoglobulin A from crystal structures at three pH values. *Biochemistry*, accepted.

Ralston, G.B. (1972). The decrease in stability of ß-lactoglobulin on blocking the sulphydryl group. *Compt. Rend. Trav. Lab. Carlsberg.*, 1972;38(25):499-512.

Rask, L., Anundi, H. and Peterson, P.A. (1979). The primary structure of

the human retinol-binding protein. *FEBS Lett.*, 1979;104:55-58.

Ravelli, R.B.G. (1997). Strategy: A program for the determination of the optimal starting spindle angle in X-ray crystallography data-collection. Department of Crystal and Structural Chemistry, Utrecht University, The Netherlands.

Reddy, M., Kella, N.K. and Kinsella, J.E. (1988). Structural and conformational basis of the resistance of ß-lactoglobulin to peptic and chymotryptic digestion. *J. Agric. Food Chem.*, 1988;36:737-742.

Robillard, K.A. and Wishnia, A. (1972). Aromatic hydrophobes and ß-lactoglobulin A: thermodynamics of binding. *Biochemistry*, 1972;11:3835-3840.

Robillard, K.A. and Wishnia, A. (1972). Aromatic hydrophobes and ß-lactoglobulin A: kinetics of binding by nuclear magnetic resonance. *Biochemistry*, 1972;11:3841-3845.

Sawyer, L. (1987). One fold among many. *Nature*, 1987;327:659-659.

Scader, C. and Schneider, R. (1991). Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins*, 1991;9:56-68.

Schubert, D. and LaCorbiere, M. (1985). Isolation of an adhesion-mediating protein from chick neural retina adherons. *J. Cell Biol.*, 1985;101:1071-1077.

Schubert, D., LaCorbiere, M. and Esch, F. (1986). A chick neural retina adhesion and survival molecule is a retinol-binding protein. *J. Cell Biol.*, 1986;102:2295-2301.

Sheldrick, G.M. (1997). SHELXL: program for crystal structure refinement, Universtiy of Göttingen, Germany.

Sjögren, B. and Svedberg, T. (1930). The molecular weight of lactalbumin. *J. Amer. Chem. Soc.*, 1930;52:3650-3654.

Smith, S. and Abraham, S. (1975). The composition and biosynthesis of milk fat. *Adv. Lipid Res.*, 1975;13:195-239.

Sørensen, C.M. and Palmer, A.H. (1938). Dissociation tendency of crystalline lactoglobulin in very weak solutions of ammonium chloride. *Compt. Rend. Trav. Lab. Carlsberg. Sér. Chim.*, 1938;21:283-290.

Spector, A.A. and Fletcher, J.E. (1970). Binding of long chain fatty acids to ß-lactoglobulin. *Lipids*, 1970;5:403-411.

Spreyer, P., Schaal, H., Kuhn, G., Rothe, T., Unterbeck, A., Olek, K. and Müller, H.W. (1990). Regeneration-associated high level expression of apolipoprotein D mRNA in endoneurial fibroblasts of peripheral nerve. *EMBO. J.*, 1990;9:2479-2484.

Sundelin, J. (1985). Amino acid sequence homologies between rabbit, rat, and human serum retinol-binding proteins. *J. Biol. Chem.*, 1985;260:6472-6480.

Tanford, C., Bunville, L.G. and Nozaki, Y. (1959a). The reversible transformation of ß-lactoglobulin at pH 7.5. *J. Amer. Chem. Soc.*, 1959;81:4032-4035.

Tanford, C. and Nozaki, Y. (1959b). Physico-chemical comparison of ß-lactoglobulins A and B. *J. Biol. Chem.*, 1959;234:2874-2877.

Tegoni, M., Ramoni, R., Bignetti, E., Spinelli, S. and Cambillau, C. (1996). Domain swapping creates a third putative combining site in bovine odorant binding protein dimer. *Nat. Struct. Biol.*, 1996;3:863-867.

Thornton, J.M. (1981). Disulphide bridges in globular proteins. *J. Mol. Biol.*, 1981;151:261-287.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 1994;22:4673-4680.

Thresher, W.C., Knighton, D.R., Otter, D.E. and Hill, J.P. (1994). Brief communications: 24th International dairy congress, Melbourne, 1994, *International Dairy Federation (Brussels)*, 1994; p101.

Timasheff, S.N., Mescanti, L., Basch, J.J. and Townend, R. (1966). Conformational Transitions of bovine ß-lactoglobulins A, B, C. *J. Biol. Chem.*, 1966;241:2496-2501.

Timasheff, S.N. and Townend, R. (1961). Molecular interaction in ß-lactoglobulin. V. The association of the genetic species of ß-lactoglobulin below the isoelectric point. *J. Amer. Chem. Soc.*, 1961;83:464-469.

Tirindelli, R., Keen, J.N., Cavaggioni, A., Eliopoulos, E.E. and Findlay, J.B. (1989). Complete amino acid sequence of pyrazine-binding protein from cow nasal mucosa. *Eur. J. Biochem.*, 1989;185:569-572.

Townend, R., Herskovits, T.T., Timasheff, S.N. and Gorbunoff, M.J. (1969). The state of amino acid residues in ß-lactoglobulin. *Arch. Biochem. Biophys.*, 1969;129:567-580.

Townend, R. and Timasheff, S.N. (1957). The molecule weight of ß-lactoglobulin. *J. Amer. Chem. Soc.*, 1957;79:3613-3614.

Townend, R. and Timasheff, S.N. (1960). Molecule interactions in ß-lactoglobulin. IV. The dissociation of ß-lactoglobulin below pH 3.5. *J. Amer. Chem. Soc.*, 1960;82:3175-3179.

Townend, R., Kumosinski, T.F. and Timasheff, S.N. (1967). The circular dichroism of variants of ß-lactoglobulin. *J. Biol. Chem.*, 1967;242:4538-4545.

Tronrud, D.E. (1995). The TNT refinement package, V5E. copyright 1994, Oregon state board of higher education, US.

Unterman, R.D., Lynch, K.R., Nakhasi, H.L., Dolan, K.P., Hamilton, J.W., Cohn, D.V. and Feigelson, P. (1981). Cloning and sequence of several α2u-

globulin cDNAs. *Proc. Natl. Acad. Sci. USA*, 1981;78:3478-3482.

Weber, K. and Osborn, M. (1969). The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. *J. Biol. Chem.*, 1969;244:4406-4412.

Weichsel, A., Anderson, J.F., Champagne, D.E., Walker, F.A. and Montfort, W.R. (1998). Crystal structures of a nitric oxide transport protein from a blood-sucking insect. *Nat. Struc. Biol.*, 1998;5:304-309.

Wilmot, C.M. and Thornton, J.M. (1990). ß-Turns and their distortions: a proposed new nomenclature. *Protein Engineering*, 1990;3:479-493.

Wishnia, A. and Pinder, T.W. Jr. (1966). Hydrophobic interactions in proteins. The alkane binding site of ß-lactoglobulins A and B. *Biochemistry*, 1966;5:1534-1542.

Yewdall, S.J. (1988). Structural studies on ß-lactoglobulin. Ph.D Thesis, University of Leeds.

Yvon, M., VanHille, I. and Pellisier, J.P. (1984). *In vivo* milk digestion in the calf abomasum. II. Milk and whey proteolysis. *Reprod. Nutr. Dev.*, 1984;24:835-846.

Zanotti, G., Berni, R. and Monaco, H.L. (1993). Crystal structure of liganded and unliganded forms of bovine plasma retinol-binding protein. *J. Biol. Chem.*, 1993;268:10728-10738.

Zanotti, G., Malpeli, G. and Berni, R. (1993). The interaction of *n*-ethyl retinamide with plasma retinol-binding protein (RBP) and the crystal structure of the retinoid-RBP complex at 1.9 Å resolution. *J. Biol. Chem.*, 1993;268:24873-24879.

Appendix

I. Script files

The following X-PLOR script file is designed for scrutinizing the contribution of waters to the R_f and R factor in a structure one by one.

```
remarks  Check  R value for every waters

topology
@/progs1/xplor3.851/toppar/tophcsdx.pro
@bro/brc.top                        {*Read topology file of ligand.*}
@water.top
 end
end
   parameter
@/progs1/xplor3.851/toppar/parhcsdx.pro                              {*Read
empirical potential*}
@bro/brc.par
@water.par
      end
    end                                      {*Append parameters for zinc.*}

{===>} structure @generate.psf end        {*Read structure file.*}

{===>} coor @b_positional_s.pdb end                  {*Read coordinates.*}
 vector do ( charge=0.0 ) ( resname LYS and
   ( name ce or name nz or name hz* ) )        {*Turn off charges on LYS.*}
 vector do ( charge=0.0 ) ( resname GLU and
   ( name cg or name cd or name oe* ) )        {*Turn off charges on GLU.*}
 vector do ( charge=0.0 ) ( resname ASP and
   ( name cb or name cg or name od* ) )        {*Turn off charges on ASP.*}
 vector do ( charge=0.0 ) ( resname ARG and
    ( name cd or name *E or name cz or name NH* or name HH* ) )
                                            {*Turn off charges on ARG.*}

 flags                                    {*In addition to the empirical  *}
   include pele pvdw xref                 {*potential energy terms, which *}
   ?                                      {*are turned on initially,  this*}
 end                                       {*statement turns on the       *}
 xrefine

{===>}
   a= 54.026   b= 54.026   c= 112.183   alpha=90. beta=90.  gamma=120.
{===>}
   symmetry=(x,y,z)   {*Operators for crystal symmetry P3221.*}
   symmetry=(-y,x-y,z+2/3)
   symmetry=(y-x,-x,z+1/3)
   symmetry=(y,x,-z)
   symmetry=(-x,y-x,2/3-z)
   symmetry=(x-y,-y,1/3-z)
       {*The following contains the atomic form factors.  A 4-Gaussian*}
         {*approximation is used.  Atoms are selected based on their    *}
```

```
       {*chemical type.  Note the use of wildcards in the selection.  *}
 SCATter ( chemical C* )
       2.31000 20.8439 1.02000 10.2075 1.58860 .568700 .865000 51.6512
.215600
   SCATter ( chemical N* )
      12.2126 .005700 3.13220 9.89330 2.01250 28.9975 1.16630 .582600 -
11.529
   SCATter ( chemical O* )
       3.04850 13.2771 2.28680 5.70110 1.54630 .323900 .867000 32.9089
.250800
   SCATter ( chemical S* )
       6.90530 1.46790 5.20340 22.2151 1.43790 .253600 1.58630 56.1720
.866900
   SCATter ( chemical Br*  or chemical BR*  )
   17.178900  2.172300  5.235800  16.579599  5.637700  0.260900  3.985100
41.432800 2.955700
 {===>}
  nreflections=10500
  reflection @br1_s.cv end                          {*Read reflections.*}
{===>}
   resolution  15.0  2.23                           {*Resolution range.*}
   reduce
    do amplitude ( fobs = fobs * heavy(fobs - 0.0*sigma))     {*Sigma
cutoff.*}
 fwind=0.01=100000
  print completeness                      {*Check completeness of data.*}
  method=FFT      {*Use the FFT method instead of direct summation.*}
 fft
  memory=1000000  {*This tells the FFT routine how much physical memory*}
  end             {*is available; the number refers to DOUBLE COMPLEX  *}
                  {*words, the memory is allocated from the HEAP.      *}
  ?                                {*This prints the current status.*}
 end                    {*This terminates the diffraction data parser.*}

    xrefin              {*This statement computes the*}
                         {*initial R value.           *}
   do (fpart=0.40*exp(-200*(s()^2)/4)*fpart)

 {* the following table lists all waters name *}
for $1 in ( 400        165 166 167 168 169
                       170 171 172 173 174
                       175 176 177 178 179
 ) loop m1

display $1
    selection=(not resid $1)
    update-fcalc
    print R               {*Please check the R value.*}
    end loop m1
 end

 stop
```

II. List of structure factor and coordinate files

All the experimental structure factor files and the corresponding structure coordinate files are backed up in the attached disc on the back cover of this book. Total 901120 bytes, it can be restored by 'tar -xvf' and followed by 'uncompress' in a UNIX environment.

1. Structure for BLGA in lattice Z at pH 7.1 and the experimental structure factors:

za10_pH7.1_final.pdb.Z (59 block), za10_pH7.1_s.cv.Z (316 block).

2. Structure for BLGA in lattice Z at pH 6.2 and the experimental structure factors:

za9_pH6.2_final.pdb.Z (61 block), za9_pH6.2_s.cv.Z (209 block).

3. Structure for BLGA in lattice Z at pH 8.2 and the experimental structure factors:

za2_pH8.2_final.pdb.Z (60 block), za2_pH8.2_s.cv.Z (249 block).

4. Structure for BLGB in lattice Z at pH 7.1 and the experimental structure factors:

zb1_pH7.1_final.pdb.Z (60 block), zb1_pH7.1_s.cv.Z (322 block).

5. Structure for BLGA-BrC12 in lattice Z at pH 7.3 and the experimental structure factors:

br1_pH7.3_final.pdb.Z (59 block), br1_pH7.3_s.cv.Z (335 block).

Key word index